



US 20170037363A1

(19) **United States**(12) **Patent Application Publication**
Whitlock et al.(10) **Pub. No.: US 2017/0037363 A1**(43) **Pub. Date: Feb. 9, 2017**(54) **AMMONIA-OXIDIZING NITROSOMONAS
EUTROPHA STRAIN D23**(71) Applicant: **AOBIOME LLC**, Cambridge, MA
(US)(72) Inventors: **David R. Whitlock**, Cambridge, MA
(US); **Spiros Jamas**, Cambridge, MA
(US); **Larry Weiss**, San Francisco, CA
(US); **Ioannis Gryllos**, Brookline, MA
(US)(21) Appl. No.: **15/304,151**(22) PCT Filed: **Apr. 15, 2015**(86) PCT No.: **PCT/US2015/025909**

§ 371 (c)(1),

(2) Date: **Oct. 14, 2016****Publication Classification**(51) **Int. Cl.****C12N 1/20** (2006.01)**A61K 35/747** (2006.01)**A61K 35/744** (2006.01)**C12Q 1/02** (2006.01)**A61K 8/99** (2006.01)**A61Q 17/04** (2006.01)**A01N 63/00** (2006.01)**C12R 1/01** (2006.01)**A61K 35/74** (2006.01)**A61K 35/745** (2006.01)(52) **U.S. Cl.**CPC **C12N 1/20** (2013.01); **A61K 35/74**
(2013.01); **A61K 35/747** (2013.01); **A61K**
35/744 (2013.01); **A61K 35/745** (2013.01);
A61K 8/99 (2013.01); **A61Q 17/04** (2013.01);
A01N 63/00 (2013.01); **C12R 1/01** (2013.01);
C12Q 1/02 (2013.01)**Related U.S. Application Data**(60) Provisional application No. 62/053,588, filed on Sep.
22, 2014, provisional application No. 62/012,811,
filed on Jun. 16, 2014, provisional application No.
62/002,084, filed on May 22, 2014.(30) **Foreign Application Priority Data**Apr. 15, 2014 (GR) 20140100217
Mar. 13, 2015 (GR) 20150100115

(57)

ABSTRACT

This disclosure provides, inter alia, an optimized strain of *Nitrosomonas eutropha* (*N. eutropha*) designated D23, D23-100, or AOB D23-100. *N. eutropha* bacteria disclosed in this application have desirable properties, e.g., optimized properties, such as the ability to suppress growth of pathogenic bacteria, and an enhanced ability to produce nitric oxide and nitric oxide precursors. The *N. eutropha* herein may be used, for instance, to treat diseases associated with low nitrite levels, skin diseases, and diseases caused by pathogenic bacteria.

Figure 1

OD 600nm

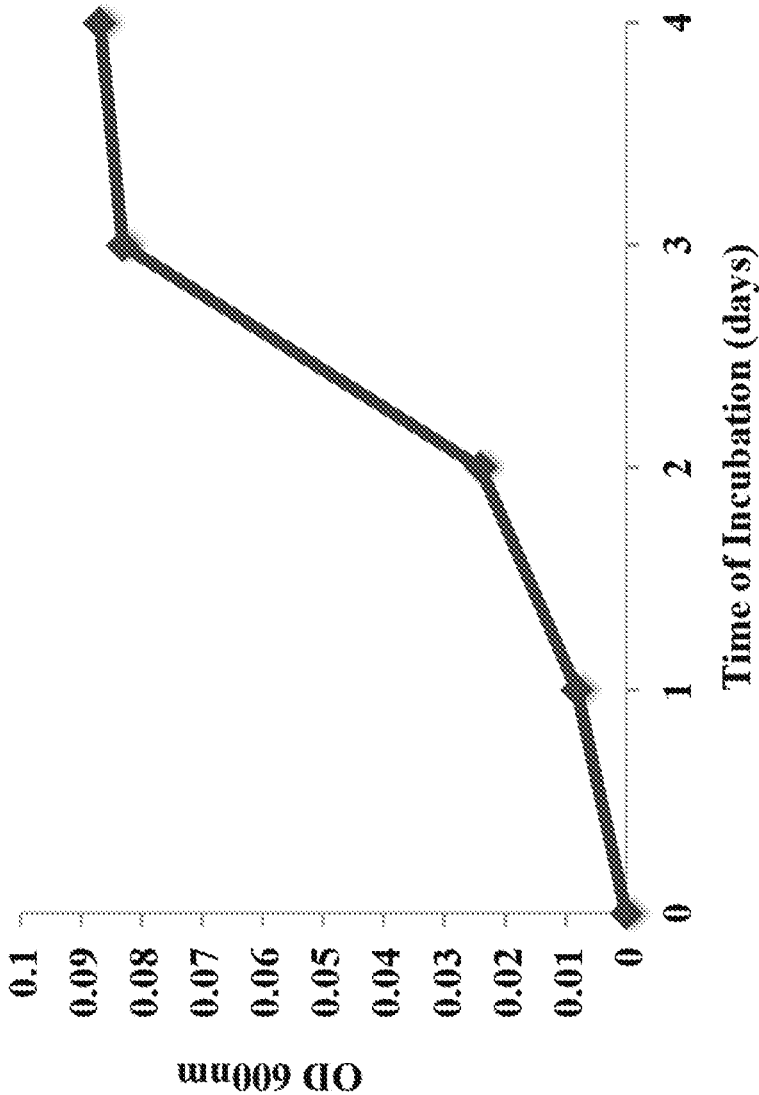


Figure 2A

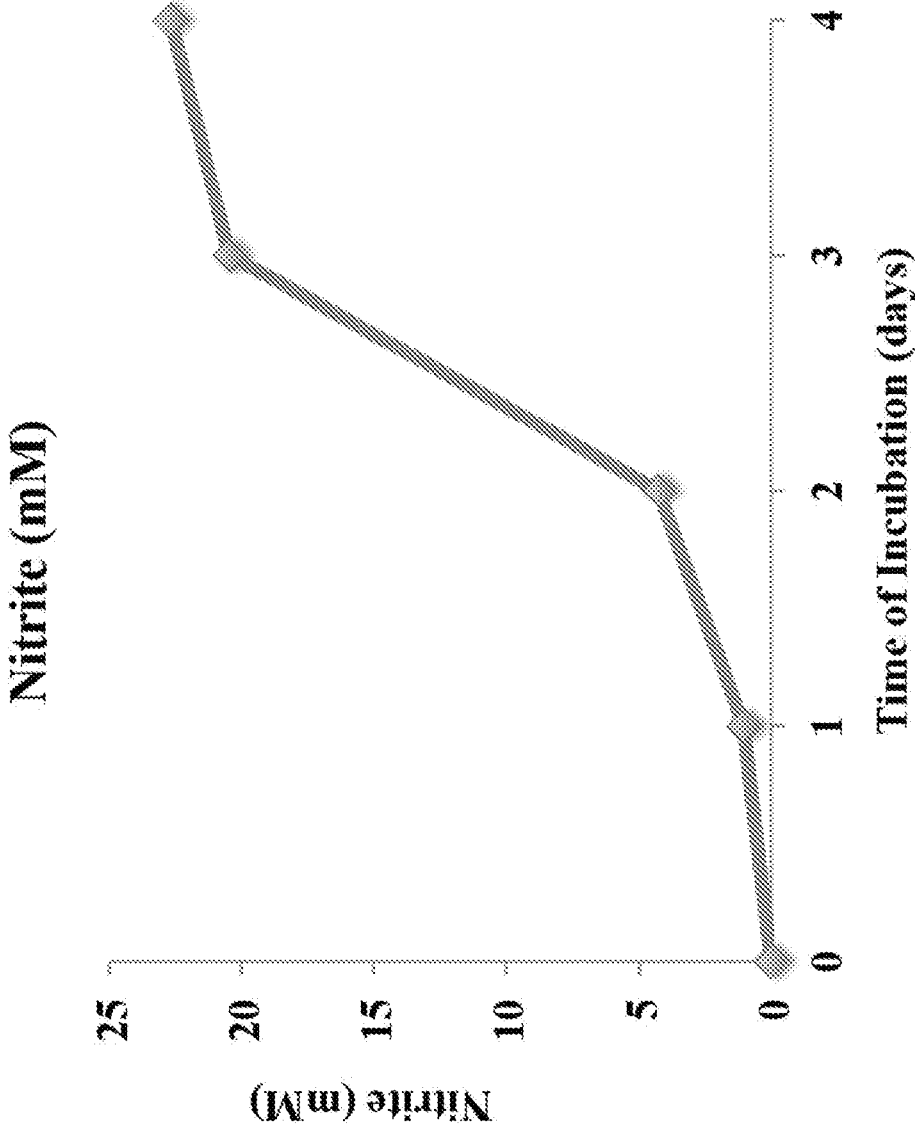


Figure 2B-I

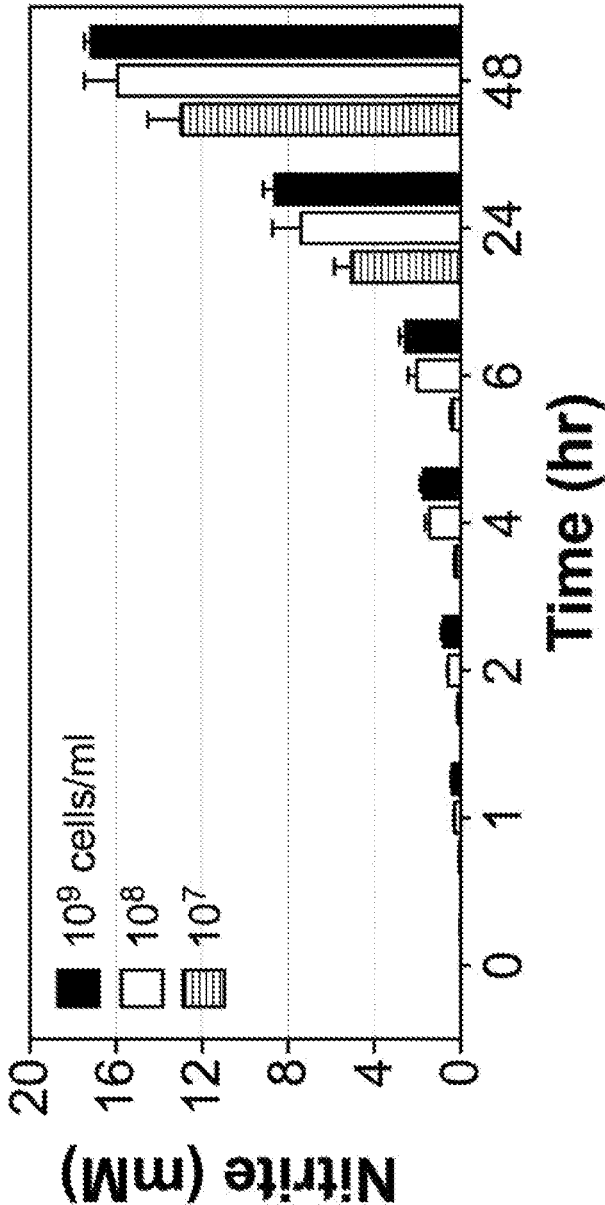


Figure 2B-II

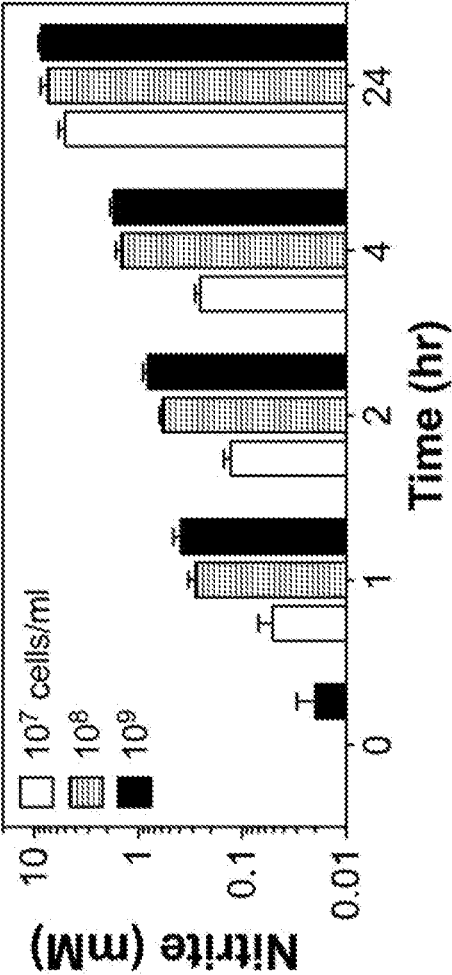


Figure 2C

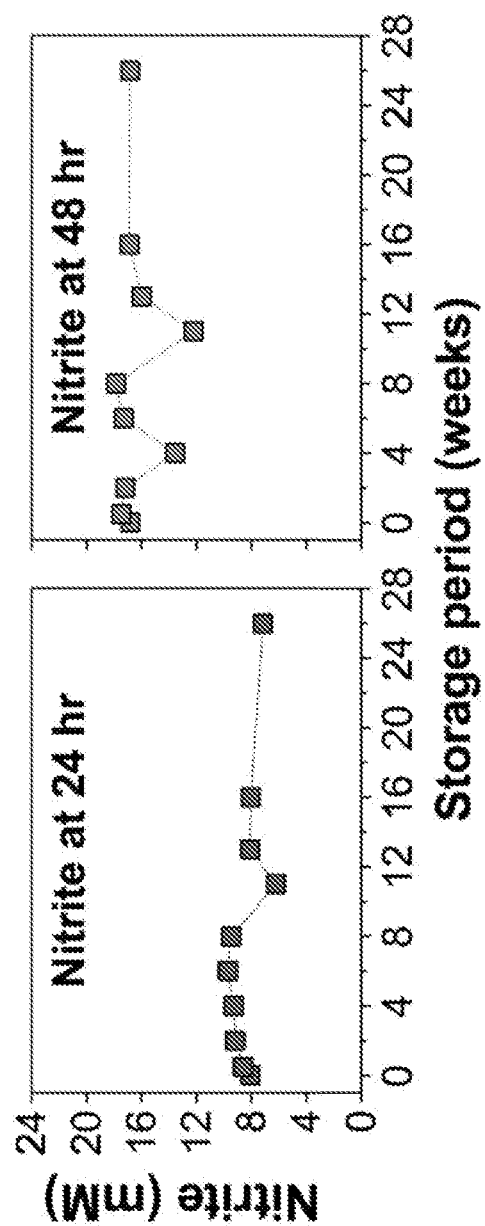


Figure 3A

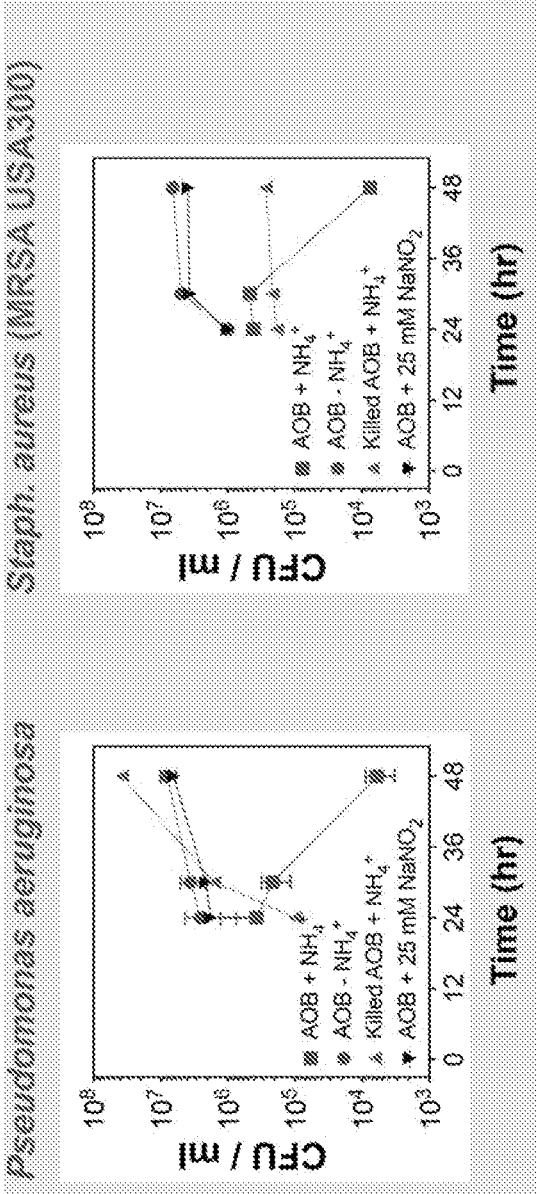
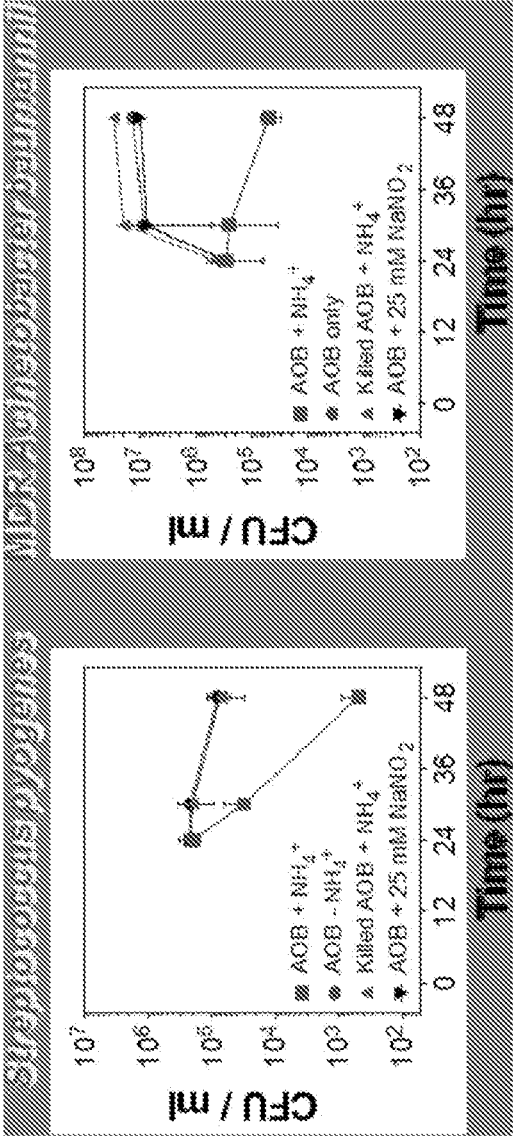


Figure 3B



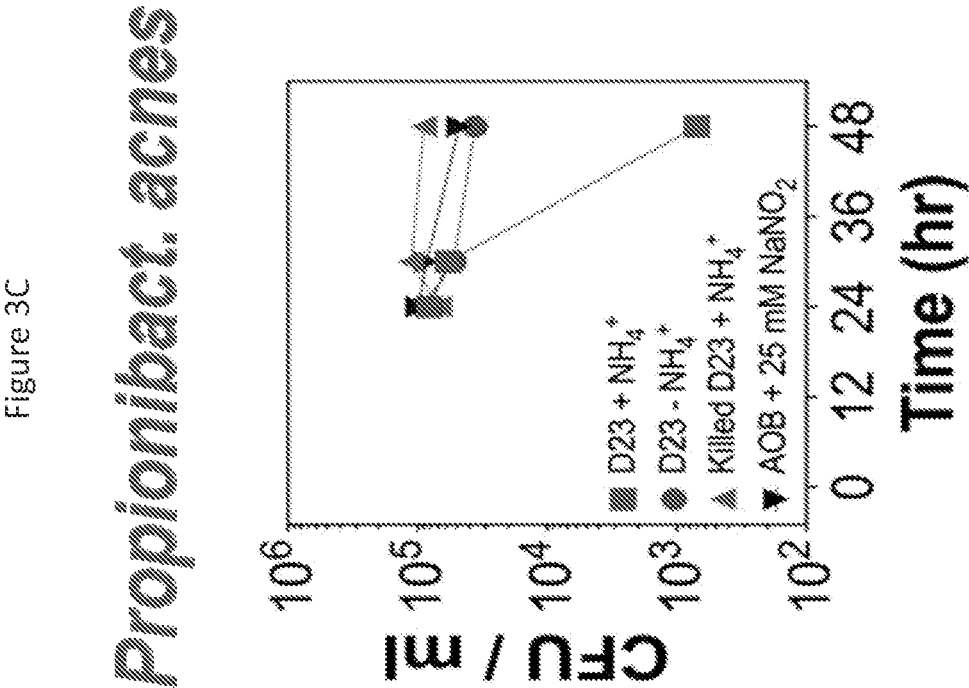


Figure 4A

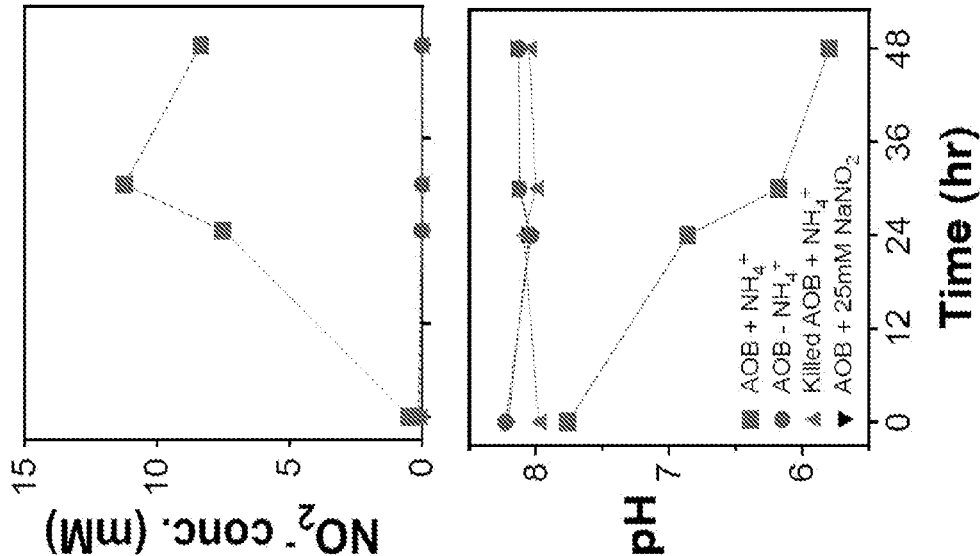
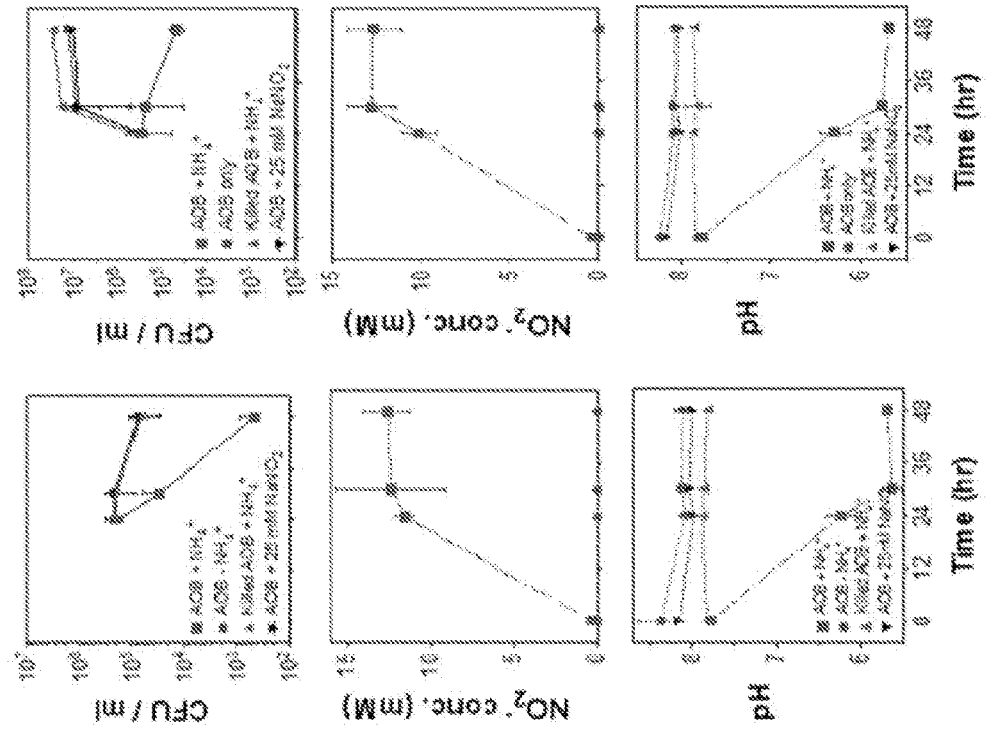


Figure 4B



S. pyogenes

MDR *A. baumannii*

Figure 4C

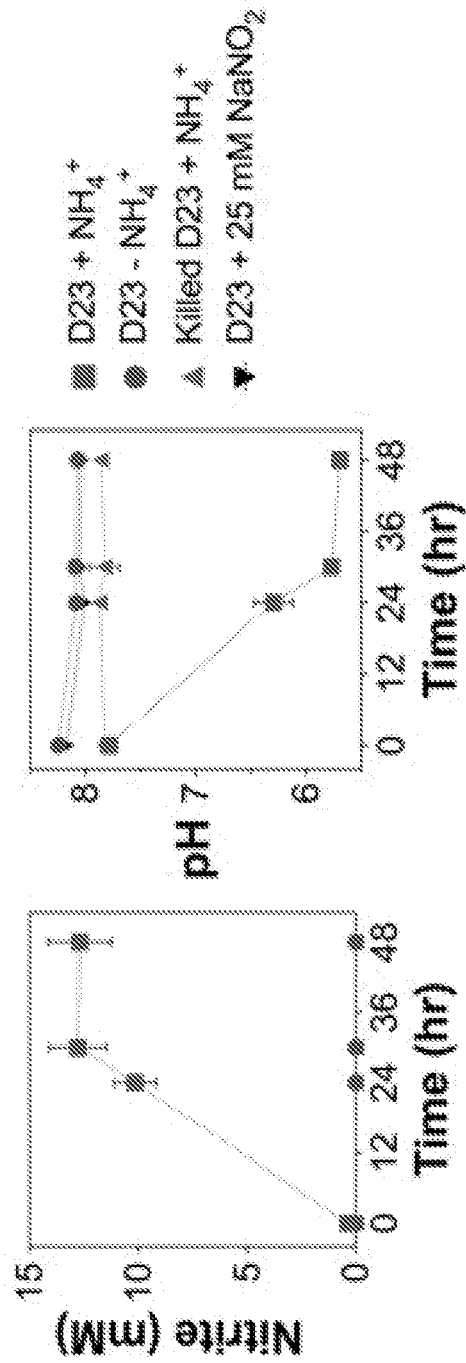


Figure 4D

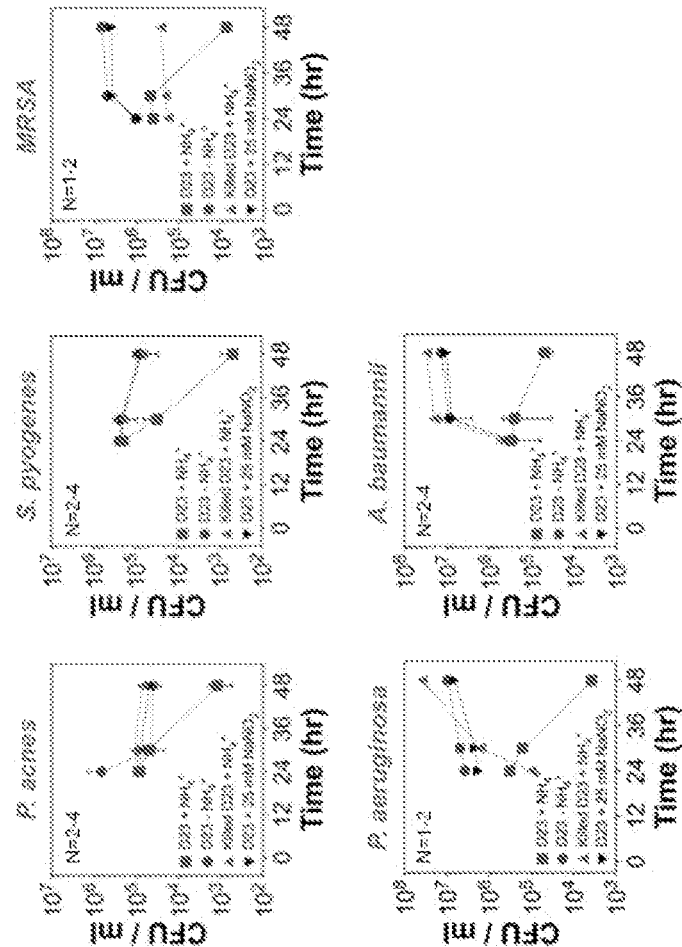


Figure 4E

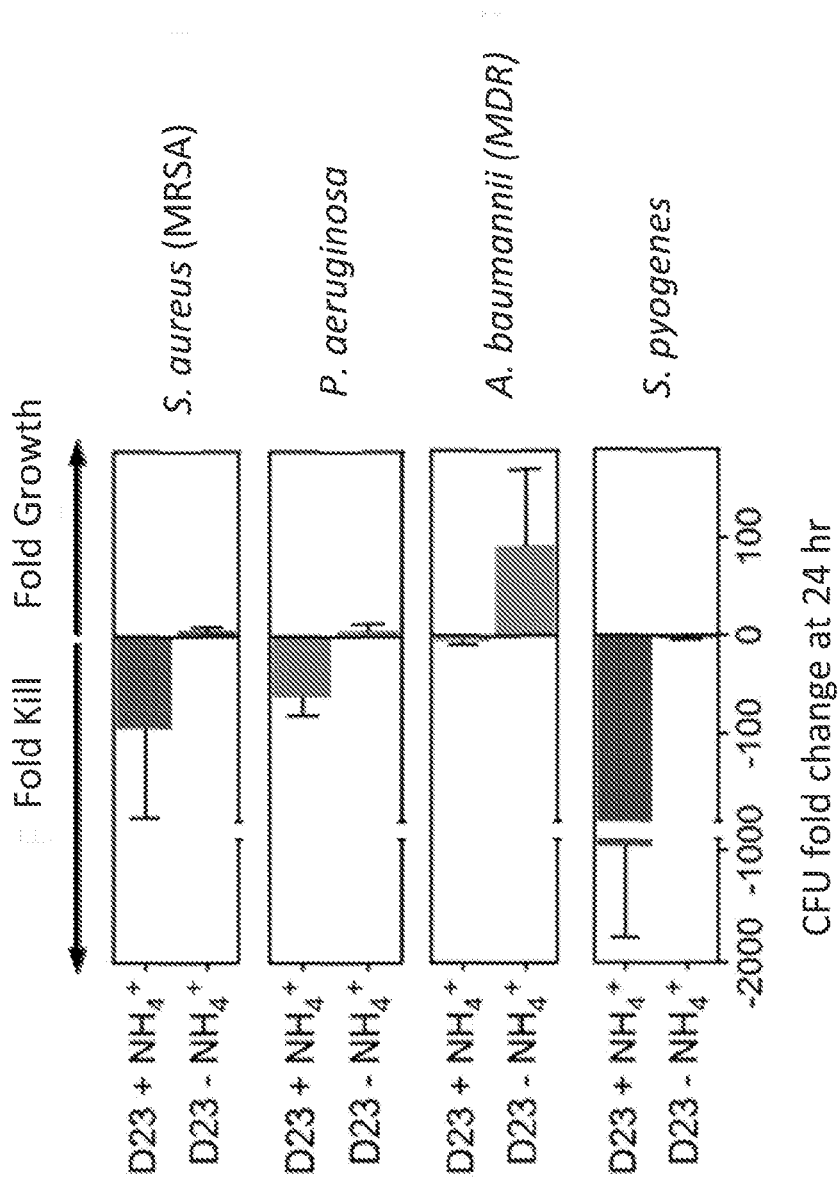
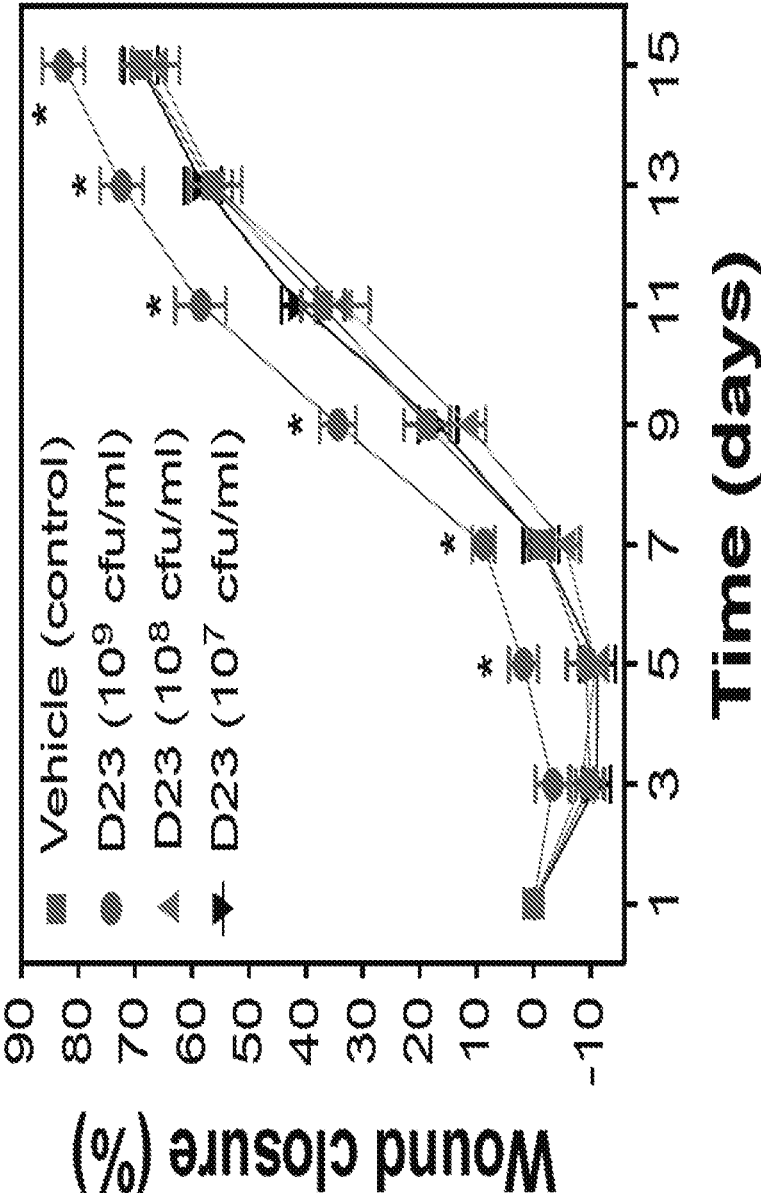


Figure 5A



File 5588

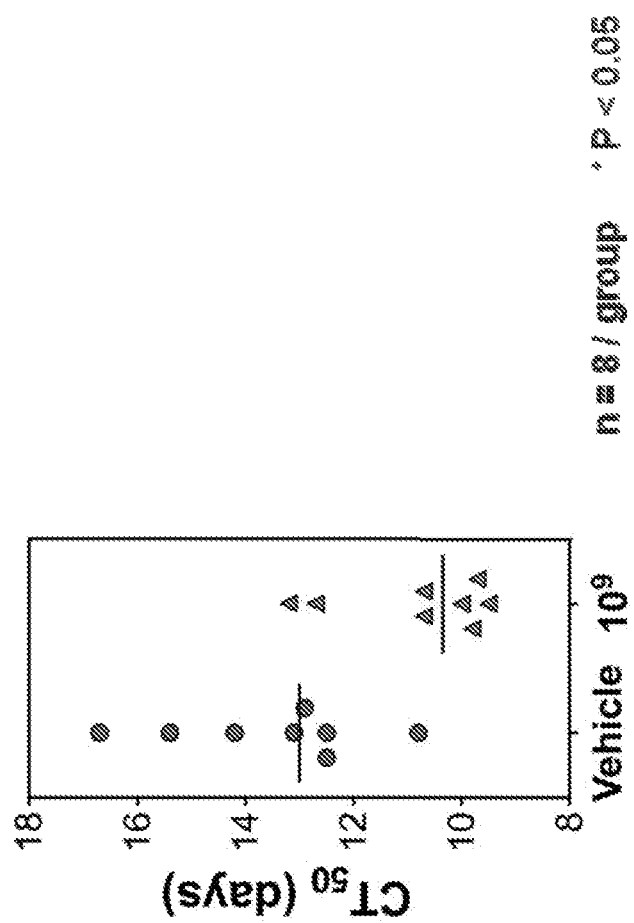
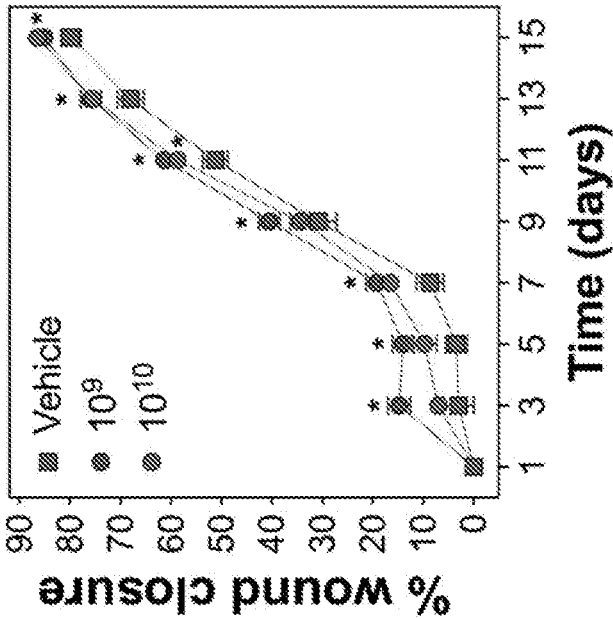


Figure 5D



N=7/group

* $p < 0.05$

Figure 5C

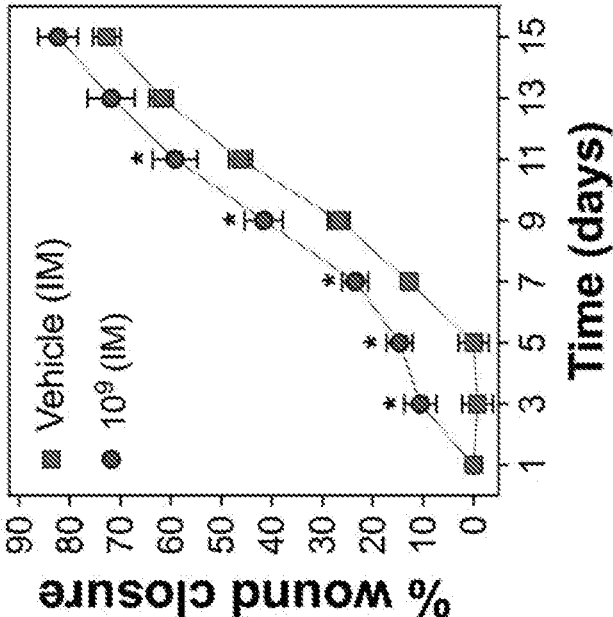


Figure 5E

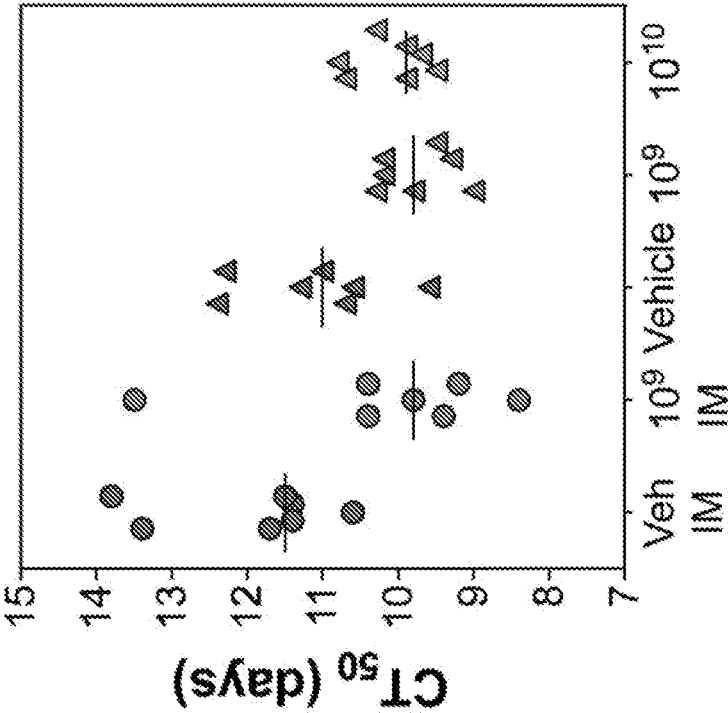


Figure 5

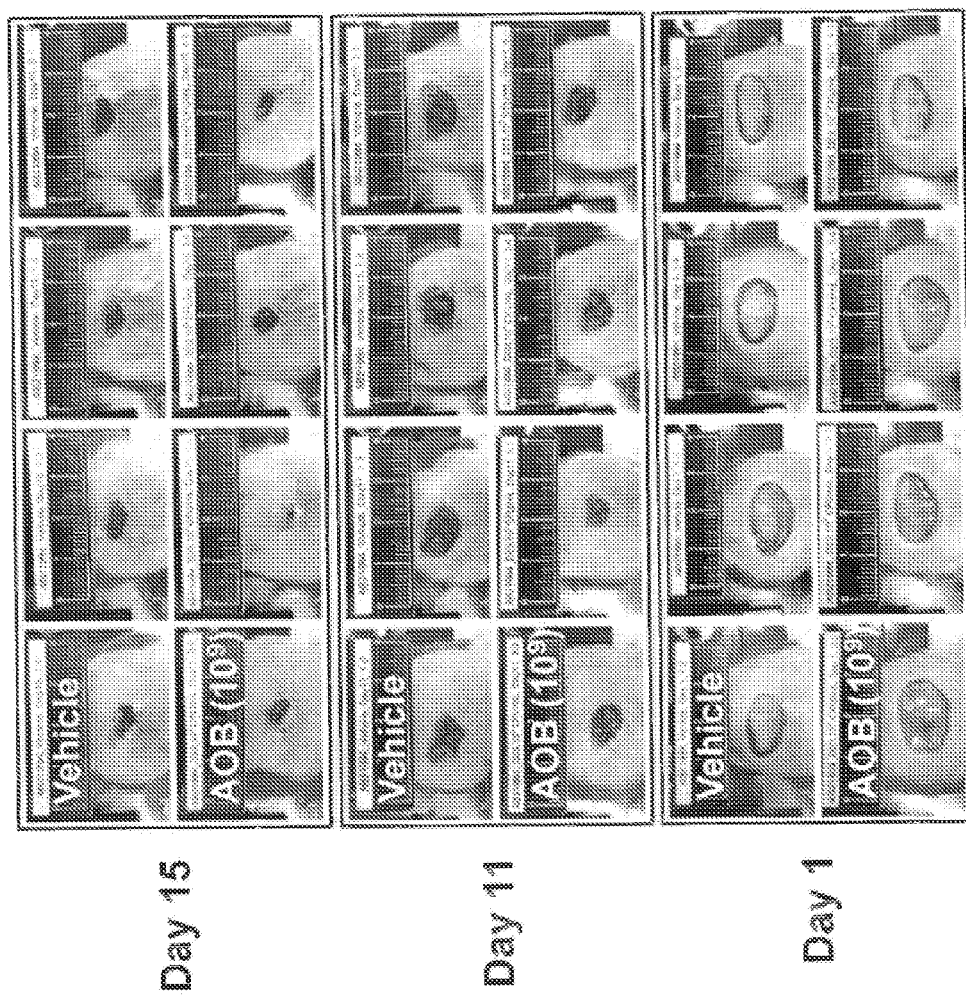
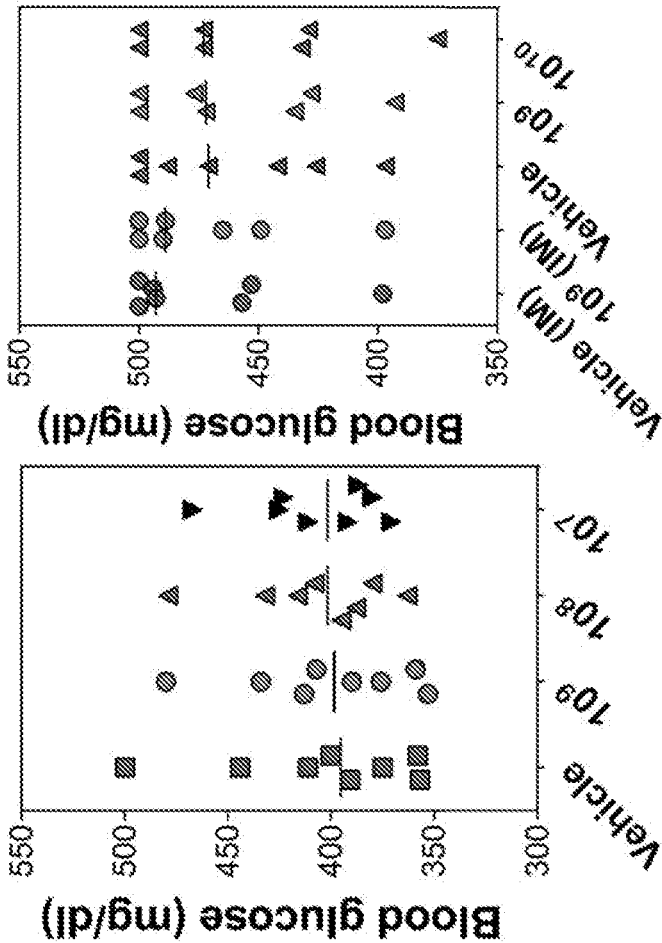


Figure 5G



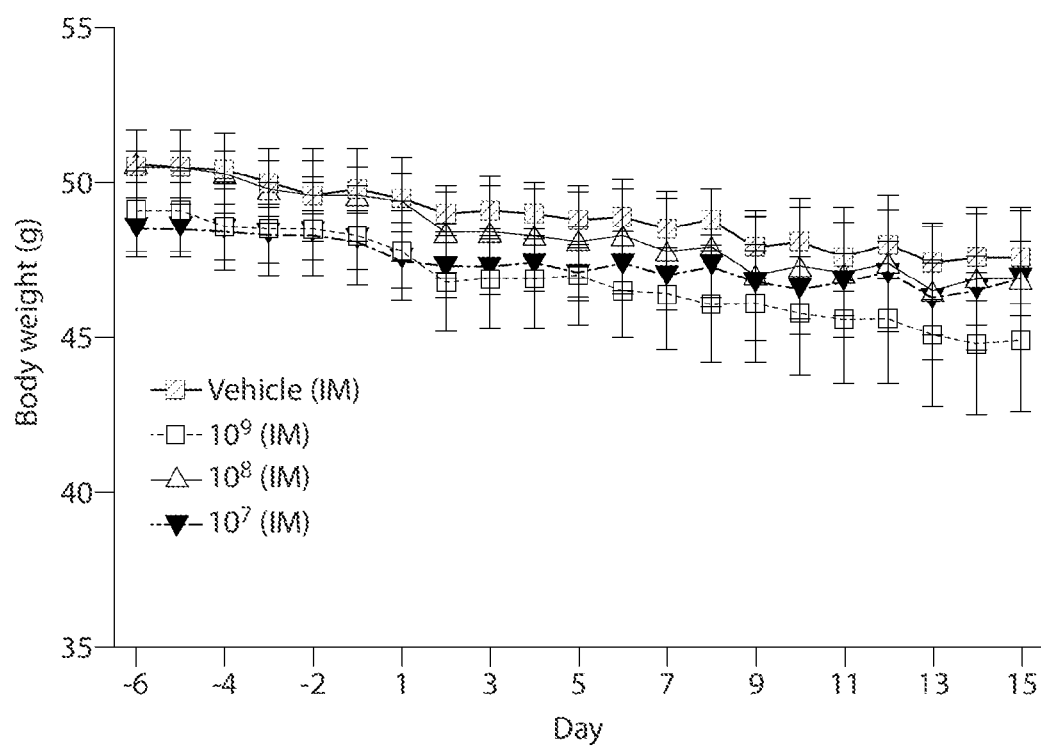


Figure 5H

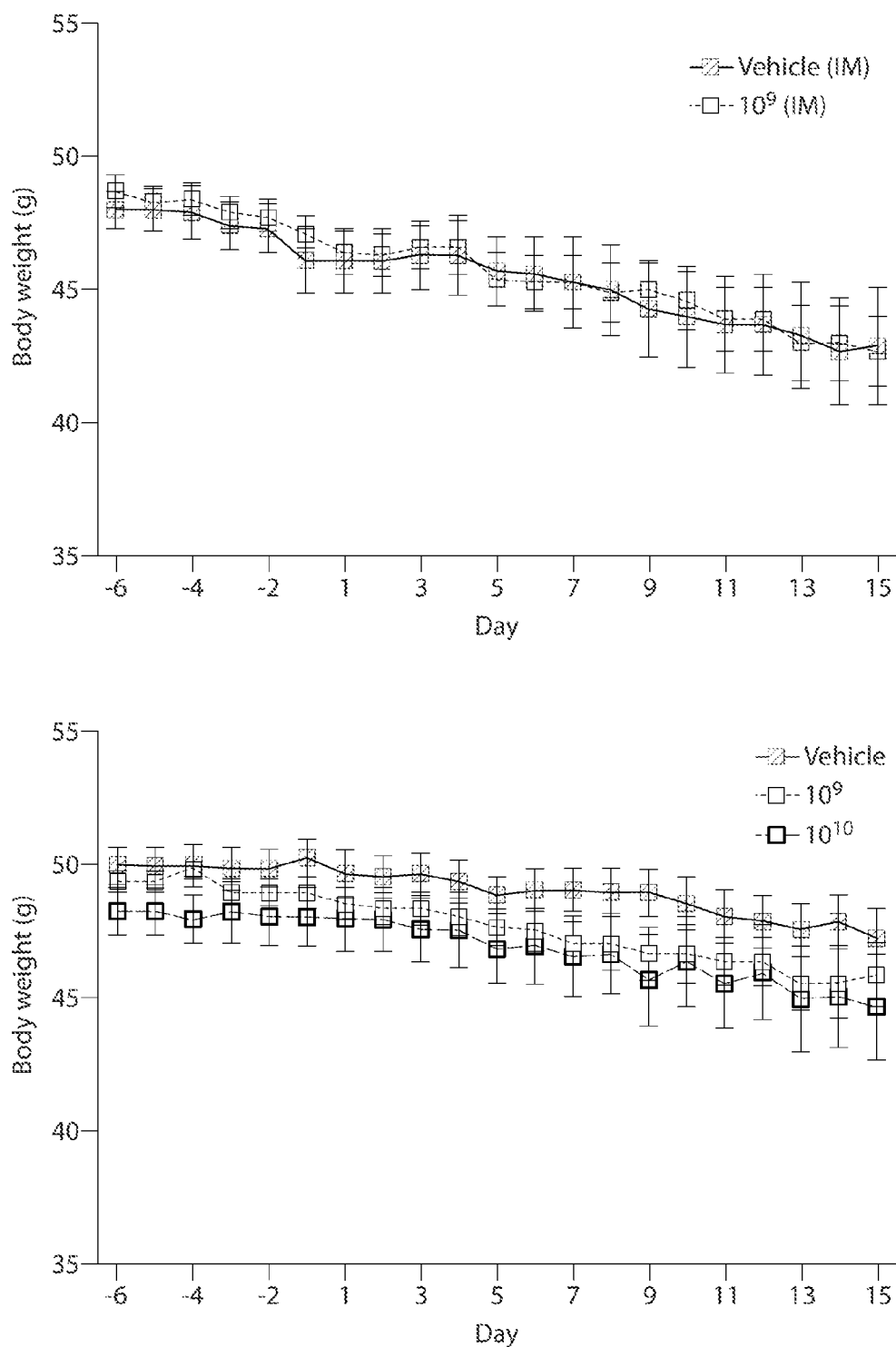


Figure 5I

Figure 5J

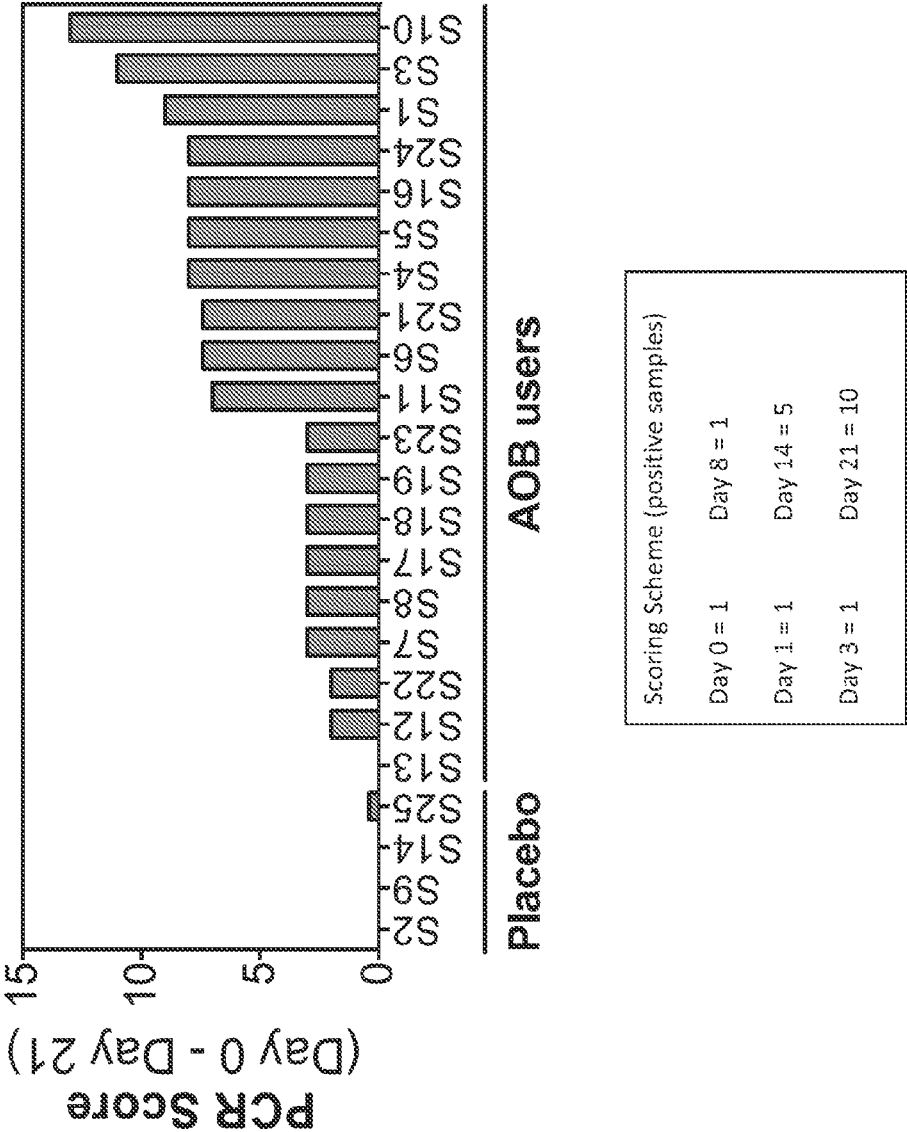


Figure 5K

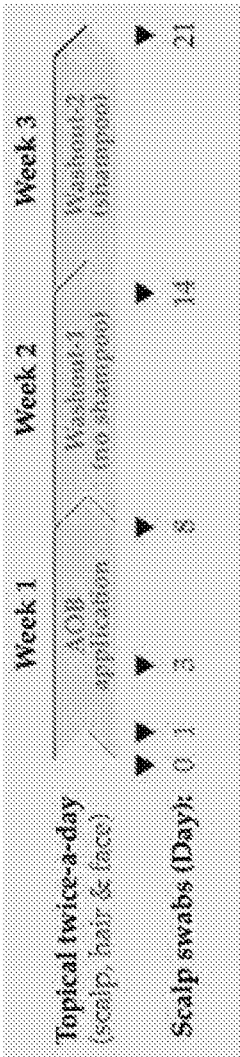


Figure 5L

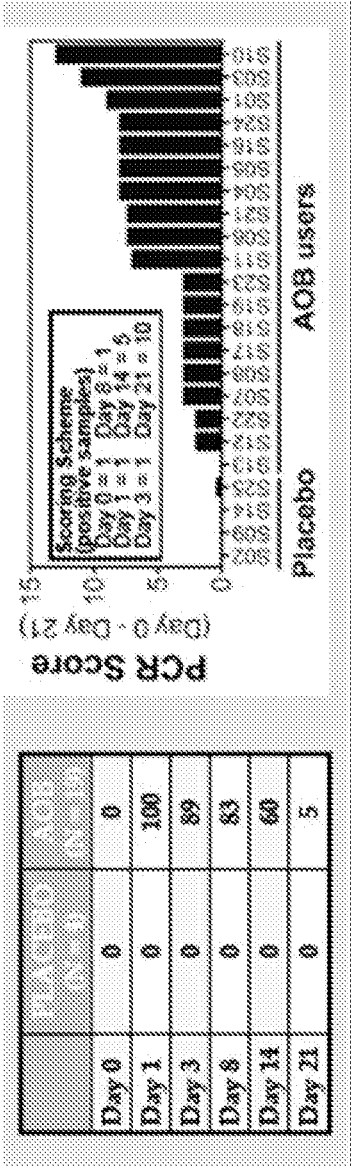
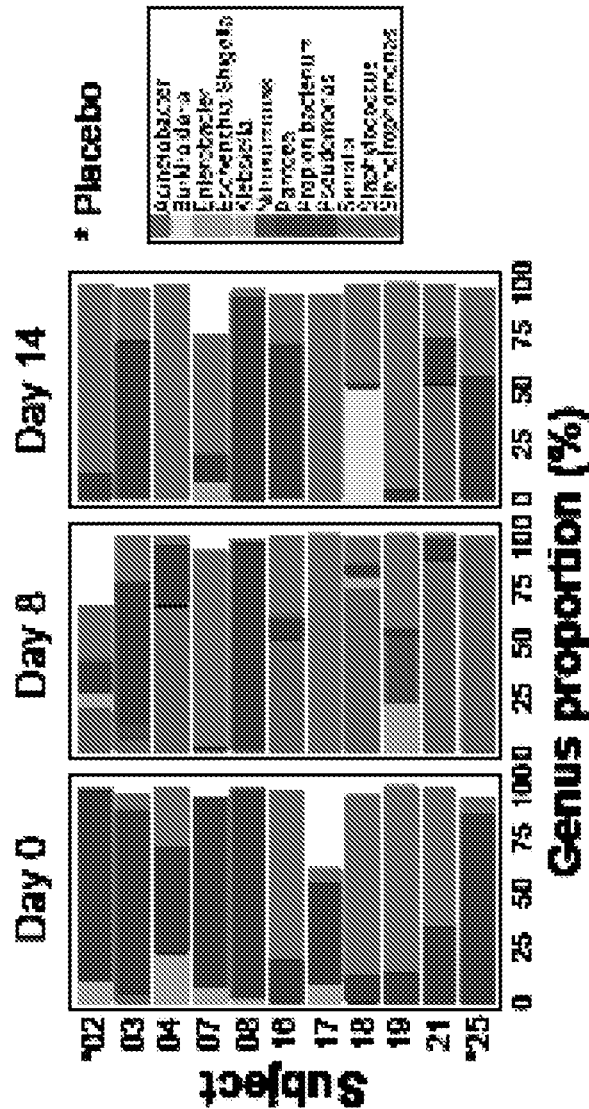


Figure 5M



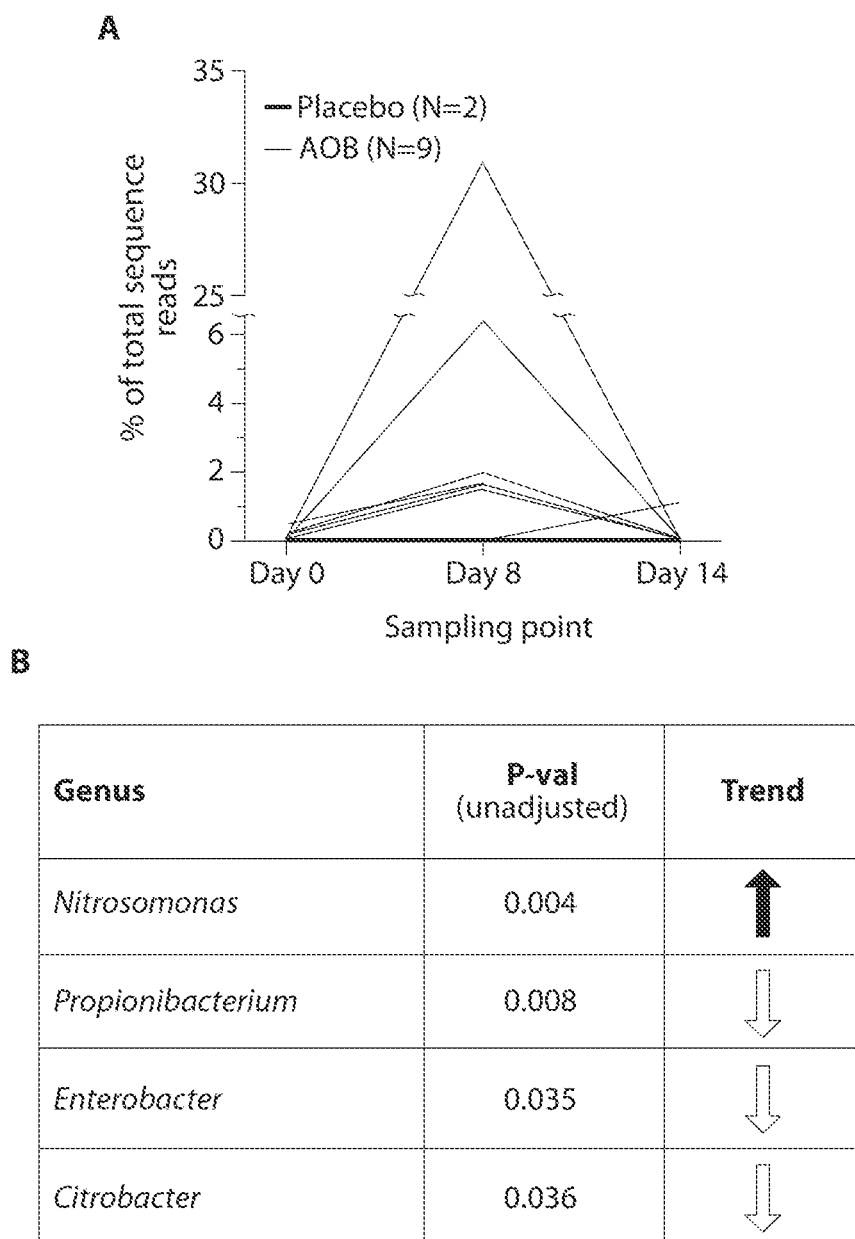


Figure 5N

Figure 50

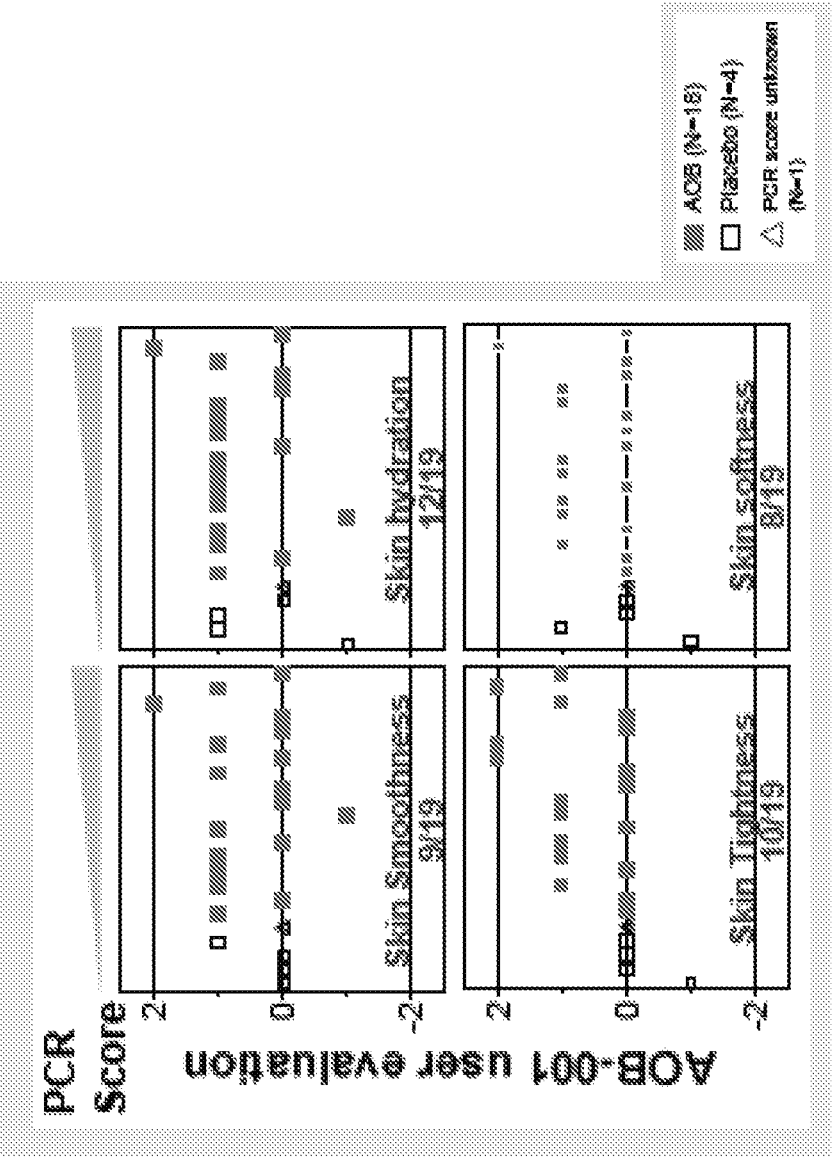


Figure 6

Unique D23 genes with assigned ORF number and function, or hypothetical above 200 base pairs: 162 genes

Feature.ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23GbkId
fig 666666666.60966.peg.23	CDS	29060	28851	-2	-	210	Death on curing protein, Doc toxin	Phd-Doc, YdcE-YdcD toxin-antitoxin (programmed cell death) systems	D23_1c0025
fig 666666666.60966.peg.33	CDS	34044	34424	3	+	381	hypothetical protein	- none -	D23_1c0034
fig 666666666.60966.peg.41	CDS	42537	43733	3	+	1197	glycosyl transferase, group 1/2 family protein	- none -	D23_1c0042
fig 666666666.60966.peg.47	CDS	47082	46804	-3	-	279	hypothetical protein	- none -	D23_1c0047
fig 666666666.60966.peg.55	CDS	51293	50880	-2	-	414	Mobile element protein	- none -	D23_1c0056
fig 666666666.60966.peg.69	CDS	64515	66023	3	+	1509	CBSS-49821.1.3.peg.1514: hypothetical protein	- none -	D23_1c0070

fig 666666666.609 66.peg.70	CDS	66074	66751	2	+	678	FIG039767: hypothetical protein	- none -	D23_1c0071
fig 666666666.609 66.peg.71	CDS	66741	70157	3	+	3417	FIG007317: hypothetical protein	- none -	D23_1c0072
fig 666666666.609 66.peg.158	CDS	157584	157859	3	+	276	DNA-3-methyladenine glycosylase II (EC 3.2.2.21)	DNA Repair Base Excision	D23_1c0165
fig 666666666.609 66.peg.207	CDS	198088	198819	1	+	732	Mobile element protein	- none -	D23_1c0213
fig 666666666.609 66.peg.215	CDS	203592	203461	-3	-	132	Phage Rha protein	- none -	D23_1c0220
fig 666666666.609 66.peg.235	CDS	219960	219271	-3	-	690	InterPro IPR001687	- none -	D23_1c0239
fig 666666666.609 66.peg.281	CDS	259254	259123	-3	-	132	cAMP-binding proteins - catabolite gene activator and regulatory subunit of cAMP-dependent protein kinases	cAMP signaling in bacteria	D23_1c0289
fig 666666666.609 66.peg.282	CDS	259543	260031	1	+	489	Cytochrome c'	- none -	D23_1c0291
fig 666666666.609 66.peg.289	CDS	263106	264041	3	+	936	hypothetical protein	- none -	D23_1c0297
fig 666666666.609 66.peg.290	CDS	264137	265633	2	+	1497	Slr1503 protein	- none -	D23_1c0298
fig 666666666.609 66.peg.302	CDS	273758	274108	2	+	351	hypothetical protein	- none -	D23_1c0307
fig 666666666.609 66.peg.471	CDS	440180	438420	-2	-	1761	Glucanase 2- dehydrogenase (EC 1.1.99.3), membrane-	D-gluconate and ketogluconates metabolism	D23_1c0475

									bound, flavoprotein		
fig 66666666.609 66.peg.474	CDS	441829	441158	-1	-	672			Gluconate 2- dehydrogenase (EC 1.1.99.3), membrane- bound, gamma subunit	D-gluconate and ketogluconates metabolism	D23_1c0478
fig 66666666.609 66.peg.487	CDS	455203	454049	-1	-	1155			hypothetical protein	- none -	D23_1c0491
fig 66666666.609 66.peg.493	CDS	462292	461345	-1	-	948			Putative DNA-binding protein in cluster with Type I restriction- modification system	Restriction- Modification System	D23_1c0497
fig 66666666.609 66.peg.508	CDS	474111	474269	3	+	159			Mobile element protein	- none -	D23_1c0513
fig 66666666.609 66.peg.510	CDS	474450	474653	3	+	204			hypothetical protein	- none -	D23_1c0514
fig 66666666.609 66.peg.577	CDS	535028	535240	2	+	213			hypothetical protein	- none -	D23_1c0578
fig 66666666.609 66.peg.579	CDS	537497	536616	-2	-	882			hypothetical protein	- none -	D23_1c0581
fig 66666666.609 66.peg.580	CDS	538547	537726	-2	-	822			hypothetical protein	- none -	D23_1c0582
fig 66666666.609 66.peg.581	CDS	539856	538789	-3	-	1068			Conserved domain protein	- none -	D23_1c0583
fig 66666666.609 66.peg.582	CDS	540712	539849	-1	-	864			Conserved domain protein	- none -	D23_1c0584
fig 66666666.609 66.peg.616	CDS	565168	564959	-1	-	210			hypothetical protein	- none -	D23_1c0618
fig 66666666.609 66.peg.677	CDS	626009	625605	-2	-	405			hypothetical protein	- none -	D23_1c0682

fig 6666666.609 66.peg.688	CDS	639034	639930	1	+	897	Quinolinate phosphoribosyltransfer ase [decarboxylating] (EC 2.4.2.19)	Mycobacterium virulence operon possibly involved in quinolinate biosynthesis; NAD and NADP cofactor biosynthesis global	D23_1c0695
fig 6666666.609 66.peg.741	CDS	687933	687400	-3	-	534	Adenylate kinase (EC 2.7.4.3)	Purine conversions	D23_1c0748
fig 6666666.609 66.peg.756	CDS	696806	696934	2	+	129	Mobile element protein	- none -	D23_1c0761
fig 6666666.609 66.peg.783	CDS	723822	723676	-3	-	147	Mobile element protein	- none -	D23_1c0788
fig 6666666.609 66.peg.785	CDS	724125	724304	3	+	180	Rubisco activation protein CbbO	CO2 uptake, carboxysome	D23_1c0791
fig 6666666.609 66.peg.850	CDS	781999	782385	1	+	387	hypothetical protein	- none -	D23_1c0856
fig 6666666.609 66.peg.851	CDS	782415	782747	3	+	333	Mobile element protein	- none -	D23_1c0857
fig 6666666.609 66.peg.855	CDS	785041	784757	-1	-	285	Mobile element protein	- none -	D23_1c0861
fig 6666666.609 66.peg.887	CDS	814624	815706	1	+	1083	glycosyltransferase	- none -	D23_1c0891
fig 6666666.609 66.peg.888	CDS	817016	816339	-2	-	678	hypothetical protein	- none -	D23_1c0892
fig 6666666.609 66.peg.988	CDS	919358	918504	-2	-	855	DNA-directed RNA polymerase alpha subunit (EC 2.7.7.6)	RNA polymerase bacterial	D23_1c0998

fig 6666666.609 66.peg.1077	CDS	100516 5	100496 2	-3	-	204	hypothetical protein	- none -	D23_1c1090
fig 6666666.609 66.peg.1169	CDS	109404 2	109341 3	-2	-	630	Nicotinamidase family protein YcaC	NAD and NADP cofactor biosynthesis global	D23_1c1176
fig 6666666.609 66.peg.1172	CDS	109533 4	109498 4	-1	-	351	hypothetical protein	- none -	D23_1c1180
fig 6666666.609 66.peg.1201	CDS	112400 5	112482 3	1	+	819	Cytochrome c family protein	- none -	D23_1c1207
fig 6666666.609 66.peg.1217	CDS	113757 6	113870 3	3	+	1128	Catalase (EC 1.11.1.6)	Oxidative stress; Photorespiration (oxidative C2 cycle); Protection from Reactive Oxygen Species	D23_1c1222
fig 6666666.609 66.peg.1225	CDS	114408 7	114341 9	-1	-	669	hypothetical protein	- none -	D23_1c1230
fig 6666666.609 66.peg.1226	CDS	114432 9	114407 8	-3	-	252	ABC-type antimicrobial peptide transport system, permease component	- none -	D23_1c1231
fig 6666666.609 66.peg.1233	CDS	115071 9	115050 7	-3	-	213	hypothetical protein	- none -	D23_1c1238
fig 6666666.609 66.peg.1237	CDS	115191 6	115227 8	3	+	363	hypothetical protein	- none -	D23_1c1242
fig 6666666.609 66.peg.1272	CDS	118135 5	118155 8	3	+	204	hypothetical protein	- none -	D23_1c1279
fig 6666666.609 66.peg.1274	CDS	118174 8	118198 1	3	+	234	Mobile element protein	- none -	D23_1c1281
fig 6666666.609 66.peg.1303	CDS	120913 9	120837 5	-1	-	765	Mobile element protein	- none -	D23_1c1311

fig 6666666.609 66.peg.1305	CDS	121046 6	121013 4	-2	-	333	hypothetical protein	- none -	D23_1c1313
fig 6666666.609 66.peg.1321	CDS	122381 5	122405 4	1	+	240	hypothetical protein	- none -	D23_1c1328
fig 6666666.609 66.peg.1341	CDS	124473 6	124353 1	-3	-	1206	Catalase (EC 1.11.1.6)	Oxidative stress; Photorespiration (oxidative C2 cycle); Protection from Reactive Oxygen Species	D23_1c1348
fig 6666666.609 66.peg.1342	CDS	124626 8	124473 9	-2	-	1530	Peroxidase (EC 1.11.1.7)	Oxidative stress; Protection from Reactive Oxygen Species	D23_1c1349
fig 6666666.609 66.peg.1343	CDS	124758 5	124628 1	-2	-	1305	Peroxidase (EC 1.11.1.7)	Oxidative stress; Protection from Reactive Oxygen Species	D23_1c1350
fig 6666666.609 66.peg.1344	CDS	124807 8	124762 3	-3	-	456	hypothetical protein	- none -	D23_1c1351
fig 6666666.609 66.peg.1346	CDS	125121 8	125315 5	2	+	1938	Choline dehydrogenase (EC 1.1.99.1)	- none -	D23_1c1353
fig 6666666.609 66.peg.1347	CDS	125605 1	125318 4	-2	-	2868	Peroxidase (EC 1.11.1.8)	- none -	D23_1c1354

fig 666666.609 66.peg.1348	CDS	125750 5	125608 1	-1	-	1425	Hemagglutinin	- none -	D23_1c1355
fig 666666.609 66.peg.1349	CDS	125825 7	125754 7	-3	-	711	hypothetical protein	- none -	D23_1c1356
fig 666666.609 66.peg.1350	CDS	125923 3	125826 2	-1	-	972	hypothetical protein	- none -	D23_1c1357
fig 666666.609 66.peg.1351	CDS	126107 8	125924 6	-1	-	1833	hypothetical protein	- none -	D23_1c1358
fig 666666.609 66.peg.1352	CDS	126277 6	126110 9	-1	-	1668	Arachidonate 15- lipoxygenase (EC 1.13.11.33)	- none -	D23_1c1359
fig 666666.609 66.peg.1353	CDS	126444 9	126284 5	-3	-	1605	putative cyclooxygenase-2	- none -	D23_1c1360
fig 666666.609 66.peg.1354	CDS	126613 4	126463 8	-2	-	1497	hypothetical protein	- none -	D23_1c1361
fig 666666.609 66.peg.1355	CDS	126638 2	126616 7	-1	-	216	hypothetical protein	- none -	D23_1c1362
fig 666666.609 66.peg.1356	CDS	126670 4	126644 4	-2	-	261	hypothetical protein	- none -	D23_1c1363
fig 666666.609 66.peg.1357	CDS	126780 4	126698 3	-1	-	822	Kazal-type serine protease inhibitor domain	- none -	D23_1c1364

fig 666666.609 66.peg.1358	CDS	126865 0	126797 0	-1	-	681	Kazal-type serine protease inhibitor domain	- none -	D23_1c1365
fig 666666.609 66.peg.1363	CDS	127133 2	127040 9	-1	-	924	alpha/beta hydrolase fold	- none -	D23_1c1370
fig 666666.609 66.peg.1364	CDS	127172 1	127132 9	-3	-	393	hypothetical protein	- none -	D23_1c1371
fig 666666.609 66.peg.1371	CDS	127903 1	127802 7	-2	-	1005	Putative DNA-binding protein in cluster with Type I restriction- modification system	Restriction- Modification System	D23_1c1378
fig 666666.609 66.peg.1373	CDS	128327 2	128298 2	-1	-	291	Type I restriction- modification system, restriction subunit R (EC 3.1.21.3)	Restriction- Modification System; Type I Restriction- Modification	D23_1c1380
fig 666666.609 66.peg.1405	CDS	131819 2	131782 7	-1	-	366	Putative metal chaperone, involved in Zn homeostasis, GTPase of COG0523 family	G3E family of P- loop GTPases (metallocenter biosynthesis); Zinc regulated enzymes	D23_1c1412
fig 666666.609 66.peg.1408	CDS	132031 4	132165 1	2	+	1338	DNA modification methyltransferase	- none -	D23_1c1415
fig 666666.609 66.peg.1409	CDS	132165 4	132253 2	1	+	879	hypothetical protein	- none -	D23_1c1416
fig 666666.609	CDS	132266	132288	1	+	216	hypothetical protein	- none -	D23_1c1417

fig 6666666.609 66.peg.1554	CDS	145674 0	145650 1	-3	-	240	hypothetical protein	- none -	D23_1c1561
fig 6666666.609 66.peg.1557	CDS	145767 8	145727 7	-2	-	402	hypothetical protein	- none -	D23_1c1564
fig 6666666.609 66.peg.1558	CDS	145813 7	145847 8	2	+	342	hypothetical protein	- none -	D23_1c1565
fig 6666666.609 66.peg.1559	CDS	145886 8	145915 8	1	+	291	hypothetical protein	- none -	D23_1c1566
fig 6666666.609 66.peg.1569	CDS	146480 6	146382 6	-2	-	981	Abortive infection bacteriophage resistance protein	- none -	D23_1c1573
fig 6666666.609 66.peg.1648	CDS	153278 7	153013 6	-3	-	2652	Malto- oligosyltrehalose synthase (EC 5.4.99.15)	- none -	D23_1c1646
fig 6666666.609 66.peg.1651	CDS	153532 7	153489 6	-2	-	432	Trehalose synthase, nucleoside diphosphate glucose dependent	- none -	D23_1c1649
fig 6666666.609 66.peg.1689	CDS	157332 0	157312 9	-3	-	192	FIG00859257: hypothetical protein	- none -	D23_1c1687
fig 6666666.609 66.peg.1694	CDS	157557 2	157536 9	-2	-	204	hypothetical protein	- none -	D23_1c1693
fig 6666666.609 66.peg.1725	CDS	160874 1	160659 7	-3	-	2145	hypothetical protein	- none -	D23_1c1723
fig 6666666.609	CDS	161007	160873	-2	-	1335	hypothetical protein	- none -	D23_1c1724

fig 6666666.609 66.peg.2001	CDS	186661 8	186632 8	-3	-	291	hypothetical protein	- none -	D23_1c2004
fig 6666666.609 66.peg.2097	CDS	194925 2	194938 6	2	+	135	Type I restriction- modification system, restriction subunit R (EC 3.1.21.3)	Restriction- Modification System; Type I Restriction- Modification	D23_1c2099
fig 6666666.609 66.peg.2111	CDS	196119 1	196151 7	1	+	327	hypothetical protein	- none -	D23_1c2112
fig 6666666.609 66.peg.2127	CDS	197651 4	197498 5	-3	-	1530	Capsular polysaccharide biosynthesis protein WcbQ	Capsular Polysaccharides Biosynthesis and Assembly	D23_1c2128
fig 6666666.609 66.peg.2132	CDS	198221 4	198123 1	-3	-	984	hypothetical protein	- none -	D23_1c2133
fig 6666666.609 66.peg.2133	CDS	198277 8	198221 8	-3	-	561	hypothetical protein	- none -	D23_1c2134
fig 6666666.609 66.peg.2137	CDS	198507 6	198485 2	-3	-	225	hypothetical protein	- none -	D23_1c2138
fig 6666666.609 66.peg.2156	CDS	199789 0	199809 3	1	+	204	Mobile element protein	- none -	D23_1c2154
fig 6666666.609 66.peg.2157	CDS	199856 5	199826 6	-1	-	300	VapC toxin protein	Toxin-antitoxin replicon stabilization systems	D23_1c2155
fig 6666666.609 66.peg.2158	CDS	199889 9	199866 6	-2	-	234	VapB protein (antitoxin to VapC)	Toxin-antitoxin replicon stabilization systems	D23_1c2156
fig 6666666.609 66.peg.2160	CDS	199962 9	199933 9	-3	-	291	transcriptional regulator, XRE family	- none -	D23_1c2158

fig 666666.609 66.peg.2161	CDS	199999 7	199969 2	-2	-	306	Phage-related protein	- none -	D23_1c2159
fig 666666.609 66.peg.2163	CDS	200022 0	200048 6	3	+	267	Mobile element protein	- none -	D23_1c2161
fig 666666.609 66.peg.2164	CDS	200050 7	200099 5	2	+	489	Mobile element protein	- none -	D23_1c2162
fig 666666.609 66.peg.2167	CDS	200201 1	200232 8	3	+	318	Mobile element protein	- none -	D23_1c2165
fig 666666.609 66.peg.2168	CDS	200352 9	200236 9	-3	-	1161	CDP-4-dehydro-6- deoxy-D-glucose 3- dehydratase (EC 4.2.1.-)	- none -	D23_1c2166
fig 666666.609 66.peg.2169	CDS	200442 6	200353 9	-3	-	888	NAD-dependent epimerase/dehydratas e family protein	CBSS- 296591.1.peg.2330	D23_1c2167
fig 666666.609 66.peg.2171	CDS	200746 7	200563 2	-2	-	1836	hypothetical protein	- none -	D23_1c2169
fig 666666.609 66.peg.2172	CDS	201056 8	200747 3	-1	-	3096	Minor teichoic acid biosynthesis protein GgaB	- none -	D23_1c2170
fig 666666.609 66.peg.2173	CDS	201216 3	201069 4	-3	-	1470	InterPro IPR001173 COGs COG0463	- none -	D23_1c2171
fig 666666.609 66.peg.2174	CDS	201582 8	201217 2	-2	-	3657	Beta-1,3- glucosyltransferase	- none -	D23_1c2172
fig 666666.609 66.peg.2175	CDS	201730 7	201614 7	-2	-	1161	Capsular polysaccharide export system inner membrane protein KpsE	Capsular Polysaccharides Biosynthesis and Assembly	D23_1c2173

fig 666666666.609 66.peg.2177	CDS	201875 7	201796 3	-3	-	795	Capsular polysaccharide ABC transporter, permease protein KpsM	Capsular Polysaccharides Biosynthesis and Assembly; Rhamnose containing glycans	D23_1c2175
fig 666666666.609 66.peg.2180	CDS	202512 1	202127 6	-1	-	3846	Capsular polysaccharide biosynthesis fatty acid synthase WcbR	- none -	D23_1c2178
fig 666666666.609 66.peg.2225	CDS	206790 7	206839 5	1	+	489	Zinc uptake regulation protein ZUR	Glycyl-tRNA synthetase containing cluster; Oxidative stress; Zinc regulated enzymes	D23_1c2222
fig 666666666.609 66.peg.2228	CDS	206881 1	206902 0	2	+	210	hypothetical protein	- none -	D23_1c2224
fig 666666666.609 66.peg.2320	CDS	215313 2	215181 6	-2	-	1317	hypothetical protein	- none -	D23_1c2315
fig 666666666.609 66.peg.2418	CDS	224023 9	224009 3	-1	-	147	Aconitate hydratase (EC 4.2.1.3)	TCA Cycle	D23_1c2410
fig 666666666.609 66.peg.2421	CDS	224183 5	224198 7	1	+	153	Mobile element protein	- none -	D23_1c2413
fig 666666666.609 66.peg.2422	CDS	224223 2	224252 2	2	+	291	hypothetical protein	- none -	D23_1c2414
fig 666666666.609 66.peg.2457	CDS	227575 8	227499 7	-3	-	762	hypothetical protein	- none -	D23_1c2447
fig 666666666.609 66.peg.2465	CDS	228271 6	228135 2	-1	-	1365	hypothetical protein	- none -	D23_1c2455
fig 666666666.609	CDS	228501	228402	-2	-	996	hypothetical protein	- none -	D23_1c2459

66.p _{eg} .2469 fig 6666666.609 66.p _{eg} .2470	CDS	228530 5	5	0	-1	-	258	hypothetical protein	- none -	D23_1c2460
fig 6666666.609 66.p _{eg} .2553	CDS	236155 1	236138 4	-2	-	-	168	protein of unknown function DUF1328	- none -	D23_1c2543
fig 6666666.609 66.p _{eg} .2576	CDS	237876 5	237834 0	-2	-	-	426	hypothetical protein	- none -	D23_1c2565
fig 6666666.609 66.p _{eg} .2594	CDS	239214 5	239256 7	2	+	-	423	Putative TEGT family carrier/transport protein	CBS-326442.4.p _{eg} .1852	D23_1c2580
fig 6666666.609 66.p _{eg} .2652	CDS	245799 8	245748 0	-2	-	-	519	HrgA protein	- none -	D23_1c2641
fig 6666666.609 66.p _{eg} .2653	CDS	245935 0	245801 3	-1	-	-	1338	hypothetical protein	- none -	D23_1c2642
fig 6666666.609 66.p _{eg} .2654	CDS	246024 9	245934 7	-3	-	-	903	hypothetical protein	- none -	D23_1c2643
fig 6666666.609 66.p _{eg} .2655	CDS	246330 5	246026 4	-2	-	-	3042	Type I restriction-modification system, restriction subunit R (EC 3.1.21.3)	Restriction-Modification System; Type I Restriction-Modification	D23_1c2644
fig 6666666.609 66.p _{eg} .2656	CDS	246394 3	246334 1	-1	-	-	603	hypothetical protein	- none -	D23_1c2645
fig 6666666.609 66.p _{eg} .2657	CDS	246534 8	246396 0	-2	-	-	1389	Type I restriction-modification system, specificity subunit S (EC 3.1.21.3)	Restriction-Modification System; Type I Restriction-Modification	D23_1c2646
fig 6666666.609 66.p _{eg} .2685	CDS	249923 0	249952 0	2	+	-	291	hypothetical protein	- none -	D23_1c2677

fig 666666666.609 66.peg.2687	CDS	250140 2	250076 7	-2	-	636	FIG01057587: hypothetical protein	- none -	D23_1c2679
fig 666666666.609 66.peg.2692	CDS	250389 3	250410 5	3	+	213	hypothetical protein	- none -	D23_1c2683
fig 666666666.609 66.peg.2693	CDS	250523 9	250459 5	-2	-	645	hypothetical protein	- none -	D23_1c2685
fig 666666666.609 66.peg.2696	CDS	250692 4	250644 5	-1	-	480	DNA primase/helicase, phage-associated	Phage replication	D23_1c2688
fig 666666666.609 66.peg.2697	CDS	250716 6	250692 1	-3	-	246	hypothetical protein	- none -	D23_1c2689
fig 666666666.609 66.peg.2698	CDS	250739 3	250716 6	-2	-	228	hypothetical protein	- none -	D23_1c2690
fig 666666666.609 66.peg.2699	CDS	250773 5	250857 7	2	+	843	hypothetical protein	- none -	D23_1c2691
fig 666666666.609 66.peg.2700	CDS	250865 0	250911 1	2	+	462	Putative bacteriophage-related protein	- none -	D23_1c2692
fig 666666666.609 66.peg.2702	CDS	251047 8	251089 7	3	+	420	elements of external origin; phage-related functions and prophages	- none -	D23_1c2694
fig 666666666.609 66.peg.2733	CDS	253738 8	253707 7	-3	-	312	hypothetical protein	- none -	D23_1c2724

Figure 7

Unique D23 genes below 200 base pairs with assigned ORF number: 164

Feature.ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23GbkId
fig 666666.609.66.peg.22	CDS	28682	28867	2	+	186	hypothetical protein	- none -	D23_1c0024
fig 666666.609.66.peg.29	CDS	31959	32078	3	+	120	hypothetical protein	- none -	D23_1c0030
fig 666666.609.66.peg.38	CDS	36580	36750	1	+	171	hypothetical protein	- none -	D23_1c0039
fig 666666.609.66.peg.52	CDS	49615	49502	-1	-	114	hypothetical protein	- none -	D23_1c0052
fig 666666.609.66.peg.64	CDS	56874	56746	-3	-	129	hypothetical protein	- none -	D23_1c0065
fig 666666.609.66.peg.66	CDS	60633	60755	3	+	123	hypothetical protein	- none -	D23_1c0067
fig 666666.609.66.peg.92	CDS	89253	89375	3	+	123	hypothetical protein	- none -	D23_1c0091
fig 666666.609.66.peg.93	CDS	89433	89579	3	+	147	hypothetical protein	- none -	D23_1c0092
fig 666666.609.66.peg.119	CDS	120193	120056	-1	-	138	hypothetical protein	- none -	D23_1c0120
fig 666666.609.66.peg.121	CDS	122376	122495	3	+	120	hypothetical protein	- none -	D23_1c0124
fig 666666.609.66.peg.120	CDS	121863	121994	3	+	132	hypothetical protein	- none -	D23_1c0124

fig 6666666.609 66.peg.123	CDS	124878	124708	-3	-	171	hypothetical protein	- none -	D23_1c0127
fig 6666666.609 66.peg.132	CDS	132190	132312	1	+	123	hypothetical protein	- none -	D23_1c0137
fig 6666666.609 66.peg.210	CDS	200398	200210	-1	-	189	hypothetical protein	- none -	D23_1c0215
fig 6666666.609 66.peg.231	CDS	215510	215623	2	+	114	hypothetical protein	- none -	D23_1c0235
fig 6666666.609 66.peg.246	CDS	231229	231411	1	+	183	hypothetical protein	- none -	D23_1c0253
fig 6666666.609 66.peg.249	CDS	232767	232585	-3	-	183	hypothetical protein	- none -	D23_1c0256
fig 6666666.609 66.peg.256	CDS	237299	237415	2	+	117	hypothetical protein	- none -	D23_1c0264
fig 6666666.609 66.peg.294	CDS	266897	266760	-2	-	138	hypothetical protein	- none -	D23_1c0300
fig 6666666.609 66.peg.304	CDS	274944	274792	-3	-	153	hypothetical protein	- none -	D23_1c0309
fig 6666666.609 66.peg.313	CDS	284258	284401	2	+	144	hypothetical protein	- none -	D23_1c0317
fig 6666666.609 66.peg.315	CDS	286328	286191	-2	-	138	hypothetical protein	- none -	D23_1c0321
fig 6666666.609 66.peg.330	CDS	302493	302368	-3	-	126	hypothetical protein	- none -	D23_1c0338
fig 6666666.609 66.peg.334	CDS	306026	305904	-2	-	123	hypothetical protein	- none -	D23_1c0342
fig 6666666.609 66.peg.349	CDS	322196	322363	2	+	168	hypothetical protein	- none -	D23_1c0356
fig 6666666.609 66.peg.351	CDS	325139	325255	2	+	117	hypothetical protein	- none -	D23_1c0358
fig 6666666.609	CDS	342384	342509	3	+	126	hypothetical protein	- none -	D23_1c0376

[illegible]

fig 6666666.609 66.peg.571	CDS	530407	530535	1	+	129	hypothetical protein	- none -	D23_1c0572
fig 6666666.609 66.peg.584	CDS	541934	541812	-2	-	123	hypothetical protein	- none -	D23_1c0586
fig 6666666.609 66.peg.591	CDS	544721	544834	2	+	114	hypothetical protein	- none -	D23_1c0591
fig 6666666.609 66.peg.600	CDS	551404	551517	1	+	114	hypothetical protein	- none -	D23_1c0600
fig 6666666.609 66.peg.630	CDS	576475	576609	1	+	135	hypothetical protein	- none -	D23_1c0633
fig 6666666.609 66.peg.667	CDS	615143	615012	-2	-	132	hypothetical protein	- none -	D23_1c0670
fig 6666666.609 66.peg.674	CDS	623711	623827	2	+	117	hypothetical protein	- none -	D23_1c0678
fig 6666666.609 66.peg.686	CDS	637484	637326	-2	-	159	hypothetical protein	- none -	D23_1c0693
fig 6666666.609 66.peg.695	CDS	647548	647402	-1	-	147	hypothetical protein	- none -	D23_1c0702
fig 6666666.609 66.peg.702	CDS	652766	652948	2	+	183	hypothetical protein	- none -	D23_1c0709
fig 6666666.609 66.peg.727	CDS	674585	674716	2	+	132	hypothetical protein	- none -	D23_1c0734
fig 6666666.609 66.peg.729	CDS	675190	674996	-1	-	195	hypothetical protein	- none -	D23_1c0736
fig 6666666.609 66.peg.778	CDS	718480	718674	1	+	195	hypothetical protein	- none -	D23_1c0783
fig 6666666.609 66.peg.784	CDS	723900	724055	3	+	156	hypothetical protein	- none -	D23_1c0790
fig 6666666.609 66.peg.800	CDS	734018	734161	2	+	144	hypothetical protein	- none -	D23_1c0807
fig 6666666.609	CDS	739386	739240	-3	-	147	hypothetical protein	- none -	D23_1c0813

fig 6666666.609 66.peg.1027	CDS	952378	952208	-1	-	171	hypothetical protein	- none -	D23_1c1036
fig 6666666.609 66.peg.1040	CDS	965355	965495	3	+	141	hypothetical protein	- none -	D23_1c1051
fig 6666666.609 66.peg.1051	CDS	973589	973461	-2	-	129	hypothetical protein	- none -	D23_1c1064
fig 6666666.609 66.peg.1069	CDS	992670	992536	-3	-	135	hypothetical protein	- none -	D23_1c1083
fig 6666666.609 66.peg.1091	CDS	1020710	1020862	2	+	153	hypothetical protein	- none -	D23_1c1103
fig 6666666.609 66.peg.1112	CDS	1043119	1043253	1	+	135	hypothetical protein	- none -	D23_1c1121
fig 6666666.609 66.peg.1113	CDS	1043368	1043511	1	+	144	hypothetical protein	- none -	D23_1c1122
fig 6666666.609 66.peg.1115	CDS	1043900	1043736	-2	-	165	hypothetical protein	- none -	D23_1c1123
fig 6666666.609 66.peg.1139	CDS	1065051	1064932	-3	-	120	hypothetical protein	- none -	D23_1c1149
fig 6666666.609 66.peg.1168	CDS	1092982	1092851	-1	-	132	hypothetical protein	- none -	D23_1c1175
fig 6666666.609 66.peg.1198	CDS	1121909	1121763	-2	-	147	hypothetical protein	- none -	D23_1c1204
fig 6666666.609 66.peg.1205	CDS	1128452	1128580	2	+	129	hypothetical protein	- none -	D23_1c1211
fig 6666666.609 66.peg.1232	CDS	1150415	1150546	2	+	132	hypothetical protein	- none -	D23_1c1237
fig 6666666.609 66.peg.1235	CDS	1150882	1151043	1	+	162	hypothetical protein	- none -	D23_1c1240
fig 6666666.609 66.peg.1250	CDS	1161926	1161813	-2	-	114	hypothetical protein	- none -	D23_1c1257
fig 6666666.609	CDS	1162096	1162263	1	+	168	hypothetical protein	- none -	D23_1c1258

66.peg.1251													
fig 6666666.609 66.peg.1252	CDS	1162260	1162415	3	+	156	hypothetical protein	- none -	D23_1c1259				
fig 6666666.609 66.peg.1273	CDS	1181609	1181776	2	+	168	hypothetical protein	- none -	D23_1c1280				
fig 6666666.609 66.peg.1275	CDS	1182136	1181990	-1	-	147	hypothetical protein	- none -	D23_1c1282				
fig 6666666.609 66.peg.1285	CDS	1189440	1189327	-3	-	114	hypothetical protein	- none -	D23_1c1290				
fig 6666666.609 66.peg.1292	CDS	1200901	1200761	-1	-	141	hypothetical protein	- none -	D23_1c1301				
fig 6666666.609 66.peg.1300	CDS	1206663	1206902	3	+	240	hypothetical protein	- none -	D23_1c1308				
fig 6666666.609 66.peg.1359	CDS	1268913	1268800	-3	-	114	hypothetical protein	- none -	D23_1c1366				
fig 6666666.609 66.peg.1387	CDS	1298101	1297967	-1	-	135	hypothetical protein	- none -	D23_1c1393				
fig 6666666.609 66.peg.1395	CDS	1303729	1303556	-1	-	174	hypothetical protein	- none -	D23_1c1402				
fig 6666666.609 66.peg.1412	CDS	1324808	1324990	2	+	183	hypothetical protein	- none -	D23_1c1419				
fig 6666666.609 66.peg.1438	CDS	1349336	1349214	-2	-	123	hypothetical protein	- none -	D23_1c1444				
fig 6666666.609 66.peg.1462	CDS	1374021	1374137	3	+	117	hypothetical protein	- none -	D23_1c1469				
fig 6666666.609 66.peg.1466	CDS	1380471	1380340	-3	-	132	hypothetical protein	- none -	D23_1c1473				
fig 6666666.609 66.peg.1475	CDS	1388559	1388395	-3	-	165	hypothetical protein	- none -	D23_1c1481				
fig 6666666.609 66.peg.1480	CDS	1394795	1394911	2	+	117	hypothetical protein	- none -	D23_1c1486				

fig 66666666.609 66.peg.1488	CDS	1402082	1401924	-2	-	159	hypothetical protein	- none -	D23_1c1494
fig 66666666.609 66.peg.1503	CDS	1415346	1415224	-3	-	123	hypothetical protein	- none -	D23_1c1510
fig 66666666.609 66.peg.1520	CDS	1430834	1430670	-2	-	165	hypothetical protein	- none -	D23_1c1526
fig 66666666.609 66.peg.1550	CDS	1453430	1453269	-2	-	162	hypothetical protein	- none -	D23_1c1557
fig 66666666.609 66.peg.1551	CDS	1453630	1453451	-1	-	180	hypothetical protein	- none -	D23_1c1558
fig 66666666.609 66.peg.1561	CDS	1459812	1459672	-3	-	141	hypothetical protein	- none -	D23_1c1567
fig 66666666.609 66.peg.1567	CDS	1462722	1462844	3	+	123	hypothetical protein	- none -	D23_1c1571
fig 66666666.609 66.peg.1573	CDS	1466953	1467108	1	+	156	hypothetical protein	- none -	D23_1c1575
fig 66666666.609 66.peg.1603	CDS	1491738	1491619	-3	-	120	hypothetical protein	- none -	D23_1c1604
fig 66666666.609 66.peg.1619	CDS	1502854	1502738	-1	-	117	hypothetical protein	- none -	D23_1c1617
fig 66666666.609 66.peg.1623	CDS	1505626	1505739	1	+	114	hypothetical protein	- none -	D23_1c1621
fig 66666666.609 66.peg.1649	CDS	1533038	1532850	-2	-	189	hypothetical protein	- none -	D23_1c1647
fig 66666666.609 66.peg.1652	CDS	1535502	1535386	-3	-	117	hypothetical protein	- none -	D23_1c1650
fig 66666666.609 66.peg.1666	CDS	1550134	1550247	1	+	114	hypothetical protein	- none -	D23_1c1664
fig 66666666.609 66.peg.1670	CDS	1553775	1553659	-3	-	117	hypothetical protein	- none -	D23_1c1668
fig 66666666.609	CDS	1573150	1573001	-1	-	150	hypothetical protein	- none -	D23_1c1686

66.peg.1638	1575055	1574942	-1	-	114	hypothetical protein	- none -	D23_1c1692
fig 6666666.609 66.peg.1693	1599507	1599391	-3	-	117	hypothetical protein	- none -	D23_1c1716
fig 6666666.609 66.peg.1718	1611922	1612056	1	+	135	hypothetical protein	- none -	D23_1c1727
fig 6666666.609 66.peg.1729	1614522	1614355	-3	-	168	hypothetical protein	- none -	D23_1c1731
fig 6666666.609 66.peg.1734	1614827	1614699	-2	-	129	hypothetical protein	- none -	D23_1c1732
fig 6666666.609 66.peg.1735	1618427	1618293	-2	-	135	hypothetical protein	- none -	D23_1c1738
fig 6666666.609 66.peg.1740	1642217	1642050	-2	-	168	hypothetical protein	- none -	D23_1c1761
fig 6666666.609 66.peg.1767	1651568	1651446	-2	-	123	hypothetical protein	- none -	D23_1c1772
fig 6666666.609 66.peg.1777	1662168	1662344	3	+	177	hypothetical protein	- none -	D23_1c1776
fig 6666666.609 66.peg.1781	1663286	1663158	-2	-	129	hypothetical protein	- none -	D23_1c1778
fig 6666666.609 66.peg.1783	1668254	1668093	-2	-	162	hypothetical protein	- none -	D23_1c1785
fig 6666666.609 66.peg.1790	1672280	1672017	-2	-	264	SSU ribosomal protein S20p	- none -	D23_1c1789
fig 6666666.609 66.peg.1794	1672473	1672634	3	+	162	hypothetical protein	- none -	D23_1c1790
fig 6666666.609 66.peg.1795	1685374	1685249	-1	-	126	hypothetical protein	- none -	D23_1c1808
fig 6666666.609 66.peg.1810	1727480	1727641	2	+	162	hypothetical protein	- none -	D23_1c1855

fig 666666.609 66.peg.1853	CDS	1727757	1727873	3	+	117	hypothetical protein	- none -	D23_1c1856
fig 666666.609 66.peg.1883	CDS	1757073	1757201	3	+	129	hypothetical protein	- none -	D23_1c1885
fig 666666.609 66.peg.1908	CDS	1784806	1784651	-1	-	156	hypothetical protein	- none -	D23_1c1913
fig 666666.609 66.peg.1952	CDS	1824892	1824704	-1	-	189	hypothetical protein	- none -	D23_1c1959
fig 666666.609 66.peg.1961	CDS	1830108	1830230	3	+	123	hypothetical protein	- none -	D23_1c1968
fig 666666.609 66.peg.1964	CDS	1831759	1831616	-1	-	144	hypothetical protein	- none -	D23_1c1971
fig 666666.609 66.peg.1969	CDS	1836022	1836150	1	+	129	hypothetical protein	- none -	D23_1c1975
fig 666666.609 66.peg.1978	CDS	1845616	1845783	1	+	168	hypothetical protein	- none -	D23_1c1984
fig 666666.609 66.peg.1981	CDS	1848084	1848206	3	+	123	hypothetical protein	- none -	D23_1c1987
fig 666666.609 66.peg.1990	CDS	1855948	1855835	-1	-	114	hypothetical protein	- none -	D23_1c1994
fig 666666.609 66.peg.1992	CDS	1856572	1856703	1	+	132	hypothetical protein	- none -	D23_1c1996
fig 666666.609 66.peg.1995	CDS	1860536	1860667	2	+	132	hypothetical protein	- none -	D23_1c1999
fig 666666.609 66.peg.2006	CDS	1870373	1870546	2	+	174	hypothetical protein	- none -	D23_1c2008
fig 666666.609 66.peg.2013	CDS	1877265	1877122	-3	-	144	hypothetical protein	- none -	D23_1c2015
fig 666666.609 66.peg.2016	CDS	1881044	1880847	-2	-	198	hypothetical protein	- none -	D23_1c2018
fig 666666.609	CDS	1888279	1888148	-1	-	132	hypothetical protein	- none -	D23_1c2028

[illegible]

fig 666666666.609 66.peg.2240	CDS	2076577	2076407	-1	-	171	hypothetical protein	- none -	D23_1c2234
fig 666666666.609 66.peg.2247	CDS	2083232	2083038	-2	-	195	hypothetical protein	- none -	D23_1c2241
fig 666666666.609 66.peg.2251	CDS	2086807	2086661	-1	-	147	hypothetical protein	- none -	D23_1c2244
fig 666666666.609 66.peg.2262	CDS	2096035	2096187	1	+	153	hypothetical protein	- none -	D23_1c2255
fig 666666666.609 66.peg.2282	CDS	2116748	2116897	2	+	150	hypothetical protein	- none -	D23_1c2276
fig 666666666.609 66.peg.2291	CDS	2126813	2126977	2	+	165	hypothetical protein	- none -	D23_1c2287
fig 666666666.609 66.peg.2293	CDS	2128367	2128218	-2	-	150	hypothetical protein	- none -	D23_1c2289
fig 666666666.609 66.peg.2316	CDS	2149225	2149353	1	+	129	hypothetical protein	- none -	D23_1c2312
fig 666666666.609 66.peg.2332	CDS	2168192	2168052	-2	-	141	hypothetical protein	- none -	D23_1c2326
fig 666666666.609 66.peg.2333	CDS	2168473	2168640	1	+	168	hypothetical protein	- none -	D23_1c2327
fig 666666666.609 66.peg.2339	CDS	2172852	2172977	3	+	126	hypothetical protein	- none -	D23_1c2334
fig 666666666.609 66.peg.2351	CDS	2184641	2184453	-2	-	189	hypothetical protein	- none -	D23_1c2346
fig 666666666.609 66.peg.2369	CDS	2201167	2201039	-1	-	129	hypothetical protein	- none -	D23_1c2365
fig 666666666.609 66.peg.2397	CDS	2225287	2225162	-1	-	126	hypothetical protein	- none -	D23_1c2390
fig 666666666.609 66.peg.2401	CDS	2229650	2229480	-2	-	171	hypothetical protein	- none -	D23_1c2393
fig 666666666.609	CDS	2233192	2233344	1	+	153	hypothetical protein	- none -	D23_1c2397

[illegible]

Figure 8

Unique D23 genes with no assigned ORF number: 219 (of which 180 are below 200 bp)

Feature.ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23Gbklid
fig 666666.609.66.peg.111	CDS	108515	108628	2	+	114	hypothetical protein	- none -	NA
fig 666666.609.66.peg.353	CDS	326760	326647	-3	-	114	hypothetical protein	- none -	NA
fig 666666.609.66.peg.404	CDS	375917	375804	-2	-	114	hypothetical protein	- none -	NA
fig 666666.609.66.peg.432	CDS	397539	397426	-3	-	114	hypothetical protein	- none -	NA
fig 666666.609.66.peg.509	CDS	474482	474369	-2	-	114	hypothetical protein	- none -	NA
fig 666666.609.66.peg.589	CDS	543008	543121	2	+	114	hypothetical protein	- none -	NA
fig 666666.609.66.peg.838	CDS	770105	770218	2	+	114	hypothetical protein	- none -	NA
fig 666666.609.66.peg.923	CDS	855159	855046	-3	-	114	hypothetical protein	- none -	NA
fig 666666.609.66.peg.943	CDS	873193	873306	1	+	114	hypothetical protein	- none -	NA
fig 666666.609.66.peg.1081	CDS	1010260	1010373	1	+	114	hypothetical protein	- none -	NA
fig 666666.609.66.peg.1333	CDS	1235666	1235779	2	+	114	hypothetical protein	- none -	NA

fig 66666666.609 66.peg.1414	CDS	1326350	1326463	2	+	114	hypothetical protein	- none -	NA
fig 66666666.609 66.peg.1695	CDS	1575540	1575653	3	+	114	hypothetical protein	- none -	NA
fig 66666666.609 66.peg.1875	CDS	1749599	1749486	-2	-	114	Mobile element protein	- none -	NA
fig 66666666.609 66.peg.2082	CDS	1936491	1936378	-3	-	114	hypothetical protein	- none -	NA
fig 66666666.609 66.peg.2205	CDS	2047715	2047828	2	+	114	hypothetical protein	- none -	NA
fig 66666666.609 66.peg.2370	CDS	2201390	2201277	-2	-	114	hypothetical protein	- none -	NA
fig 66666666.609 66.peg.2413	CDS	2236764	2236651	-3	-	114	hypothetical protein	- none -	NA
fig 66666666.609 66.peg.2577	CDS	2379043	2378930	-1	-	114	hypothetical protein	- none -	NA
fig 66666666.609 66.peg.421	CDS	392634	392518	-3	-	117	hypothetical protein	- none -	NA
fig 66666666.609 66.peg.431	CDS	397234	397350	1	+	117	hypothetical protein	- none -	NA
fig 66666666.609 66.peg.1104	CDS	1033773	1033889	3	+	117	hypothetical protein	- none -	NA
fig 66666666.609 66.peg.1856	CDS	1730271	1730155	-3	-	117	hypothetical protein	- none -	NA
fig 66666666.609 66.peg.2044	CDS	1907891	1907775	-2	-	117	hypothetical protein	- none -	NA
fig 66666666.609 66.peg.2046	CDS	1908400	1908516	1	+	117	hypothetical protein	- none -	NA
fig 66666666.609 66.peg.2250	CDS	2086632	2086516	-3	-	117	hypothetical protein	- none -	NA
fig 66666666.609	CDS	2209362	2209246	-3	-	117	hypothetical protein	- none -	NA

66.peg.2378	CDS	31960	31841	-1	-	120	hypothetical protein	- none -	NA
fig 66666666.609.66.peg.28	CDS	496598	496717	2	+	120	hypothetical protein	- none -	NA
fig 66666666.609.66.peg.536	CDS	794273	794392	2	+	120	hypothetical protein	- none -	NA
fig 66666666.609.66.peg.866	CDS	956342	956223	-2	-	120	hypothetical protein	- none -	NA
fig 66666666.609.66.peg.1031	CDS	1152349	1152230	-1	-	120	hypothetical protein	- none -	NA
fig 66666666.609.66.peg.1238	CDS	1184453	1184572	2	+	120	hypothetical protein	- none -	NA
fig 66666666.609.66.peg.1279	CDS	1462529	1462410	-2	-	120	hypothetical protein	- none -	NA
fig 66666666.609.66.peg.1566	CDS	1492079	1491960	-2	-	120	hypothetical protein	- none -	NA
fig 66666666.609.66.peg.1604	CDS	1495100	1495219	2	+	120	hypothetical protein	- none -	NA
fig 66666666.609.66.peg.1610	CDS	1820165	1820046	-2	-	120	hypothetical protein	- none -	NA
fig 66666666.609.66.peg.1938	CDS	1820392	1820511	1	+	120	hypothetical protein	- none -	NA
fig 66666666.609.66.peg.1940	CDS	2040715	2040596	-1	-	120	hypothetical protein	- none -	NA
fig 66666666.609.66.peg.2194	CDS	2069192	2069073	-2	-	120	hypothetical protein	- none -	NA
fig 66666666.609.66.peg.2229	CDS	2355482	2355601	2	+	120	hypothetical protein	- none -	NA
fig 66666666.609.66.peg.2542	CDS	444630	444508	-3	-	123	hypothetical protein	- none -	NA
fig 66666666.609.66.peg.477	CDS								

fig 6666666.609 66.peg.634	CDS	579242	579364	2	+	123	hypothetical protein	- none -	NA
fig 6666666.609 66.peg.927	CDS	859107	859229	3	+	123	hypothetical protein	- none -	NA
fig 6666666.609 66.peg.1015	CDS	941818	941940	1	+	123	hypothetical protein	- none -	NA
fig 6666666.609 66.peg.1388	CDS	1298114	1298236	2	+	123	hypothetical protein	- none -	NA
fig 6666666.609 66.peg.1755	CDS	1633195	1633073	-1	-	123	hypothetical protein	- none -	NA
fig 6666666.609 66.peg.2224	CDS	2067753	2067631	-3	-	123	hypothetical protein	- none -	NA
fig 6666666.609 66.peg.2271	CDS	2103666	2103788	3	+	123	hypothetical protein	- none -	NA
fig 6666666.609 66.peg.2336	CDS	2170340	2170462	2	+	123	hypothetical protein	- none -	NA
fig 6666666.609 66.peg.2415	CDS	2237130	2237008	-3	-	123	Hydroxycylglutathione hydrolase (EC 3.1.2.6)	CBSS- 228410.1.pe g.134; CBSS- 342610.3.pe g.1536; Glutathione : Non-redox reactions; Methylglyox al Metabolism	NA
fig 6666666.609 66.peg.82	CDS	79801	79676	-1	-	126	hypothetical protein	- none -	NA
fig 6666666.609	CDS	541998	542123	3	+	126	hypothetical protein	- none -	NA

[illegible]

fig 6666666.609 66.peg.1730	CDS	1612183	1612311	1	+	129	hypothetical protein	- none -	NA
fig 6666666.609 66.peg.2328	CDS	2162434	2162306	-1	-	129	hypothetical protein	- none -	NA
fig 6666666.609 66.peg.122	CDS	124194	124063	-3	-	132	hypothetical protein	- none -	NA
fig 6666666.609 66.peg.154	CDS	155514	155645	3	+	132	hypothetical protein	- none -	NA
fig 6666666.609 66.peg.874	CDS	801546	801677	3	+	132	hypothetical protein	- none -	NA
fig 6666666.609 66.peg.894	CDS	823808	823939	2	+	132	hypothetical protein	- none -	NA
fig 6666666.609 66.peg.960	CDS	886681	886550	-1	-	132	hypothetical protein	- none -	NA
fig 6666666.609 66.peg.1188	CDS	1108861	1108730	-1	-	132	hypothetical protein	- none -	NA
fig 6666666.609 66.peg.1578	CDS	1468706	1468837	2	+	132	hypothetical protein	- none -	NA
fig 6666666.609 66.peg.1950	CDS	1824304	1824435	1	+	132	hypothetical protein	- none -	NA
fig 6666666.609 66.peg.1983	CDS	1850338	1850469	1	+	132	hypothetical protein	- none -	NA
fig 6666666.609 66.peg.2066	CDS	1923319	1923450	1	+	132	hypothetical protein	- none -	NA
fig 6666666.609 66.peg.2445	CDS	2263603	2263472	-1	-	132	hypothetical protein	- none -	NA
fig 6666666.609 66.peg.2570	CDS	2376200	2376331	2	+	132	hypothetical protein	- none -	NA
fig 6666666.609 66.peg.296	CDS	267656	267790	2	+	135	hypothetical protein	- none -	NA
fig 6666666.609	CDS	1016149	1016015	-1	-	135	hypothetical protein	- none -	NA

[illegible]

fig 666666666.609 66.peg.1867	CDS	1739233	1739370	1	+	138	Error-prone, lesion bypass DNA polymerase V (UmuC)	- none -	NA
fig 666666666.609 66.peg.2117	CDS	1966387	1966250	-1	-	138	hypothetical protein	- none -	NA
fig 666666666.609 66.peg.2227	CDS	2068598	2068735	2	+	138	hypothetical protein	- none -	NA
fig 666666666.609 66.peg.163	CDS	160970	160830	-2	-	141	hypothetical protein	- none -	NA
fig 666666666.609 66.peg.173	CDS	168825	168965	3	+	141	hypothetical protein	- none -	NA
fig 666666666.609 66.peg.1286	CDS	1189481	1189621	2	+	141	hypothetical protein	- none -	NA
fig 666666666.609 66.peg.1416	CDS	1328534	1328674	2	+	141	hypothetical protein	- none -	NA
fig 666666666.609 66.peg.1422	CDS	1330733	1330873	2	+	141	hypothetical protein	- none -	NA
fig 666666666.609 66.peg.1759	CDS	1635699	1635839	3	+	141	hypothetical protein	- none -	NA
fig 666666666.609 66.peg.1775	CDS	1649712	1649572	-3	-	141	Mobile element protein	- none -	NA
fig 666666666.609 66.peg.1984	CDS	1850712	1850572	-3	-	141	hypothetical protein	- none -	NA
fig 666666666.609 66.peg.2003	CDS	1866906	1867046	3	+	141	hypothetical protein	- none -	NA
fig 666666666.609 66.peg.2080	CDS	1935668	1935808	2	+	141	hypothetical protein	- none -	NA
fig 666666666.609 66.peg.2487	CDS	2301272	2301132	-2	-	141	hypothetical protein	- none -	NA
fig 666666666.609 66.peg.369	CDS	342367	342224	-1	-	144	hypothetical protein	- none -	NA
fig 666666666.609	CDS	579509	579652	2	+	144	hypothetical protein	- none -	NA

66.peg.635		928837	928980	1	+		144		hypothetical protein	- none -		NA
fig 6666666.609.66.peg.1001	CDS											
fig 6666666.609.66.peg.1642	CDS	1524746	1524603	-2	-		144		hypothetical protein	- none -		NA
fig 6666666.609.66.peg.2114	CDS	1962886	1963029	1	+		144		hypothetical protein	- none -		NA
fig 6666666.609.66.peg.2593	CDS	2391861	2392004	3	+		144		hypothetical protein	- none -		NA
fig 6666666.609.66.peg.205	CDS	197036	196890	-2	-		147		Integrase	- none -		NA
fig 6666666.609.66.peg.284	CDS	260995	261141	1	+		147		hypothetical protein	- none -		NA
fig 6666666.609.66.peg.1885	CDS	1759287	1759141	-3	-		147		hypothetical protein	- none -		NA
fig 6666666.609.66.peg.2453	CDS	2272001	2271855	-2	-		147		hypothetical protein	- none -		NA
fig 6666666.609.66.peg.2620	CDS	2417974	2417828	-1	-		147		hypothetical protein	- none -		NA
fig 6666666.609.66.peg.2623	CDS	2420545	2420399	-1	-		147		hypothetical protein	- none -		NA
fig 6666666.609.66.peg.108	CDS	106117	106266	1	+		150		hypothetical protein	- none -		NA
fig 6666666.609.66.peg.813	CDS	746234	746085	-2	-		150		hypothetical protein	- none -		NA
fig 6666666.609.66.peg.1296	CDS	1205130	1205279	3	+		150		hypothetical protein	- none -		NA
fig 6666666.609.66.peg.1334	CDS	1235992	1235843	-1	-		150		hypothetical protein	- none -		NA
fig 6666666.609.66.peg.2300	CDS	2134634	2134485	-2	-		150		hypothetical protein	- none -		NA

fig 6666666.609 66.peg.35	CDS	36062	35910	-2	-	153	hypothetical protein	- none -	NA
fig 6666666.609 66.peg.1312	CDS	1215961	1216113	1	+	153	hypothetical protein	- none -	NA
fig 6666666.609 66.peg.2143	CDS	1990257	1990105	-3	-	153	hypothetical protein	- none -	NA
fig 6666666.609 66.peg.359	CDS	335237	335392	2	+	156	hypothetical protein	- none -	NA
fig 6666666.609 66.peg.546	CDS	506105	505950	-2	-	156	FIG00858972: hypothetical protein	- none -	NA
fig 6666666.609 66.peg.1174	CDS	1095978	1095823	-3	-	156	Mobile element protein	- none -	NA
fig 6666666.609 66.peg.1384	CDS	1296435	1296280	-3	-	156	hypothetical protein	- none -	NA
fig 6666666.609 66.peg.2575	CDS	2378324	2378169	-2	-	156	hypothetical protein	- none -	NA
fig 6666666.609 66.peg.89	CDS	85217	85059	-2	-	159	hypothetical protein	- none -	NA
fig 6666666.609 66.peg.556	CDS	516983	516825	-2	-	159	hypothetical protein	- none -	NA
fig 6666666.609 66.peg.562	CDS	523304	523462	2	+	159	hypothetical protein	- none -	NA
fig 6666666.609 66.peg.1099	CDS	1029431	1029589	2	+	159	hypothetical protein	- none -	NA
fig 6666666.609 66.peg.1571	CDS	1466563	1466405	-1	-	159	hypothetical protein	- none -	NA
fig 6666666.609 66.peg.613	CDS	562226	562065	-2	-	162	hypothetical protein	- none -	NA
fig 6666666.609 66.peg.886	CDS	814562	814401	-2	-	162	hypothetical protein	- none -	NA
fig 6666666.609	CDS	1043720	1043559	-2	-	162	hypothetical protein	- none -	NA

66.peg.1114	CDS	265661	265825	2	+				FIG00856904: hypothetical protein	- none -	NA
fig 6666666.609 66.peg.291	CDS	1934012	1933848	-2	-				hypothetical protein	- none -	NA
fig 6666666.609 66.peg.2077	CDS	568234	568401	1	+				Methyltransferase (EC 2.1.1.-)	- none -	NA
fig 6666666.609 66.peg.620	CDS	2149354	2149521	1	+				Mobile element protein	- none -	NA
fig 6666666.609 66.peg.2317	CDS	2241817	2241650	-1	-				hypothetical protein	- none -	NA
fig 6666666.609 66.peg.856	CDS	785317	785147	-1	-				hypothetical protein	- none -	NA
fig 6666666.609 66.peg.1487	CDS	1401876	1401706	-3	-				FIG00858878: hypothetical protein	- none -	NA
fig 6666666.609 66.peg.2286	CDS	2119703	2119873	2	+				hypothetical protein	- none -	NA
fig 6666666.609 66.peg.2707	CDS	2514260	2514090	-2	-				hypothetical protein	- none -	NA
fig 6666666.609 66.peg.862	CDS	790885	790712	-1	-				hypothetical protein	- none -	NA
fig 6666666.609 66.peg.1450	CDS	1360555	1360382	-1	-				Mobile element protein	- none -	NA
fig 6666666.609 66.peg.1572	CDS	1466783	1466956	2	+				hypothetical protein	- none -	NA
fig 6666666.609 66.peg.1608	CDS	1494659	1494835	2	+				hypothetical protein	- none -	NA
fig 6666666.609 66.peg.2416	CDS	2237303	2237127	-2	-				hypothetical protein	- none -	NA
fig 6666666.609 66.peg.2668	CDS	2479940	2479764	-2	-				hypothetical protein	- none -	NA

fig 6666666.609 66.peg.748	CDS	694556	694377	-2	-	180	hypothetical protein	- none -	NA
fig 6666666.609 66.peg.1667	CDS	1550475	1550296	-3	-	180	hypothetical protein	- none -	NA
fig 6666666.609 66.peg.2116	CDS	1965779	1965600	-2	-	180	hypothetical protein	- none -	NA
fig 6666666.609 66.peg.2144	CDS	1990362	1990541	3	+	180	Mobile element protein	- none -	NA
fig 6666666.609 66.peg.2400	CDS	2229074	2228895	-2	-	180	hypothetical protein	- none -	NA
fig 6666666.609 66.peg.287	CDS	262264	262446	1	+	183	hypothetical protein	- none -	NA
fig 6666666.609 66.peg.422	CDS	392649	392831	3	+	183	hypothetical protein	- none -	NA
fig 6666666.609 66.peg.1562	CDS	1459862	1460044	2	+	183	hypothetical protein	- none -	NA
fig 6666666.609 66.peg.2388	CDS	2218996	2219178	1	+	183	hypothetical protein	- none -	NA
fig 6666666.609 66.peg.338	CDS	308508	308323	-3	-	186	hypothetical protein	- none -	NA
fig 6666666.609 66.peg.1842	CDS	1720173	1719988	-3	-	186	hypothetical protein	- none -	NA
fig 6666666.609 66.peg.286	CDS	262147	261959	-1	-	189	hypothetical protein	- none -	NA
fig 6666666.609 66.peg.1202	CDS	1125016	1124828	-1	-	189	hypothetical protein	- none -	NA
fig 6666666.609 66.peg.1304	CDS	1209919	1209731	-1	-	189	FIG00858878: hypothetical protein	- none -	NA
fig 6666666.609 66.peg.2395	CDS	2224520	2224332	-2	-	189	hypothetical protein	- none -	NA
fig 6666666.609	CDS	233011	233202	1	+	192	hypothetical protein	- none -	NA

66.peg.250				-2	-	192		hypothetical protein	- none -	NA
fig 66666666.609 66.peg.794	CDS	729662	729471	2	+	192		hypothetical protein	- none -	NA
fig 66666666.609 66.peg.1764	CDS	1639415	1639606	-3	-	195		hypothetical protein	- none -	NA
fig 66666666.609 66.peg.808	CDS	741018	740824	1	+	198		Mobile element protein	- none -	NA
fig 66666666.609 66.peg.276	CDS	255673	255870	1	+	198		hypothetical protein	- none -	NA
fig 66666666.609 66.peg.749	CDS	694591	694788	-1	-	198		hypothetical protein	- none -	NA
fig 66666666.609 66.peg.2718	CDS	2523733	2523536	1	+	201		Mobile element protein	- none -	NA
fig 66666666.609 66.peg.217	CDS	631759	631959	1	+	201		FIG00859622: hypothetical protein	- none -	NA
fig 66666666.609 66.peg.208	CDS	199189	199392	1	+	204		Cold shock protein CspA	Cold shock, CspA family of proteins	NA
fig 66666666.609 66.peg.380	CDS	348880	348677	-1	-	204		SSU ribosomal protein S16p	- none -	NA
fig 66666666.609 66.peg.2149	CDS	1995706	1995503	-1	-	204		dTDP-4-dehydrothiamine reductase (EC 1.1.1.133)	Rhamnose containing glycans; dTDP- rhamnose synthesis	NA
fig 66666666.609 66.peg.1148	CDS	1073050	1073259	1	+	210		hypothetical protein	- none -	NA
fig 66666666.609	CDS	2281343	2281134	-2	-	210		hypothetical protein	- none -	NA

66.peg.2464	CDS	484104	484316	3	+	213	LSU ribosomal protein 129p (L35e)	- none -	NA
fig 66666666.609.66.peg.522	CDS	1395113	1394901	-2	-	213	hypothetical protein	- none -	NA
fig 66666666.609.66.peg.1481	CDS	265809	266024	3	+	216	hypothetical protein	- none -	NA
fig 66666666.609.66.peg.511	CDS	475254	475469	3	+	216	SSU ribosomal protein S12p (S23e)	Mycobacterium virulence operon involved in protein synthesis (SSU ribosomal proteins); Ribosomal protein S12p Asp methylthiotransferase	NA
fig 66666666.609.66.peg.529	CDS	490129	490347	1	+	219	Translation initiation factor 1	Translation initiation factors bacterial	NA
fig 66666666.609.66.peg.2374	CDS	2207422	2207640	1	+	219	hypothetical protein	- none -	NA
fig 66666666.609.66.peg.2105	CDS	1956719	1956940	2	+	222	hypothetical protein	- none -	NA
fig 66666666.609.66.peg.2105	CDS	2502317	2502538	2	+	222	hypothetical protein	- none -	NA

[illegible]

fig 666666666.609 66.peg.2579	CDS	2380195	2380446	1	+	252	Mobile element protein	- none -	NA
fig 666666666.609 66.peg.523	CDS	484294	484548	1	+	255	SSU ribosomal protein S17p (S11e)	- none -	NA
fig 666666666.609 66.peg.1022	CDS	948522	948782	3	+	261	Stringent starvation protein B	Carbon Starvation	NA
fig 666666666.609 66.peg.1957	CDS	1826664	1826936	3	+	273	hypothetical protein	- none -	NA
fig 666666666.609 66.peg.2273	CDS	2108052	2108336	3	+	285	putative DNA transport competence protein	- none -	NA
fig 666666666.609 66.peg.46	CDS	46807	46481	-1	-	327	Conserved domain protein	- none -	NA
fig 666666666.609 66.peg.2109	CDS	1959886	1960239	1	+	354	hypothetical protein	- none -	NA
fig 666666666.609 66.peg.2429	CDS	2248841	2248488	-2	-	354	hypothetical protein	- none -	NA
fig 666666666.609 66.peg.1579	CDS	1468834	1469334	1	+	501	hypothetical protein	- none -	NA
fig 666666666.609 66.peg.293	CDS	266114	266689	2	+	576	hypothetical protein	- none -	NA

Figure 9

Neut_0006, Neut_0007, Neut_0008, Neut_0009, Neut_0019, Neut_0020, Neut_0023, Neut_0028, Neut_0031, Neut_0032, Neut_0033, Neut_0034, Neut_0035, Neut_0036, Neut_0039, Neut_0042, Neut_0044, Neut_0046, Neut_0048, Neut_0052, Neut_0053, Neut_0069, Neut_0070, Neut_0071, Neut_0072, Neut_0073, Neut_0074, Neut_0075, Neut_0076, Neut_0077, Neut_0080, Neut_0082, Neut_0083, Neut_0085, Neut_0086, Neut_0087, Neut_0088, Neut_0089, Neut_0090, Neut_0091, Neut_0093, Neut_0094, Neut_0096, Neut_0097, Neut_0098, Neut_0099, Neut_0100, Neut_0101, Neut_0102, Neut_0103, Neut_0104, Neut_0106, Neut_0107, Neut_0109, Neut_0110, Neut_0111, Neut_0112, Neut_0113, Neut_0114, Neut_0115, Neut_0118, Neut_0119, Neut_0120, Neut_0121, Neut_0125, Neut_0126, Neut_0154, Neut_0155, Neut_0157, Neut_0158, Neut_0161, Neut_0354, Neut_0358, Neut_0359, Neut_0365, Neut_0366, Neut_0368, Neut_0538, Neut_0539, Neut_0577, Neut_0581, Neut_0585, Neut_0617, Neut_0627, Neut_0632, Neut_0633, Neut_0636, Neut_0665, Neut_0666, Neut_0667, Neut_0668, Neut_0669, Neut_0670, Neut_0671, Neut_0672, Neut_0673, Neut_0674, Neut_0675, Neut_0676, Neut_0677, Neut_0678, Neut_0679, Neut_0680, Neut_0693, Neut_0737, Neut_0738, Neut_0880, Neut_0885, Neut_0904, Neut_0984, Neut_1012, Neut_1013, Neut_1014, Neut_1015, Neut_1016, Neut_1057, Neut_1058, Neut_1060, Neut_1062, Neut_1065, Neut_1066, Neut_1068, Neut_1073, Neut_1075, Neut_1084, Neut_1115, Neut_1120, Neut_1121, Neut_1122, Neut_1123, Neut_1124, Neut_1125, Neut_1305, Neut_1306, Neut_1307, Neut_1308, Neut_1310, Neut_1311, Neut_1312, Neut_1314, Neut_1319, Neut_1320, Neut_1322, Neut_1324, Neut_1327, Neut_1334, Neut_1335, Neut_1367, Neut_1369, Neut_1370, Neut_1372, Neut_1399, Neut_1414, Neut_1476, Neut_1482, Neut_1484, Neut_1485, Neut_1487, Neut_1490, Neut_1597, Neut_1674, Neut_1725, Neut_1728, Neut_1731, Neut_1733, Neut_1755, Neut_1758, Neut_1759, Neut_1818, Neut_1820, Neut_1846, Neut_1919, Neut_1948, Neut_1980, Neut_2111, Neut_2114, Neut_2150, Neut_2153, Neut_2200, Neut_2201, Neut_2212, Neut_2215, Neut_2216, Neut_2217, Neut_2218, Neut_2219, Neut_2220, Neut_2221, Neut_2222, Neut_2223, Neut_2224, Neut_2225, Neut_2226, Neut_2227, Neut_2228, Neut_2229, Neut_2230, Neut_2231, Neut_2252, Neut_2253, Neut_2254, Neut_2255, Neut_2256, Neut_2257, Neut_2258, Neut_2259, Neut_2260, Neut_2261, Neut_2262, Neut_2263, Neut_2264, Neut_2265, Neut_2266, Neut_2352, Neut_2353, Neut_2354, Neut_2356, Neut_2411, Neut_2412, Neut_2416, Neut_2504, Neut_2505, Neut_2533, Neut_2533a, Neut_2534, Neut_2535, Neut_2536, Neut_2540, Neut_2541, Neut_2543, Neut_2547, Neut_2549, Neut_2551, Neut_2554, Neut_2558, Neut_2559, Neut_2560, Neut_2561, Neut_2562, Neut_2564, Neut_2565, Neut_2566, Neut_2568, Neut_2569, Neut_2570, Neut_2571, Neut_2572, Neut_2573, Neut_2574, Neut_2576, Neut_2578, Neut_2579, Neut_2582, Neut_2583, Neut_2585, Neut_2588, Neut_2589, Neut_2590, Neut_2593, Neut_2595, Neut_2596, Neut_2597, Neut_2598, Neut_2599, Neut_2602, Neut_2603, Neut_2607, Neut_2609, Neut_2612, Neut_2613, Neut_2614, Neut_2617, Neut_2618, Neut_2619, Neut_2620, Neut_2621, Neut_2623, Neut_2625, Neut_2626, Neut_2627, Neut_2628, Neut_2629, Neut_2630, Neut_2631, Neut_2632, Neut_2633, Neut_2634, Neut_2635, Neut_2638, Neut_2639, Neut_2641, rplN, rplQ, rplR, rplV, rpmD, rpmI, rpsH, rpsM, rpsN.

Figure 10

CLUSTAL 2.1 multiple sequence alignment

```
D23-amoA1      VSIPTREEILKAAMPPEAVHMSRLIDAVYFPILVLLVGTYYHMHFMLLAGDWDFWMDWK 60
D23-amoA2      VSIPTREEILKAAMPPEAVHMSRLIDAVYFPILVLLVGTYYHMHFMLLAGDWDFWMDWK 60
C91-amoA1      MSIPTREEILKAAMPPEAVHMSRLIDAVYFPILVLLVGTYYHMHFMLLAGDWDFWMDWK 60
C91-amoA2      MSIPTREEILKAAMPPEAVHMSRLIDAVYFPILVLLVGTYYHMHFMLLAGDWDFWMDWK 60
                ;*****
D23-amoA1      DRQWPFVVTPIVGITYCSAIMYYLWVNYRQPFGATLCVVCLLIGEMLTRYWGFYWWSHYP 120
D23-amoA2      DRQWPFVVTPIVGITYCSAIMYYLWVNYRQPFGATLCVVCLLIGEMLTRYWGFYWWSHYP 120
C91-amoA1      DRQWPFVVTPIVGITYCSAIMYYLWVNYRQPFGATLCVVCLLIGEMLTRYWGFYWWSHYP 120
C91-amoA2      DRQWPFVVTPIVGITYCSAIMYYLWVNYRQPFGATLCVVCLLIGEMLTRYWGFYWWSHYP 120
                *****
D23-amoA1      LNFVTPGIMLPGALMLDFTMYLTRNNLVTALVGGGFEGLLFYPCNWAIFGPTHLPIVEG 180
D23-amoA2      LNFVTPGIMLPGALMLDFTMYLTRNNLVTALVGGGFEGLLFYPCNWAIFGPTHLPIVEG 180
C91-amoA1      LNFVTPGIMLPGALMLDFTMYLTRNNLVTALVGGGFEGLLFYPCNWAIFGPTHLPIVEG 180
C91-amoA2      LNFVTPGIMLPGALMLDFTMYLTRNNLVTALVGGGFEGLLFYPCNWAIFGPTHLPIVEG 180
                *****;*****
D23-amoA1      TLLSMADYMGCHLYVRTCTPEYVVRHIEQGSLSRTFGCHTTVIAAFFAFAVSNLMFAVWWYL 240
D23-amoA2      TLLSMADYMGCHLYVRTCTPEYVVRHIEQGSLSRTFGCHTTVIAAFFAFAVSNLMFAVWWYL 240
C91-amoA1      TLLSMADYMGCHLYVRTCTPEYVVRHIEQGSLSRTFGCHTTVIAAFFAFAVSNLMFAVWWYL 240
C91-amoA2      TLLSMADYMGCHLYVRTCTPEYVVRHIEQGSLSRTFGCHTTVIAAFFAFAVSNLMFAVWWYL 240
                *****
D23-amoA1      KVVCTAFFYVKCKRGRIVQRNDVTAFGEGFPEGIK 276
D23-amoA2      KVVCTAFFYVKCKRGRIVQRNDVTAFGEGFPEGIK 276
C91-amoA1      KVVCTAFFYVKCKRGRIVQRNDVTAFGEGFPEGIK 276
C91-amoA2      KVVCTAFFYVKCKRGRIVQRNDVTAFGEGFPEGIK 276
                *****
```

Figure 11

```

D23--amob1      MGLKRLIYKCGMGLCGGVAVYAAALTWVTLGQSTVAAGCEKCEFFLAKKTVQYQVNR 60
D23--amob2      MGLKRLIYKCGMGLCGGVAVYAAALTWVTLGQSTVAAGCEKCEFFLAKKTVQYQVNR 80
C91--amob1      MGLKRLIYKCGMGLCGGVAVYAAALTWVTLGQSTVAAGCEKCEFFLAKKTVQYQVNR 60
C91--amob2      MGLKRLIYKCGMGLCGGVAVYAAALTWVTLGQSTVAAGCEKCEFFLAKKTVQYQVNR 60
*****
D23--amob1      GPEVTVKVENAGITGKTFPLAEDWRAAARCPAFEFVWSPSPSVYVWLSVXNGHPEFISG 120
D23--amob2      GPEVTVKVENAGITGKTFPLAEDWRAAARCPAFEFVWSPSPSVYVWLSVXNGHPEFISG 120
C91--amob1      GPEVTVKVENAGITGKTFPLAEDWRAAARCPAFEFVWSPSPSVYVWLSVXNGHPEFISG 120
C91--amob2      GPEVTVKVENAGITGKTFPLAEDWRAAARCPAFEFVWSPSPSVYVWLSVXNGHPEFISG 120
*****
D23--amob1      PLOIGRQVAFVYQVQABRYSGHSHHMLNKNKAGFAGCAHMLTGSKQNTNPKLIT 180
D23--amob2      PLOIGRQVAFVYQVQABRYSGHSHHMLNKNKAGFAGCAHMLTGSKQNTNPKLIT 180
C91--amob1      PLOIGRQVAFVYQVQABRYSGHSHHMLNKNKAGFAGCAHMLTGSKQNTNPKLIT 180
C91--amob2      PLOIGRQVAFVYQVQABRYSGHSHHMLNKNKAGFAGCAHMLTGSKQNTNPKLIT 180
*****
D23--amob1      GETIDSETNLSNGITFNRILNMSIGITFNGIFVAREPFIKPSRVLAYKQKLLDPEKK 240
D23--amob2      GETIDSETNLSNGITFNRILNMSIGITFNGIFVAREPFIKPSRVLAYKQKLLDPEKK 240
C91--amob1      GETIDSETNLSNGITFNRILNMSIGITFNGIFVAREPFIKPSRVLAYKQKLLDPEKK 240
C91--amob2      GETIDSETNLSNGITFNRILNMSIGITFNGIFVAREPFIKPSRVLAYKQKLLDPEKK 240
*****
D23--amob1      ITWVLAILTLAIVNGVATSTHEVTPVTCAGQSKVAFIPVAPHPVAIKITDARYVFG 300
D23--amob2      ITWVLAILTLAIVNGVATSTHEVTPVTCAGQSKVAFIPVAPHPVAIKITDARYVFG 300
C91--amob1      ITWVLAILTLAIVNGVATSTHEVTPVTCAGQSKVAFIPVAPHPVAIKITDARYVFG 300
C91--amob2      ITWVLAILTLAIVNGVATSTHEVTPVTCAGQSKVAFIPVAPHPVAIKITDARYVFG 300
*****
D23--amob1      RALRVSMETVNGTPTVTFGEFTTAGIHFVNSGTGRVLEPQFRELAVAGLNFDDGAIQ 360
D23--amob2      RALRVSMETVNGTPTVTFGEFTTAGIHFVNSGTGRVLEPQFRELAVAGLNFDDGAIQ 360
C91--amob1      RALRVSMETVNGTPTVTFGEFTTAGIHFVNSGTGRVLEPQFRELAVAGLNFDDGAIQ 360
C91--amob2      RALRVSMETVNGTPTVTFGEFTTAGIHFVNSGTGRVLEPQFRELAVAGLNFDDGAIQ 360
*****
D23--amob1      PGETKQLRHEARDALMEYQRLMALLGDFSRFQGLNHSKGRHINHSIAGFVTEVTR 420
D23--amob2      PGETKQLRHEARDALMEYQRLMALLGDFSRFQGLNHSKGRHINHSIAGFVTEVTR 420
C91--amob1      PGETKQLRHEARDALMEYQRLMALLGDFSRFQGLNHSKGRHINHSIAGFVTEVTR 420
C91--amob2      PGETKQLRHEARDALMEYQRLMALLGDFSRFQGLNHSKGRHINHSIAGFVTEVTR 420
*****
D23--amob1      2. 421
D23--amob2      1. 421
C91--amob1      2. 421
C91--amob2      1. 421
*
```

Figure 12

```
C91~amoc1      MATTIQTSSASVS6SRCYDMSIWDKRYKFKLITMLIVAFPWVYQRYFAYSHQDMS 60
C91~amoc2      MATTIQTSSASVS6SRCYDMSIWDKRYKFKLITMLIVAFPWVYQRYFAYSHQDMS 60
D23~amoc2      MATTIQTSSASVS6SRCYDMSIWDKRYKFKLITMLIVAFPWVYQRYFAYSHQDMS 60
D23~amoc1      MATTIQTSSASVS6SRCYDMSIWDKRYKFKLITMLIVAFPWVYQRYFAYSHQDMS 60
*****
C91~amoc1      FEFDEVHNGIWEVHMAIMPLFALVYTWQWIMXTDTEZLNLDPELEIKRYFYIMMWLEY 120
C91~amoc2      FEFDEVHNGIWEVHMAIMPLFALVYTWQWIMXTDTEZLNLDPELEIKRYFYIMMWLEY 120
D23~amoc2      FEFDEVHNGIWEVHMAIMPLFALVYTWQWIMXTDTEZLNLDPELEIKRYFYIMMWLEY 120
D23~amoc1      FEFDEVHNGIWEVHMAIMPLFALVYTWQWIMXTDTEZLNLDPELEIKRYFYIMMWLEY 120
*****
C91~amoc1      YIFGVYWGCSFPTTEQDASNHQVLIIRDSTFTTEHVVVYFGSFPNVIYGVATYLIYAMTRLP 180
C91~amoc2      YIFGVYWGCSFPTTEQDASNHQVLIIRDSTFTTEHVVVYFGSFPNVIYGVATYLIYAMTRLP 180
D23~amoc2      YIFGVYWGCSFPTTEQDASNHQVLIIRDSTFTTEHVVVYFGSFPNVIYGVATYLIYAMTRLP 180
D23~amoc1      YIFGVYWGCSFPTTEQDASNHQVLIIRDSTFTTEHVVVYFGSFPNVIYGVATYLIYAMTRLP 180
*****
C91~amoc1      LFHRSISFFPLVMAIAGFLMILPNVGLNMGHAPWFNHELFSAPLHNSFVVLGNAGLPQGG 240
C91~amoc2      LFHRSISFFPLVMAIAGFLMILPNVGLNMGHAPWFNHELFSAPLHNSFVVLGNAGLPQGG 240
D23~amoc2      LFHRSISFFPLVMAIAGFLMILPNVGLNMGHAPWFNHELFSAPLHNSFVVLGNAGLPQGG 240
D23~amoc1      LFHRSISFFPLVMAIAGFLMILPNVGLNMGHAPWFNHELFSAPLHNSFVVLGNAGLPQGG 240
*****
C91~amoc1      VAAQIITHYISMLTTPVYVNNQSEILNKKVA 271
C91~amoc2      VAAQIITHYISMLTTPVYVNNQSEILNKKVA 271
D23~amoc2      VAAQIITHYISMLTTPVYVNNQSEILNKKVA 271
D23~amoc1      VAAQIITHYISMLTTPVYVNNQSEILNKKVA 271
*****
```


Figure 13

C91~amoc3	MATNILKDKAACQVADKPTYDKSEWFDKYYKFGLLPILAVAVMWVYFORTYAYSHGMS	60
D23~amoc3	MATNILKDKAACQVADKPTYDKSEWFDKYYKFGLLPILAVAVMWVYFORTYAYSHGMS	60
	*****	*****
C91~amoc3	MEPEFDRIWMGLNRVQMAVLPLIALFTWGWLYKTRNTAEQLANLTPKQEIKRIFYFLMWL	120
D23~amoc3	MEPEFDRIWMGLNRVQMAVLPLIALFTWGWLYKTRNTAEQLANLTPKQEIKRIFYFLMWL	120
	*****,******	*****
C91~amoc3	GVYIFAVYWCSSFFTEQDASWHQVIIIRDTSTFTPSHIPLEFYCSFPVYIIMCVSMIYYANTR	180
D23~amoc3	GVYIFAVYWCSSFFTEQDASWHQVIIIRDTSTFTPSHIPLEFYCSFPVYIIMCVSMIYYANTR	180
	*****	*****
C91~amoc3	LPLYNKGWSFPLIMTVAGPLMSLPNVGLNEWGHAFWFMEELFSAPLHWGFVILAWAALFQ	240
D23~amoc3	LPLYNKGWSFPLIMTVAGPLMSLPNVGLNEWGHAFWFMEELFSAPLHWGFVILAWAALFQ	240
	*****	*****
C91~amoc3	GGLAVQIIARFSNLLDVEWNKQDRAILLDDVITAP	274
D23~amoc3	GGLAVQIIARFSNLLDVEWNKQDRAILLDDVITAP	274
	*****	*****

Fig. 14A

D23-hao2 MRLGEYLGMLLCAGLLIGPVQADISVPOETYEALKIDRSKATPKETDYLAVKRYKDP 60
D23-hao3 MRLGEYLGMLLCAGLLIGPVQADISVPOETYEALKIDRSKATPKETDYLAVKRYKDP 60
D23-hao1 MRLGEYLGMLLCAGLLIGPVQADISVPOETYEALKIDRSKATPKETDYLAVKRYKDP 60
C91-hao2 MRLGEYLGMLLCAGLLIGPVQADISVPOETYEALKIDRSKATPKETDYLAVKRYKDP 60
C91-hao3 MRLGEYLGMLLCAGLLIGPVQADISVPOETYEALKIDRSKATPKETDYLAVKRYKDP 60
C91-hao1 MRLGEYLGMLLCAGLLIGPVQADISVPOETYEALKIDRSKATPKETDYLAVKRYKDP 60

D23-hao2 AHGAGKGTMGDYWEPIALSIYMDPSTFYKPPVSPKKEIAERKDCVECHSDETFVWVRAWKR 120
D23-hao3 AHGAGKGTMGDYWEPIALSIYMDPSTFYKPPVSPKKEIAERKDCVECHSDETFVWVRAWKR 120
D23-hao1 AHGAGKGTMGDYWEPIALSIYMDPSTFYKPPVSPKKEIAERKDCVECHSDETFVWVRAWKR 120
C91-hao2 AHGAGKGTMGDYWEPIALSIYMDPNTFYKPPVSPKKEIAERKDCVECHSDETFVWVRAWKR 120
C91-hao3 AHGAGKGTMGDYWEPIALSIYMDPNTFYKPPVSPKKEIAERKDCVECHSDETFVWVRAWKR 120
C91-hao1 AHGAGKGTMGDYWEPIALSIYMDPNTFYKPPVSPKKEIAERKDCVECHSDETFVWVRAWKR 120

D23-hao2 STHANLKIIRNLKPEDPLFYKKGKLEEVENNLRSMLGKKEKALKEVGCIDCHVDINAKK 180
D23-hao3 STHANLKIIRNLKPEDPLFYKKGKLEEVENNLRSMLGKKEKALKEVGCIDCHVDINAKK 180
D23-hao1 STHANLKIIRNLKPEDPLFYKKGKLEEVENNLRSMLGKKEKALKEVGCIDCHVDINAKK 180
C91-hao2 STHANLKIIRNLKPEDPLFYKKGKLEEVENNLRSMLGKKEKALKEVGCIDCHVDINAKK 180
C91-hao3 STHANLKIIRNLKPEDPLFYKKGKLEEVENNLRSMLGKKEKALKEVGCIDCHVDINAKK 180
C91-hao1 STHANLKIIRNLKPEDPLFYKKGKLEEVENNLRSMLGKKEKALKEVGCIDCHVDINAKK 180

D23-hao2 KADHTKDVRMPTADVCGTCHLREFAERESERDTMIWPNQWQPDGRPSHALDYTANIEYTV 240
D23-hao3 KADHTKDVRMPTADVCGTCHLREFAERESERDTMIWPNQWQPDGRPSHALDYTANIEYTV 240
D23-hao1 KADHTKDVRMPTADVCGTCHLREFAERESERDTMIWPNQWQPDGRPSHALDYTANIEYTV 240
C91-hao2 KADHTKDVRMPTADVCGTCHLREFAERESERDTMIWPNQWQPDGRPSHALDYTANIEYTV 240
C91-hao3 KADHTKDVRMPTADVCGTCHLREFAERESERDTMIWPNQWQPDGRPSHALDYTANIEYTV 240
C91-hao1 KADHTKDVRMPTADVCGTCHLREFAERESERDTMIWPNQWQPDGRPSHALDYTANIEYTV 240

D23-hao2 WAAMPQREVAEGCTMCHTNQKNCNCHTRHEFSAAESRKEPEACATCHSGVDHNNYEAYIM 300
D23-hao3 WAAMPQREVAEGCTMCHTNQKNCNCHTRHEFSAAESRKEPEACATCHSGVDHNNYEAYIM 300
D23-hao1 WAAMPQREVAEGCTMCHTNQKNCNCHTRHEFSAAESRKEPEACATCHSGVDHNNYEAYIM 300
C91-hao2 WAAMPQREVAEGCTMCHTNQKNCNCHTRHEFSAAESRKEPEACATCHSGVDHNNYEAYIM 300
C91-hao3 WAAMPQREVAEGCTMCHTNQKNCNCHTRHEFSAAESRKEPEACATCHSGVDHNNYEAYIM 300
C91-hao1 WAAMPQREVAEGCTMCHTNQKNCNCHTRHEFSAAESRKEPEACATCHSGVDHNNYEAYIM 300

Fig. 14B

D23-hac2	SKHGKLAEMNRKNNWVRLKDAFSGGQGTAPTCACCHMEYEGSYTHNITRKTRWANYTF 360
D23-hac3	SKHGKLAEMNRKNNWVRLKDAFSGGQGTAPTCACCHMEYEGSYTHNITRKTRWANYTF 360
D23-hac1	SKHGKLAEMNRKNNWVRLKDAFSGGQGTAPTCACCHMEYEGSYTHNITRKTRWANYTF 360
C91-hac2	SKHGKLAEMNRKNNWVRLKDAFSGGQGTAPTCACCHMEYEGSYTHNITRKTRWANYTF 360
C91-hac3	SKHGKLAEMNRKNNWVRLKDAFSGGQGTAPTCACCHMEYEGSYTHNITRKTRWANYTF 360
C91-hac1	SKHGKLAEMNRKNNWVRLKDAFSGGQGTAPTCACCHMEYEGSYTHNITRKTRWANYTF 360

D23-hac2	VPGIAENITSDWSEARLDSWVVTCTQCHSERFASLYLDMDKSTLEGLAKYQERAAIVHK 420
D23-hac3	VPGIAENITSDWSEARLDSWVVTCTQCHSERFASLYLDMDKSTLEGLAKYQERAAIVHK 420
D23-hac1	VPGIAENITSDWSEARLDSWVVTCTQCHSERFASLYLDMDKSTLEGLAKYQERAAIVHK 420
C91-hac2	VPGIAENITSDWSEARLDSWVVTCTQCHSERFASLYLDMDKSTLEGLAKYQERAAIVHK 420
C91-hac3	VPGIAENITSDWSEARLDSWVVTCTQCHSERFASLYLDMDKSTLEGLAKYQERAAIVHK 420
C91-hac1	VPGIAENITSDWSEARLDSWVVTCTQCHSERFASLYLDMDKSTLEGLAKYQERAAIVHK 420

D23-hac2	MYEDGTLTGQKTNRPENPFAPKPGGIFTCQFWKGNPFASLELKVLENAENLAKMHVG 480
D23-hac3	MYEDGTLTGQKTNRPENPFAPKPGGIFTCQFWKGNPFASLELKVLENAENLAKMHVG 480
D23-hac1	MYEDGTLTGQKTNRPENPFAPKPGGIFTCQFWKGNPFASLELKVLENAENLAKMHVG 480
C91-hac2	MYEDGTLTGQKTNRPENPFAPKPGGIFTCQFWKGNPFASLELKVLENAENLAKMHVG 480
C91-hac3	MYEDGTLTGQKTNRPENPFAPKPGGIFTCQFWKGNPFASLELKVLENAENLAKMHVG 480
C91-hac1	MYEDGTLTGQKTNRPENPFAPKPGGIFTCQFWKGNPFASLELKVLENAENLAKMHVG 480

D23-hac2	LAHVFGGWTYTEGKPNRRAVVEIQDEYTKMQENTALQARVNLKGYTSLLDLKGAGE 540
D23-hac3	LAHVFGGWTYTEGKPNRRAVVEIQDEYTKMQENTALQARVNLKGYTSLLDLKGAGE 540
D23-hac1	LAHVFGGWTYTEGKPNRRAVVEIQDEYTKMQENTALQARVNLKGYTSLLDLKGAGE 540
C91-hac2	LAHVFGGWTYTEGKPNRRAVVEIQDEYTKMQENTALQARVNLKGYTSLLDLKGAGE 540
C91-hac3	LAHVFGGWTYTEGKPNRRAVVEIQDEYTKMQENTALQARVNLKGYTSLLDLKGAGE 540
C91-hac1	LAHVFGGWTYTEGKPNRRAVVEIQDEYTKMQENTALQARVNLKGYTSLLDLKGAGE 540

D23-hac2	KISLGIGGGMLLAGAIALICWRKXKQQA 570
D23-hac3	KISLGIGGGMLLAGAIALICWRKXKQQA 570
D23-hac1	KISLGIGGGMLLAGAIALICWRKXKQQA 570
C91-hac2	KISLGIGGGMLLAGAIALICWRKXKQQA 570
C91-hac3	KISLGIGGGMLLAGAIALICWRKXKQQA 570
C91-hac1	KISLGIGGGMLLAGAIALICWRKXKQQA 570

Figure 15

```
D23--cycA2      NKIIACGLVAAALFTLTSGQSLAADAFEGGRKKCSSCHKPQAQSWKHTAHAKAMESLKP 60
D23--cycA3      NKIIACGLVAAALFTLTSGQSLAADAFEGGRKKCSSCHKPQAQSWKHTAHAKAMESLKP 60
D23--cycA1      NKIIACGLVAAALFTLTSGQSLAADAFEGGRKKCSSCHKPQAQSWKHTAHAKAMESLKP 60
C91--cycA3      NKIIACGLVAAALFTLTSGQSLAADAFEGGRKKCSSCHKPQAQSWKHTAHAKAMESLKP 60
C91--cycA2      NKIIACGLVAAALFTLTSGQSLAADAFEGGRKKCSSCHKPQAQSWKHTAHAKAMESLKP 60
C91--cycA1      NKIIACGLVAAALFTLTSGQSLAADAFEGGRKKCSSCHKPQAQSWKHTAHAKAMESLKP 60
*****
D23--cycA2      NVKVEAKQKAKLDPTKDXTQDKDCVGCCHVDGFGQKGGYTTIDSPKPMLTGVCCSCHGPGR 120
D23--cycA3      NVKVEAKQKAKLDPTKDXTQDKDCVGCCHVDGFGQKGGYTTIDSPKPMLTGVCCSCHGPGR 120
D23--cycA1      NVKVEAKQKAKLDPTKDXTQDKDCVGCCHVDGFGQKGGYTTIDSPKPMLTGVCCSCHGPGR 120
C91--cycA3      NVKVEAKQKAKLDPAKDYTDQDKDCVGCCHVDGFGQKGGYTTIDSPKPMLTGVCCSCHGPGR 120
C91--cycA2      NVKVEAKQKAKLDPAKDYTDQDKDCVGCCHVDGFGQKGGYTTIDSPKPMLTGVCCSCHGPGR 120
C91--cycA1      NVKVEAKQKAKLDPAKDYTDQDKDCVGCCHVDGFGQKGGYTTIDSPKPMLTGVCCSCHGPGR 120
*****;*****
D23--cycA2      KYRGDHRKAGQAFESGKKAPRKXTLASKGQDFNFEERCSACHLNYEGSPWKGAKPPYTPY 180
D23--cycA3      KYRGDHRKAGQAFESGKKAPRKXTLASKGQDFNFEERCSACHLNYEGSPWKGAKPPYTPY 180
D23--cycA1      KYRGDHRKAGQAFESGKKAPRKXTLASKGQDFNFEERCSACHLNYEGSPWKGAKPPYTPY 180
C91--cycA3      KYRGDHRKAGQAFESGKKAPRKXTLASKGQDFNFEERCSACHLNYEGSPWKGAKPPYTPY 180
C91--cycA2      KYRGDHRKAGQAFESGKKAPRKXTLASKGQDFNFEERCSACHLNYEGSPWKGAKPPYTPY 180
C91--cycA1      KYRGDHRKAGQAFESGKKAPRKXTLASKGQDFNFEERCSACHLNYEGSPWKGAKPPYTPY 180
*****
D23--cycA2      TPEVDPKYTFKFDENVKDVKAMHEHYKLDGVFDGEPKFKFHDQFQANAKTAKKKG 235
D23--cycA3      TPEVDPKYTFKFDENVKDVKAMHEHYKLDGVFDGEPKFKFHDQFQANAKTAKKKG 235
D23--cycA1      TPEVDPKYTFKFDENVKDVKAMHEHYKLDGVFDGEPKFKFHDQFQANAKTAKKKG 235
C91--cycA3      TPEVDPKYTFKFDENVKDVKAMHEHYKLDGVFDGEPKFKFHDQFQANAKTAKKKG 235
C91--cycA2      TPEVDPKYTFKFDENVKDVKAMHEHYKLDGVFDGEPKFKFHDQFQANAKTAKKKG 235
C91--cycA1      TPEVDKTYTFKFDENVKDVKAMHEHYKLDGVFDGEPKFKFHDQFQANAKTAKKKG 235
*****,*****
```

Figure 16

D23--cycB1	MTRLQKGSIGTLLTGALLGIALVAVVFGGEAALSTEEFCTSCHSMSYPQSELKESTHYGA	60
D23--cycB2	MTRLQKGSIGTLLTGALLGIALVAVVFGGEAALSTEEFCTSCHSMSYPQSELKESTHYGA	60
C91--cycB1	MTRLQKGSIGTLLTGALLGIALVAVVFGGEAALSTEEFCTSCHSMSYPQSELKESTHYGA	60
C91--cycB2	MTRLQKGSIGTLLTGALLGIALVAVVFGGEAALSTEEFCTSCHSMSYPQSELKESTHYGA	60

D23--cycB1	LGVNPTCKDCHIPQGIENFHLAVATHVVDGARELWLEMVNDYSTLEKFNERRLEMAHDAR	120
D23--cycB2	LGVNPTCKDCHIPQGIENFHLAVATHVVDGARELWLEMVNDYSTLEKFNERRLEMAHDAR	120
C91--cycB1	LGINPTCKDCHIPQGIENFHLAVATHVVDGARELWLEMVNDYSTLEKFNERRLEMAHDAR	120
C91--cycB2	LGINPTCKDCHIPQGIENFHLAVATHVVDGARELWLEMVNDYSTLEKFNERRLEMAHDAR	120
	**;	
D23--cycB1	MNLKKWDSITCRTCHVKPAPPGESAQAEHKKKMETEGATCIDCHQNLVHEEAPMTDLNASL	180
D23--cycB2	MNLKKWDSITCRTCHVKPAPPGESAQAEHKKKMETEGATCIDCHQNLVHEEAPMTDLNASL	180
C91--cycB1	MNLKKWDSITCRTCHVKPAPPGESAQAEHKKKMETEGATCIDCHQNLVHEEAPMTDLNASL	180
C91--cycB2	MNLKKWDSITCRTCHVKPAPPGESAQAEHKKKMETEGATCIDCHQNLVHEEAPMTDLNASL	180

D23--cycB1	AAGKLVLKPEEGDGGDDDDVDVDDDEEEVEVEEETETATDDSDSASSSNHDDSDDE	239
D23--cycB2	AAGKLVLKPEEGDGGDDDDVDVDDDEEEVEVEEETETATDDSDSASSSNHDDSDDE	239
C91--cycB1	AAGKLVLKSEEGDD---DVEDDDE---DEEVEVEEETETATDDSDSASSSNHDDSDDE	234
C91--cycB2	AAGKLVLKSEEGDDVDVDDVDDDE---DEEVEVEEETETATDDSDSASSSNHDDSDDE	237
	*****,*****,* **** *	

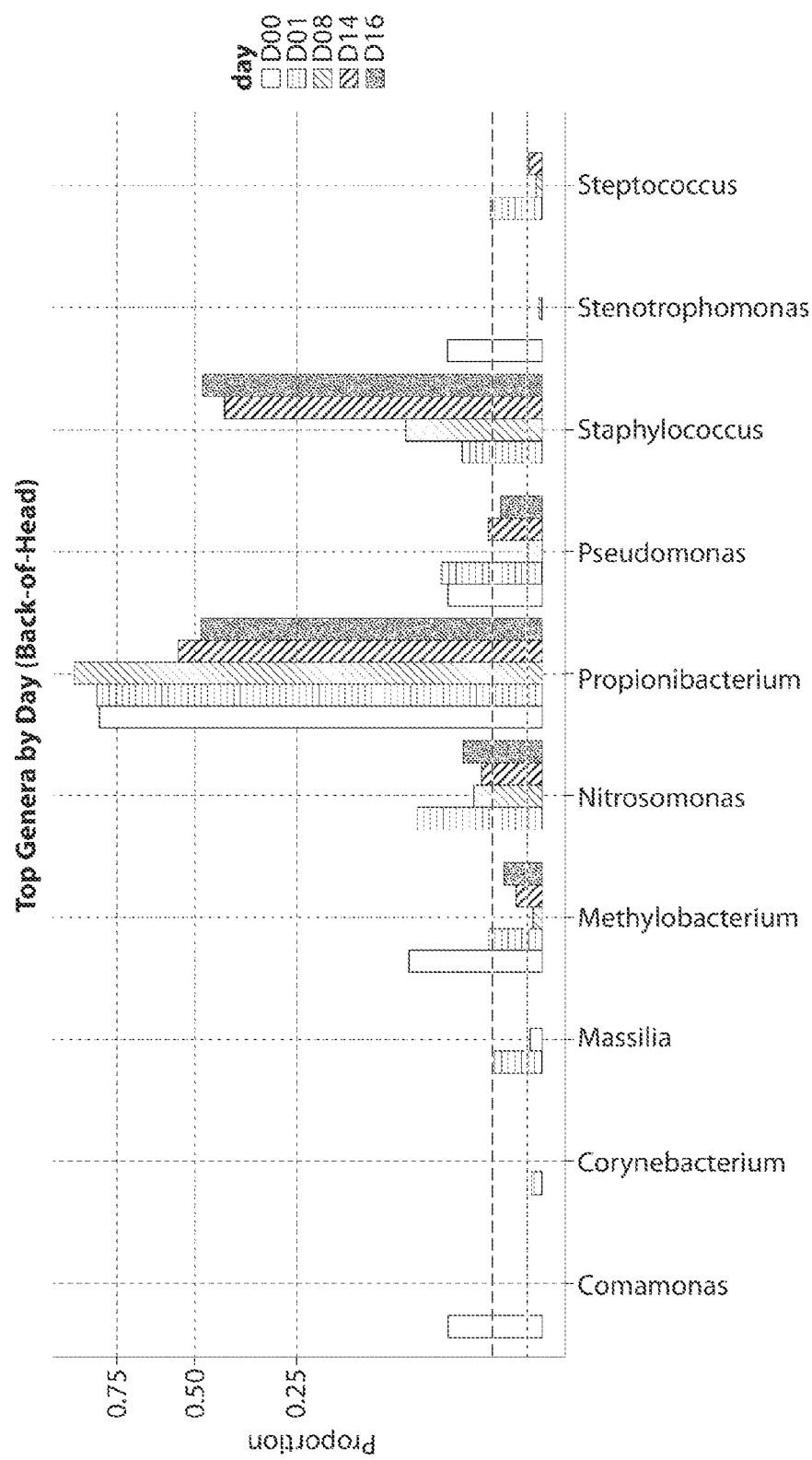


Figure 17

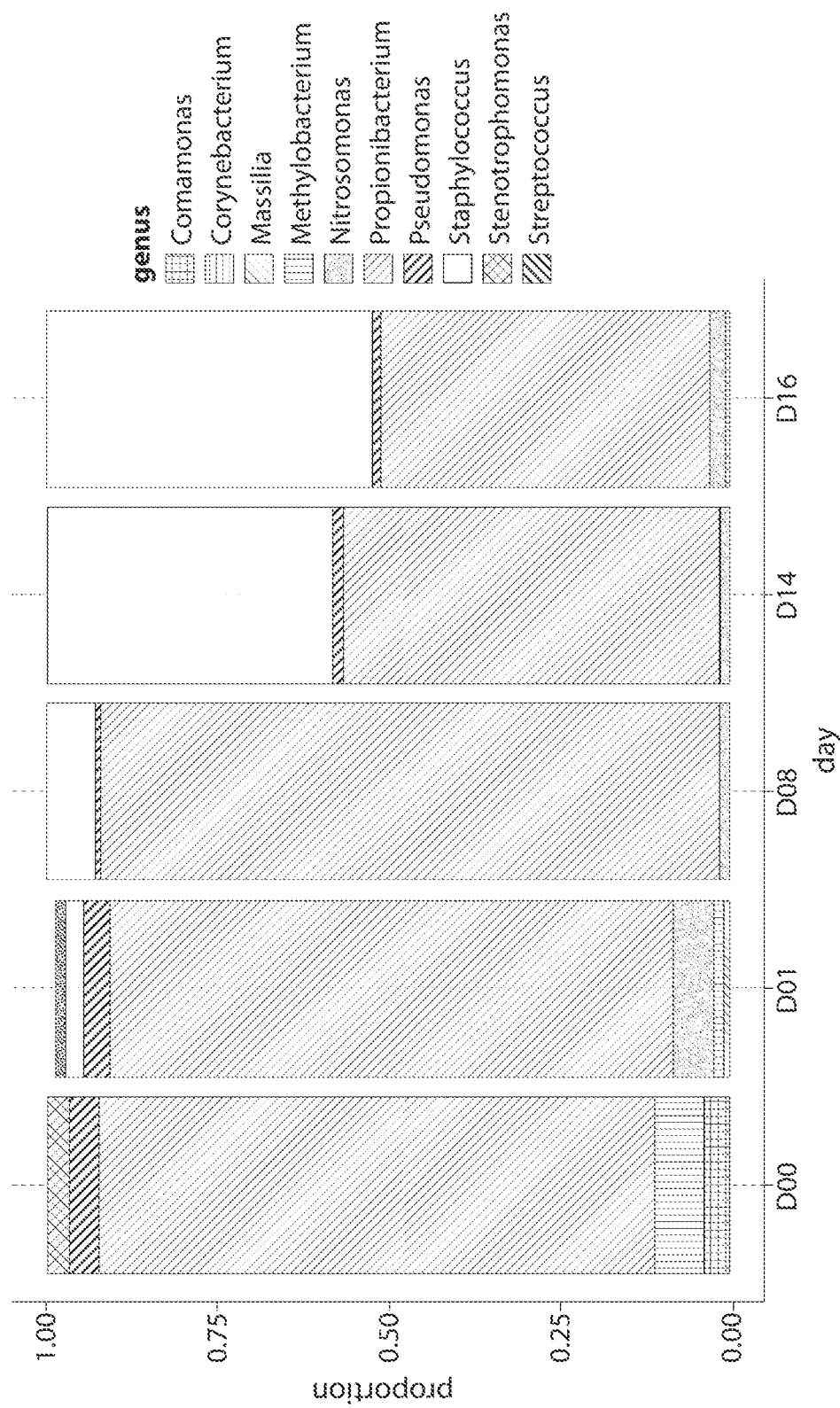


Figure 18

AMMONIA-OXIDIZING NITROSOMONAS EUTROPHA STRAIN D23

[0001] This application claims priority to Greek Patent Application Number 20140100217, filed Apr. 15, 2014, U.S. Provisional Application No. 62/002,084, filed May 22, 2014, U.S. Provisional Application No. 62/012,811, filed Jun. 16, 2014, U.S. Provisional Application No. 62/053,588, filed Sep. 22, 2014, and Greek Patent Application Number 20150100115, filed Mar. 13, 2015, the contents of which are incorporated herein by reference in their entireties.

SEQUENCE LISTING

[0002] The instant application contains a Sequence Listing which has been submitted electronically in ASCII format and is hereby incorporated by reference in its entirety. Said ASCII copy, created on Apr. 13, 2015, is named N2060-7001WO.txt and is 3,590,980 bytes in size.

BACKGROUND

[0003] Beneficial bacteria can be used to suppress the growth of pathogenic bacteria. Bacteria and other microorganisms are ubiquitous in the environment. The discovery of pathogenic bacteria and the germ theory of disease have had a tremendous effect on health and disease states. Bacteria are a normal part of the environment of all living things. In the gut, these bacteria are not pathogenic under normal conditions, and in fact improve health by rendering the normal intestinal contents less hospitable for disease causing organisms. Disease prevention is accomplished in a number of ways: nutrients are consumed, leaving less for pathogens; conditions are produced, such as pH and oxygen tension, which are not hospitable for pathogens; compounds are produced that are toxic to pathogens; pathogens are consumed as food by these microorganisms; less physical space remains available for pathogens; and specific binding sites are occupied leaving fewer binding sites available for pathogens. The presence of these desirable bacteria is seen as useful in preventing disease states.

[0004] There is a need in the art for improved beneficial bacteria that can suppress the growth of pathogenic bacteria.

SUMMARY

[0005] This disclosure provides, inter alia, an optimized strain of *Nitrosomonas eutropha* (*N. eutropha*) designated D23, D23-100 or AOB D23-100, the terms which may be used interchangeably throughout the disclosure.

[0006] Ammonia oxidizing bacterial of the genus *Nitrosomonas* are ubiquitous Gram-negative obligate chemolithoautotrophic bacteria with a unique capacity to generate energy exclusively from the conversion of ammonia to nitrite.

[0007] *N. eutropha* bacteria disclosed in this application have desirable, e.g. optimized, properties such as the ability to suppress growth of pathogenic bacteria, and an enhanced ability to produce nitric oxide (NO) and nitric oxide (NO_2^-) precursors. The *N. eutropha*, e.g., optimized *N. eutropha*, e.g., purified preparations of optimized *N. eutropha* herein may be used, for instance, to treat diseases, e.g., diseases associated with low nitrite levels, skin disorders, and diseases caused by pathogenic bacteria. When referring to *N.*

eutropha throughout the disclosure, it may be referring to an optimized strain of *N. eutropha* or a purified preparation of optimized *N. eutropha*.

[0008] The present disclosure provides, inter alia, a *Nitrosomonas eutropha* (*N. eutropha*) bacterium, e.g., an optimized *N. eutropha*, e.g., a purified preparation of optimized *N. eutropha*, having at least one property selected from:

[0009] an optimized growth rate;

[0010] an optimized NH_4^+ oxidation rate; and

[0011] an optimized resistance to ammonium ion (NH_4^+).

[0012] The bacterium is optionally axenic.

[0013] In embodiments, the optimized growth rate is a rate allowing a continuous culture of *N. eutropha* at an OD600 (optical density at 600 nm) of about 0.15-0.18 to reach an OD600 of about 0.5-0.6 in about 1-2 days. In embodiments, optimized growth rate is a doubling time of about 8 hours when cultured under batch culture conditions. In embodiments, the optimized NH_4^+ oxidation rate is a rate of at least about 125 micromoles per minute of oxidizing NH_4^+ to NO_2^- . In embodiments, the optimized resistance to NH_4^+ is an ability to grow in medium comprising about 200 mM NH_4^+ for at least about 48 hours.

[0014] In some embodiments, the purified preparation of optimized *N. eutropha* bacterium (which is optionally axenic) has at least two properties selected from an optimized growth rate, an optimized NH_4^+ oxidation rate, and an optimized resistance to NH_4^+ . In some embodiments, the purified preparation of optimized *N. eutropha* bacterium (which is optionally axenic) has an optimized growth rate, an optimized NH_4^+ oxidation rate, and an optimized resistance to NH_4^+ . In some embodiments, the purified preparation of optimized *N. eutropha* bacterium (which is optionally axenic) comprises a chromosome that hybridizes under very high stringency to SEQ ID NO: 1.

[0015] In some embodiments, the purified preparation of optimized *N. eutropha* bacterium (which is optionally axenic) comprises an AmoA protein having an identity to SEQ ID NO: 6 or 12 selected from at least about 80%, 85%, 90%, 95%, 98%, 99%, 99.5%, and 100% identical, an AmoB protein having an identity to SEQ ID NO: 8 or 14 selected from at least about 80%, 85%, 90%, 95%, 98%, 99%, 99.5%, and 100% identical, an amoC gene having an identity to SEQ ID NO: 4, 10, or 16 selected from at least about 80%, 85%, 90%, 95%, 98%, 99%, 99.5%, and 100% identical, a hydroxylamine oxidoreductase protein having an identity to SEQ ID NO: 18, 20, or 22 selected from at least about 80%, 85%, 90%, 95%, 98%, 99%, 99.5%, and 100% identical, a cytochrome c554 protein having an identity to SEQ ID NO: 24, 26, or 28 selected from at least about 80%, 85%, 90%, 95%, 98%, 99%, 99.5%, and 100% identical, or a cytochrome c_M552 protein having an identity to SEQ ID NO: 30 or 32 selected from at least about 80%, 85%, 90%, 95%, 98%, 99%, 99.5%, and 100% identical.

[0016] In some embodiments, the purified preparation of optimized *N. eutropha* bacterium (which is optionally axenic) comprises 1-5, 5-10, 10-15, 15-20, 20-25, 25-30, or all of the sequence characteristics of Table 2. For instance, in some embodiments, the bacterium or preparation comprises an AmoA1 or AmoA2 protein having (or gene encoding) a mutation relative to *N. eutropha* strain C91 at position 1, e.g., a V at position 1. In some embodiments, the bacterium or preparation comprises an AmoA1 or AmoA2

protein having (or gene encoding) a mutation relative to *N. eutropha* strain C91 at position 160, e.g., an L at position 160. In some embodiments, the bacterium or preparation comprises an AmoA1 or AmoA2 protein having (or gene encoding) a mutation relative to *N. eutropha* strain C91 at position 167, e.g., an A at position 167. In some embodiments, the bacterium or preparation comprises an AmoB1 or AmoB2 protein having (or gene encoding) a mutation relative to *N. eutropha* strain C91 at position 33, e.g., a V at position 33. In some embodiments, the bacterium or preparation comprises an AmoB1 or AmoB2 protein having (or gene encoding) a mutation relative to *N. eutropha* strain C91 at position 165, e.g., an I at position 165. In some embodiments, the bacterium or preparation comprises an AmoC3 protein having (or gene encoding) a mutation relative to *N. eutropha* strain C91 at position 79, e.g., an A at position 79. In some embodiments, the bacterium or preparation comprises an AmoC3 protein having (or gene encoding) a mutation relative to *N. eutropha* strain C91 at position 271, e.g., a V at position 271. In some embodiments, the bacterium or preparation comprises a Hao1, Hao2, or Hao3 protein having (or gene encoding) a mutation relative to *N. eutropha* strain C91 at position 85, e.g., an S at position 85. In some embodiments, the bacterium or preparation comprises a Hao1, Hao2, or Hao3 protein having (or gene encoding) a mutation relative to *N. eutropha* strain C91 at position 312, e.g., an E at position 312. In some embodiments, the bacterium or preparation comprises a Hao1 protein having (or gene encoding) a mutation relative to *N. eutropha* strain C91 at position 163, e.g., an A at position 163. In some embodiments, the bacterium or preparation comprises a c554 CycA1, c554 CycA2, or c554 CycA3 protein having (or gene encoding) a mutation relative to *N. eutropha* strain C91 at position 65, e.g., a T at position 65. In some embodiments, the bacterium or preparation comprises a c554 CycA1 protein having (or gene encoding) a mutation relative to *N. eutropha* strain C91 at position 186, e.g., a T at position 186. In some embodiments, the bacterium or preparation comprises a c_M552 CycB1 or c_M552 CycB2 protein having (or gene encoding) a mutation relative to *N. eutropha* strain C91 at position 63, e.g., a V at position 63. In some embodiments, the bacterium or preparation comprises a c_M552 CycB1 or c_M552 CycB2 protein having (or gene encoding) a mutation relative to *N. eutropha* strain C91 at position 189, e.g., a P at position 189. In some embodiments, the bacterium or preparation comprises a c_M552 CycB1 or c_M552 CycB2 protein having (or gene encoding) a mutation relative to *N. eutropha* strain C91 at position 206, e.g., an insE at position 206. In some embodiments, the bacterium or preparation comprises a c_M552 CycB1 or c_M552 CycB2 protein having (or gene encoding) a mutation relative to *N. eutropha* strain C91 at position 207, e.g., an insE at position 207. In some embodiments, the bacterium or preparation comprises a c_M552 CycB1 protein having (or gene encoding) a mutation relative to *N. eutropha* strain C91 at position 195, e.g., an insD at position 195. In some embodiments, the bacterium or preparation comprises a c_M552 CycB1 protein having (or gene encoding) a mutation relative to *N. eutropha* strain C91 at position 196, e.g., an insD at position 196. In some embodiments, the bacterium or preparation comprises a c_M552 CycB1 protein having (or gene encoding) a mutation relative to *N. eutropha* strain C91 at position 197, e.g., an insD at position 197.

[0017] Combinations of two or more sequence characteristics of Table 2 are also described. The two or more sequence characteristics may be in the same gene or different genes. The two or more sequence characteristics may be in the same protein or different proteins. For instance, in some embodiments, the bacterium or preparation comprises an AmoA1 or AmoA2 protein having (or gene encoding) a mutation relative to *N. eutropha* strain C91 at position 1, e.g., a V at position 1 and a mutation relative to *N. eutropha* strain C91 at position 160, e.g., an L at position 160. In some embodiments, the bacterium or preparation comprises an AmoA1 or AmoA2 protein having (or gene encoding) a mutation relative to *N. eutropha* strain C91 at position 1, e.g., a V at position 1 and a mutation relative to *N. eutropha* strain C91 at position 167, e.g., an A at position 167. In some embodiments, the bacterium or preparation comprises an AmoA1 or AmoA2 protein having (or gene encoding) a mutation relative to *N. eutropha* strain C91 at position 160, e.g., an L at position 160 and a mutation relative to *N. eutropha* strain C91 at position 167, e.g., an A at position 167.

[0018] In some embodiments, the bacterium or preparation comprises an AmoB1 or AmoB2 protein having (or gene encoding) a mutation relative to *N. eutropha* strain C91 at position 33, e.g., a V at position 33 and a mutation relative to *N. eutropha* strain C91 at position 165, e.g., an I at position 165.

[0019] In some embodiments, the bacterium or preparation comprises an AmoC3 protein having (or gene encoding) a mutation relative to *N. eutropha* strain C91 at position 79, e.g., an A at position 79 and a mutation relative to *N. eutropha* strain C91 at position 271, e.g., a V at position 271.

[0020] In some embodiments, the bacterium or preparation comprises a Hao1, Hao2, or Hao3 protein having (or gene encoding) a mutation relative to *N. eutropha* strain C91 at position 85, e.g., an S at position 85 and a mutation relative to *N. eutropha* strain C91 at position 312, e.g., an E at position 312. In some embodiments, the bacterium or preparation comprises a Hao1 protein having (or gene encoding) a mutation relative to *N. eutropha* strain C91 at position 85, e.g., an S at position 85 and a mutation relative to *N. eutropha* strain C91 at position 163, e.g., an A at position 163. In some embodiments, the bacterium or preparation comprises a Hao1 protein having (or gene encoding) a mutation relative to *N. eutropha* strain C91 at position 312, e.g., an E at position 312 and a mutation relative to *N. eutropha* strain C91 at position 163, e.g., an A at position 163.

[0021] In some embodiments, the bacterium or preparation comprises a c554 CycA1 protein having (or gene encoding) a mutation relative to *N. eutropha* strain C91 at position 65, e.g., a T at position 65 and a mutation relative to *N. eutropha* strain C91 at position 186, e.g., a T at position 186.

[0022] In some embodiments, the bacterium or preparation comprises a c_M552 CycB1 protein having (or gene encoding) mutations at any two or more of the following amino acid positions: 63, 189, 194, 195, 196, 197, 206, and 207. For instance, the two or more amino acid positions may comprise: 63 and 189, 63 and 194, 63 and 195, 63 and 196, 63 and 197, 63 and 206, 63 and 207, 189 and 194, 189 and 195, 189 and 196, 189 and 194, 189 and 195, 189 and 196, 189 and 197, 189 and 206, 189 and 207, 194 and 195, 194 and 196, 194 and 197, 194 and 206, 194 and 207, 195 and

196, 195 and 197, 195 and 206, 195 and 207, 196 and 197, 196 and 206, 196 and 207, 197 and 206, 197 and 207, or 206 and 207. In some embodiments, the bacterium or preparation comprises a c_M552 CycB1 protein having (or gene encoding) any two or more mutations selected from the group consisting of: I63V, S189P, D194G, 195insD, 196insD, 197insD, 206insE, and 207insE. For instance, the two or more mutations can be selected from the group consisting of: I63V and S189P, I63V and D194G, I63V and 195insD, I63V and 196insD, I63V and 197insD, I63V and 206insE, I63V and 207insE, S189P and D194G, S189P and 195insD, S189P and 196insD, S189P and 197insD, S189P and 206insE, S189P and 207insE, D194G and 195insD, D194G and 196insD, D194G and 197insD, D194G and 206insE, D194G and 207insE, 195insD and 196insD, 195insD and 197insD, 195insD and 206insE, 195insD and 207insE, 196insD and 197insD, 196insD and 206insE, 196insD and 207insE, 197insD and 206insE, 197insD and 207insE, and 206insE and 207insE.

[0023] In some embodiments, the bacterium or preparation comprises a c_M552 CycB2 protein having (or gene encoding) mutations at any two or more of the following amino acid positions: 63, 189, 206, and 207. For instance, the two or more amino acid positions may comprise: 63 and 189, 63 and 206, 63 and 207, 189 and 206, 189 and 207, or 206 and 207. In some embodiments, the bacterium or preparation comprises a c_M552 CycB2 protein having (or gene encoding) any two or more mutations selected from the group consisting of: I63V, S189P, 206insE, and 207insE. For instance, the two or more mutations can be selected from the group consisting of: I63V and S189P, I63V and 206insE, I63V and 207insE, S189P and 206insE, S189P and 207insE, and 206insE and 207insE.

[0024] Combinations of three or more sequence characteristics of Table 2 are also described. For instance, in some embodiments, the bacterium or preparation comprises an AmoA1 or AmoA2 protein having (or gene encoding) a mutation relative to *N. eutropha* strain C91 at position 1, e.g., a V at position 1 and a mutation relative to *N. eutropha* strain C91 at position 160, e.g., an L at position 160, and a mutation relative to *N. eutropha* strain C91 at position 167, e.g., an A at position 167.

[0025] In some embodiments, the bacterium or preparation comprises a Hao1 protein having (or gene encoding) a mutation relative to *N. eutropha* strain C91 at position 85, e.g., an S at position 85 and a mutation relative to *N. eutropha* strain C91 at position 312, e.g., an E at position 312, and a mutation relative to *N. eutropha* strain C91 at position 163, e.g., an A at position 163.

[0026] In some embodiments, the bacterium or preparation comprises a c_M552 CycB1 protein having (or gene encoding) mutations at any three or more (e.g., 4, 5, 6, 7, or all) of the following amino acid positions: 63, 189, 194, 195, 196, 197, 206, and 207. For instance, the three mutations may be at positions 195, 196, and 197. In some embodiments, the bacterium or preparation comprises a c_M552 CycB1 protein having (or gene encoding) any three or more (e.g., 4, 5, 6, 7, or all) mutations selected from the group consisting of: I63V, S189P, D194G, 195insD, 196insD, 197insD, 206insE, and 207insE. For instance, the three mutations may be 195insD, 196insD, and 197insD.

[0027] In some embodiments, the bacterium or preparation comprises a c_M552 CycB2 protein having (or gene encoding) mutations at any three or more (e.g., all) of the

following amino acid positions: 63, 189, 206, and 207. In some embodiments, the bacterium or preparation comprises a c_M552 CycB2 protein having (or gene encoding) any three or more (e.g., all) mutations selected from the group consisting of: I63V, S189P, 206insE, and 207insE.

[0028] In some embodiments, the bacterium or preparation comprises mutations relative to *N. eutropha* strain C91 in at least two genes, e.g., at least two genes listed in Table 2. The two genes may be, for instance, AmoA1 and AmoA2, AmoA1 and AmoB1, AmoA1 and AmoB2, AmoA1 and AmoC1, AmoA1 and AmoC2, AmoA1 and AmoC3, AmoA1 and Hao1, AmoA1 and Hao2, AmoA1 and Hao3, AmoA1 and c554 CycA1, AmoA1 and c554 CycA2, AmoA1 and c554 CycA3, AmoA1 and cM552 CycB1, AmoA1 and cM552 CycB2, AmoA2 and AmoB1, AmoA2 and AmoB2, AmoA2 and AmoC1, AmoA2 and AmoC2, AmoA2 and AmoC3, AmoA2 and Hao1, AmoA2 and Hao2, AmoA2 and Hao3, AmoA2 and c554 CycA1, AmoA2 and c554 CycA2, AmoA2 and c554 CycA3, AmoA2 and cM552 CycB1, AmoA2 and cM552 CycB2, AmoB1 and AmoB2, AmoB1 and AmoC1, AmoB1 and AmoC2, AmoB1 and AmoC3, AmoB1 and Hao1, AmoB1 and Hao2, AmoB1 and Hao3, AmoB1 and c554 CycA1, AmoB1 and c554 CycA2, AmoB1 and c554 CycA3, AmoB1 and cM552 CycB1, AmoB1 and cM552 CycB2, AmoB2 and AmoC1, AmoB2 and AmoC2, AmoB2 and AmoC3, AmoB2 and Hao1, AmoB2 and Hao2, AmoB2 and Hao3, AmoB2 and c554 CycA1, AmoB2 and c554 CycA2, AmoB2 and c554 CycA3, AmoB2 and cM552 CycB1, AmoB2 and cM552 CycB2, AmoC1 and AmoC2, AmoC1 and AmoC3, AmoC1 and Hao1, AmoC1 and Hao2, AmoC1 and Hao3, AmoC1 and c554 CycA1, AmoC1 and c554 CycA2, AmoC1 and c554 CycA3, AmoC1 and cM552 CycB1, AmoC1 and cM552 CycB2, AmoC2 and AmoC3, AmoC2 and Hao1, AmoC2 and Hao2, AmoC2 and Hao3, AmoC2 and c554 CycA1, AmoC2 and c554 CycA2, AmoC2 and c554 CycA3, AmoC2 and cM552 CycB1, AmoC2 and cM552 CycB2, AmoC3 and Hao1, AmoC3 and Hao2, AmoC3 and Hao3, AmoC3 and c554 CycA1, AmoC3 and c554 CycA2, AmoC3 and c554 CycA3, AmoC3 and cM552 CycB1, AmoC3 and cM552 CycB2, Hao1 and Hao2, Hao1 and Hao3, Hao1 and c554 CycA1, Hao1 and c554 CycA2, Hao1 and c554 CycA3, Hao1 and cM552 CycB1, Hao1 and cM552 CycB2, Hao2 and Hao3, Hao2 and c554 CycA1, Hao2 and c554 CycA2, Hao2 and c554 CycA3, Hao2 and cM552 CycB1, Hao2 and cM552 CycB2, Hao3 and c554 CycA1, Hao3 and c554 CycA2, Hao3 and c554 CycA3, Hao3 and cM552 CycB1, Hao3 and cM552 CycB2, c554 CycA1 and c554 CycA2, c554 CycA1 and c554 CycA3, c554 CycA1 and cM552 CycB1, c554 CycA1 and cM552 CycB2, c554 CycA2 and c554 CycA3, c554 CycA2 and cM552 CycB1, c554 CycA2 and cM552 CycB2, c554 CycA3 and cM552 CycB1, c554 CycA3 and cM552 CycB2, or cM552 CycB1 and cM552 CycB2.

[0029] In some embodiments, the bacterium or preparation comprises mutations relative to *N. eutropha* strain C91 in at least three genes, e.g., at least three (e.g., 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, or all) genes listed in Table 2. The three genes may be, for instance AmoA1 and AmoA2 and AmoA3; AmoC1 and AmoC2 and AmoC3; or Hao1 and Hao2 and Hao3.

[0030] In some embodiments, the bacterium or preparation comprises at least one structural difference, e.g., at least one mutation, relative to a wild-type bacterium such as *N. eutropha* strain C91. In some embodiments, the bacterium or

preparation comprises a nucleic acid that can be amplified using a pair of primers described herein, e.g., a primer comprising a sequence of SEQ ID NO: 64 and a primer comprising a sequence of SEQ ID NO: 65. In some embodiments, the bacterium or preparation comprises a nucleic acid or protein at least 80%, 85%, 90%, 95%, 98%, 99%, 99.5%, or 100% identical to a gene of FIG. 6, 7, or 8, or a protein encoded by a gene of FIG. 6, 7, or 8. In some embodiments, the bacterium or preparation comprises a nucleic acid or protein at least 80%, 85%, 90%, 95%, 98%, 99%, 99.5%, or 100% identical to a sequence of any of SEQ ID NOS: 64-66 or a protein encoded by a sequence of any of SEQ ID NOS: 64-66.

[0031] In some aspects, the present disclosure provides, inter alia, an *N. eutropha* bacterium, or a purified preparation thereof, comprising a mutation in an ammonia monooxygenase gene, a hydroxylamine oxidoreductase gene, a cytochrome c554 gene, or a cytochrome c_M552 gene. The mutation may be relative to a wild-type bacterium such as *N. eutropha* strain C91. The mutation may be in one or more of the amoA1 gene, the amoA2 gene, amoB1 gene, the amoB2 gene, and the amoC3 gene. The *N. eutropha* bacterium, or a purified preparation thereof may have a mutation at a position described herein, e.g., in Table 2. The *N. eutropha* bacterium, or a purified preparation thereof may have a mutation wherein said mutation is a mutation described herein, e.g., in Table 2.

[0032] In some embodiments, the mutation may be in one or more of the hao1 gene, the hao2 gene, or the hao3 gene. The *N. eutropha* bacterium, or a purified preparation thereof may have a mutation at a position described herein, e.g., in Table 2. The *N. eutropha* bacterium, or a purified preparation thereof may have a mutation wherein said mutation is a mutation described herein, e.g., in Table 2.

[0033] In some embodiments, the mutation may be in one or more of the c554 cycA1 gene, the c554 cycA2 gene, and the c554 cycA3 gene. The *N. eutropha* bacterium, or a purified preparation thereof may have a mutation at a position described herein, e.g., in Table 2. The *N. eutropha* bacterium, or a purified preparation thereof may have a mutation wherein said mutation is a mutation described herein, e.g., in Table 2.

[0034] In some embodiments, the mutation may be in one or more of the c_M552 cycB1 gene and the c_M552 cycB2 gene. The *N. eutropha* bacterium, or a purified preparation thereof may have a mutation at a position described herein, e.g., in Table 2. The *N. eutropha* bacterium, or a purified preparation thereof may have a mutation wherein said mutation is a mutation described herein, e.g., in Table 2.

[0035] In certain aspects the *N. eutropha* bacterium, or a purified preparation thereof, described in the preceding four paragraphs may be based on a *N. eutropha* bacterium, e.g., an optimized *N. eutropha*, e.g., a purified preparation of optimized *N. eutropha*, having at least one property selected from:

[0036] an optimized growth rate;

[0037] an optimized NH₄⁺ oxidation rate; and

[0038] an optimized resistance to ammonium ion (NH₄⁺).

[0039] In certain aspects, the *N. eutropha* bacterium, or a purified preparation thereof, described in the preceding five paragraphs may have a mutation in at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, or 35 positions of one

or more of amoA1 gene, amoA2 gene, amoB1 gene, amoB2 gene, amoC3 gene, hao1 gene, hao2 gene, hao3 gene, c554 cycA1 gene, c554 cycA2 gene, c554 cycA3 gene, c_M552 cycB1 gene, and c554 cycB2 gene.

[0040] In some embodiments, the *N. eutropha* bacterium has an optimized growth rate, e.g., an optimized growth rate described herein, and a structural difference such as a mutation (e.g., relative to a wild-type strain such as *N. eutropha* strain C91), e.g., a mutation described herein, e.g., a mutation of Table 2. In some embodiments, the *N. eutropha* bacterium has an optimized NH₄⁺ oxidation rate, e.g., an optimized NH₄⁺ oxidation rate described herein, and a structural difference such as a mutation (e.g., relative to a wild-type strain such as *N. eutropha* strain C91), e.g., a mutation described herein, e.g., a mutation of Table 2. In some embodiments, the *N. eutropha* bacterium has an optimized resistance to NH₄⁺, e.g., an optimized resistance to NH₄⁺ described herein, and a structural difference such as a mutation (e.g., relative to a wild-type strain such as *N. eutropha* strain C91), e.g., a mutation described herein, e.g., a mutation of Table 2.

[0041] In some embodiments, the *N. eutropha* bacterium comprises a nucleic acid that can be amplified using a pair of primers described herein, e.g., a primer comprising a sequence of SEQ ID NO: 64 and a primer comprising a sequence of SEQ ID NO: 65.

[0042] In certain aspects, this disclosure provides an *N. eutropha* bacterium (which is optionally axenic) comprising a chromosome that hybridizes at high stringency to SEQ ID NO: 1.

[0043] In embodiments, the chromosome hybridizes at very high stringency to SEQ ID NO: 1. In embodiments, the *N. eutropha* bacterium (which is optionally axenic) comprises a gene that is at least about 80%, 85%, 90%, 95%, 97%, 98%, 99%, or 100% identical to one or more genes of FIGS. 6-8 (e.g., 10, 20, 30, 40, 50, 100, or all genes of any one or more of FIGS. 6, 7, and 8).

[0044] In embodiments, the *N. eutropha* bacterium (which is optionally axenic) lacks any plasmid that is at least about 80%, 85%, 90%, 95%, 97%, 98%, 99%, or 100% identical to SEQ ID NO: 2 (pNeut1) or SEQ ID NO: 3 (pNeut2), as described by Stein et al. Whole-genome analysis of the ammonia-oxidizing bacterium, *Nitrosomonas eutropha* C91: implications for niche adaptation. Environmental Microbiology (2007) 9(12), 2993-3007. In embodiments, the *N. eutropha* (which is optionally axenic) lacks one or more genes present on the plasmids of SEQ ID NO: 2 or SEQ ID NO: 3. For instance, the *N. eutropha* (which is optionally axenic) may lack at least 2, 3, 4, 5, 10, 15, or 20 genes present on one or both of pNeut1 and pNeut2. pNeut1 contains 55 protein-coding sequences while pNeut2 contains 52 protein-coding sequences. In embodiments, the *N. eutropha* bacterium (which is optionally axenic) lacks any plasmid.

[0045] In certain aspects, this disclosure provides an *N. eutropha* bacterium (which is optionally axenic) comprising one or more of an amoA1 gene at least about 98.9% identical to SEQ ID NO: 7 and an amoA2 gene at least about 98.8% identical to SEQ ID NO: 13.

[0046] In certain aspects, this disclosure provides an *N. eutropha* bacterium (which is optionally axenic) comprising one or more of an AmoA1 protein at least about 99.0% identical to SEQ ID NO: 6 and an AmoA2 protein at least about 99.0% identical to SEQ ID NO: 12.

[0049] In certain aspects, this disclosure provides an *N. eutropha* bacterium (which is optionally axenic) comprising one or more of an AmoB1 protein at least about 99.6% identical to SEQ ID NO: 8 and an AmoB1 protein at least about 99.6% identical to SEQ ID NO: 14.

[0051] In certain aspects, this disclosure provides an *N. eutropha* bacterium (which is optionally axenic) comprising one or more of an amoC1 gene at least about 99.9% identical to SEQ ID NO: 5, an amoC2 gene at least about 99.9% identical to SEQ ID NO: 11, and an amoC3 gene at least about 99.0% identical to SEQ ID NO: 17.

[0052] In embodiments, the *N. eutropha* bacterium (which is optionally axenic) further comprises one or more of an amoA1 gene at least about 98.9% identical to SEQ ID NO: 7, an amo2 gene at least about 98.9% identical to SEQ ID NO: 13, an amoB1 gene at least about 99.2% identical to SEQ ID NO: 9, and an amoB2 gene at least about 99.2% identical to SEQ ID NO: 15.

[0054] In embodiments, the *N. eutropha* bacterium (which is optionally axenic) further comprises one or more of an AmoA1 protein at least about 99.0% identical to SEQ ID NO: 6, an AmoA2 protein at least about 99.0% identical to SEQ ID NO: 12, an AmoB1 protein at least about 99.6% identical to SEQ ID NO: 8, and an AmoB1 protein at least about 99.6% identical to SEQ ID NO: 14.

[0055] In certain aspects, this disclosure provides an *N. eutropha* bacterium (which is optionally axenic) comprising one or more of a hao1 gene at least about 99.1% identical to SEQ ID NO: 19, a hao2 gene at least about 99.5% identical to SEQ ID NO: 21, and a hao3 gene at least about 99.3% identical to SEQ ID NO: 23.

[0056] In embodiments, the *N. eutropha* bacterium (which is optionally axenic) further comprises one or more of an amoA1 gene at least about 98.9% identical to SEQ ID NO: 7, an amo2 gene at least about 98.9% identical to SEQ ID NO: 13, an amoB1 gene at least about 99.2% identical to SEQ ID NO: 9, an amoB2 gene at least about 99.2% identical to SEQ ID NO: 15, an amoC1 gene at least about 99.9% identical to SEQ ID NO: 5, an amoC2 gene at least about 99.9% identical to SEQ ID NO: 11, and an amoC3 gene at least about 99.0% identical to SEQ ID NO: 17.

identical to SEQ ID NO: 20, and a Hao3 protein at least about 99.7% identical to SEQ ID NO: 22.

[0058] In embodiments, the *N. eutropha* bacterium (which is optionally axenic) further comprises an AmoA1 protein at least about 99.0% identical to SEQ ID NO: 6, an AmoA2 protein at least about 99.0% identical to SEQ ID NO: 12, an AmoB1 protein at least about 99.6% identical to SEQ ID NO: 8, an AmoB1 protein at least about 99.6% identical to SEQ ID NO: 14, or an AmoC3 protein at least about 99.4% identical to SEQ ID NO: 16.

[0059] In certain aspects, this disclosure provides an *N. eutropha* bacterium (which is optionally axenic) comprising one or more of a *cycA1* gene at least about 98.1% identical to SEQ ID NO: 25, a *cycA2* gene at least about 98.8% identical to SEQ ID NO: 27, and a *cycA3* gene at least about 99.4% identical to SEQ ID NO: 28.

[0060] In embodiments, the *N. eutropha* bacterium (which is optionally axenic) further comprises one or more of an amoA1 gene at least about 98.9% identical to SEQ ID NO: 7, an amo2 gene at least about 98.9% identical to SEQ ID NO: 13, an amoB1 gene at least about 99.2% identical to SEQ ID NO: 9, an amoB2 gene at least about 99.2% identical to SEQ ID NO: 15, an amoC1 gene at least about 99.9% identical to SEQ ID NO: 5, an amoC2 gene at least about 99.9% identical to SEQ ID NO: 11, an amoC3 gene at least about 99.0% identical to SEQ ID NO: 17, a hao1 gene at least about 99.1% identical to SEQ ID NO: 19, a hao2 gene at least about 99.5% identical to SEQ ID NO: 21, and a hao3 gene at least about 99.3% identical to SEQ ID NO: 23.

[0061] In certain aspects, this disclosure provides an *N. eutropha* bacterium (which is optionally axenic) comprising one or more of a CycA1 protein at least about 99.2% identical to SEQ ID NO: 24, a CycA2 protein at least about 99.7% identical to SEQ ID NO: 26, and a CycA3 protein at least about 99.7% identical to SEQ ID NO: 28.

[0062] In embodiments, the *N. eutropha* bacterium (which is optionally axenic) further comprises one or more of an AmoA1 protein at least about 99.0% identical to SEQ ID NO: 6, an AmoA2 protein at least about 99.0% identical to SEQ ID NO: 12, an AmoB1 protein at least about 99.6% identical to SEQ ID NO: 8, an AmoB1 protein at least about 99.6% identical to SEQ ID NO: 14, an AmoC3 protein at least about 99.4% identical to SEQ ID NO: 16, a Hao1 protein at least about 99.6% identical to SEQ ID NO: 18, a Hao2 protein at least about 99.7% identical to SEQ ID NO: 20, and a Hao3 protein at least about 99.7% identical to SEQ ID NO: 22.

[0063] In certain aspects, this disclosure provides an *N. eutropha* bacterium (which is optionally axenic) comprising one or more of a cycB1 gene at least about 96.8% identical to SEQ ID NO: 31 and a cycB2 gene at least about 97.2% identical to SEQ ID NO: 33.

[0064] In embodiments, the *N. eutropha* bacterium (which is optionally axenic) further comprises one or more of an amoA1 gene at least about 98.9% identical to SEQ ID NO: 7, an amo2 gene at least about 98.9% identical to SEQ ID NO: 13, an amoB1 gene at least about 99.2% identical to SEQ ID NO: 9, an amoB2 gene at least about 99.2% identical to SEQ ID NO: 15, an amoC1 gene at least about 99.9% identical to SEQ ID NO: 5, an amoC2 gene at least about 99.9% identical to SEQ ID NO: 11, an amoC3 gene at least about 99.0% identical to SEQ ID NO: 17, a hao1 gene at least about 99.1% identical to SEQ ID NO: 19, a hao2

gene at least about 99.5% identical to SEQ ID NO: 21, a hao3 gene at least about 99.3% identical to SEQ ID NO: 23, a cycA1 gene at least about 98.1% identical to SEQ ID NO: 25, a cycA2 gene at least about 98.8% identical to SEQ ID NO: 27, and a cycA3 gene at least about 99.4% identical to SEQ ID NO: 28.

[0065] In certain aspects, this disclosure provides an *N. eutropha* bacterium (which is optionally axenic) comprising one or more of a CycB1 protein at least about 97.2% identical to SEQ ID NO: 30 or a CycB2 protein at least about 98.8% identical to SEQ ID NO: 32.

[0066] In embodiments, the *N. eutropha* bacterium (which is optionally axenic) further comprises one or more of an AmoA1 protein at least about 99.0% identical to SEQ ID NO: 6, an AmoA2 protein at least about 99.0% identical to SEQ ID NO: 12, an AmoB1 protein at least about 99.6% identical to SEQ ID NO: 8, an AmoB1 protein at least about 99.6% identical to SEQ ID NO: 14, an AmoC3 protein at least about 99.4% identical to SEQ ID NO: 16, a Hao1 protein at least about 99.6% identical to SEQ ID NO: 18, a Hao2 protein at least about 99.7% identical to SEQ ID NO: 20, a Hao3 protein at least about 99.7% identical to SEQ ID NO: 22, a CycA1 protein at least about 99.2% identical to SEQ ID NO: 24, a CycA2 protein at least about 99.7% identical to SEQ ID NO: 26, and a CycA3 protein at least about 99.7% identical to SEQ ID NO: 28.

[0067] In certain aspects, this disclosure provides an *N. eutropha* bacterium (which is optionally axenic) comprising one or more genes according to SEQ ID NOS: 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, and 33.

[0068] In certain aspects, this disclosure provides an *N. eutropha* bacterium (which is optionally axenic) comprising one or more proteins according to SEQ ID NOS: 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, and 32.

[0069] In certain aspects, this disclosure provides an *N. eutropha* bacterium (which is optionally axenic) comprising a protein that is mutant relative to *N. eutropha* strain C91 at at least 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, or all of the amino acid positions listed in Table 2.

[0070] In certain aspects, this disclosure provides an *N. eutropha* bacterium (which is optionally axenic) comprising proteins that are mutant relative to *N. eutropha* strain C91 at all of the amino acid positions listed in Table 2.

[0071] In certain aspects, this disclosure provides an *N. eutropha* bacterium (which is optionally axenic) of strain D23, 25 vials of said bacterium, designated AOB D23-100, having been deposited with the ATCC patent depository on Apr. 8, 2014 under ATCC accession number PTA-121157.

[0072] In embodiments, the *N. eutropha* bacterium (which is optionally axenic) is transgenic.

[0073] In embodiments, the *N. eutropha* bacterium (which is optionally axenic) has at least one property selected from an optimized growth rate, an optimized NH_4^+ oxidation rate, and an optimized resistance to NH_4^+ .

[0074] In embodiments, the *N. eutropha* bacterium (which is optionally axenic) has at least two properties selected from an optimized growth rate, an optimized NH_4^+ oxidation rate, and an optimized resistance to NH_4^+ .

[0075] In embodiments, the *N. eutropha* bacterium (which is optionally axenic) has an optimized growth rate, an optimized NH_4^+ oxidation rate, and an optimized resistance to NH_4^+ .

[0076] In embodiments, the *N. eutropha* bacterium as described herein (e.g., strain D23) is substantially free of

bacteria, other ammonia oxidizing bacteria, fungi, viruses, or pathogens (e.g., animal pathogens, e.g., human pathogens), or any combination thereof.

[0077] In certain aspects, this disclosure provides a composition comprising the *N. eutropha* bacterium as described herein (e.g., strain D23), wherein the composition is substantially free of other organisms.

[0078] In certain aspects, this disclosure provides a composition comprising the *N. eutropha* bacterium as described herein (e.g., strain D23) and further comprising a second organism (e.g., a second strain or species), wherein the composition is substantially free of other organisms (e.g., strains or species). In embodiments, the second organism is an ammonia oxidizing bacterium. In embodiments, the second organism is selected from the group consisting of *Nitrosomonas*, *Nitrosococcus*, *Nitrosospria*, *Nitrosocystis*, *Nitrosolobus*, *Nitrosovibrio*, *Lactobacillus*, *Streptococcus*, and *Bifidobacter*, and combinations thereof.

[0079] This disclosure also provides a composition comprising the *N. eutropha* bacterium as described herein (e.g., strain D23) and further comprising a second and a third organism (e.g., of other strains or species), wherein the composition is substantially free of other organisms (e.g., strains or species). This disclosure also provides a composition comprising the *N. eutropha* bacterium as described herein (e.g., strain D23) and further comprising 2, 3, 4, 5, 6, 7, 8, 9, or 10 other organisms (e.g., of other strains or species), wherein the composition is substantially free of other organisms (e.g., strains or species).

[0080] In some aspects, this disclosure provides a composition comprising a cell suspension of an actively dividing culture of *N. eutropha* bacteria having an OD600 of at least about 0.1, 0.15, 0.2, 0.25, 0.3, 0.35, 0.4, 0.45, 0.5, 0.6, 0.7, or 0.8, wherein the composition is substantially free of other organisms.

[0081] In some aspects, this disclosure provides a composition for topical administration, comprising the *N. eutropha* bacterium as described herein (e.g., strain D23) and a pharmaceutically or cosmetically acceptable excipient suitable for topical administration. In embodiments, the composition is substantially free of other organisms. In embodiments, the composition further comprises a second organism (e.g., of another strain or specie). In embodiments, the composition further comprises 2, 3, 4, 5, 6, 7, 8, 9, or 10 other organisms (e.g., of other strains or species). The second organism may be, for example, an ammonia oxidizing bacterium. In embodiments, the second organism is selected from the group consisting of *Nitrosomonas*, *Nitrosococcus*, *Nitrosospria*, *Nitrosocystis*, *Nitrosolobus*, *Nitrosovibrio*, *Lactobacillus*, *Streptococcus*, and *Bifidobacter*, and combinations thereof.

[0082] In embodiments, the composition is a powder, cosmetic, cream, stick, aerosol, salve, wipe, or bandage. In embodiments, the composition further comprises a moisturizing agent, deodorizing agent, scent, colorant, insect repellent, cleansing agent, or UV-blocking agent. In embodiments, the excipient is an anti-adherent, binder, coat, disintegrant, filler, flavor, color, lubricant, glidant, sorbent, preservative, or sweetener. In embodiments, the concentration of *N. eutropha* in the composition is about 10^{11} - 10^{13} CFU/L. In embodiments, the concentration of *N. eutropha* in the composition is about 10^9 CFU/ml. In embodiments, the mass ratio of *N. eutropha* to pharmaceutical excipient may be about 0.1 gram per liter to about 100 grams per liter. In

some embodiments, the mass ratio of *N. eutropha* to pharmaceutical excipient is 1 gram per liter.

[0083] In some aspects the composition and/or excipient may be in the form of one or more of a liquid, a solid, or a gel. For example, liquid suspensions may include, but are not limited to, water, saline, phosphate-buffered saline, or an ammonia oxidizing storage buffer. Gel formulations may include, but are not limited to, agar, silica, polyacrylic acid (for example Carbopol®), carboxymethyl cellulose, starch, guar gum, alginate or chitosan. In some embodiments, the formulation may be supplemented with an ammonia source including, but not limited to ammonium chloride or ammonium sulfate.

In some aspects, this disclosure provides a composition comprising at least about 10, 20, 50, 100, 200, 500, 1,000, 2,000, or 10,000 L, e.g., at about 10^{11} CFU/L, 10^{12} CFU/L, 10^{13} CFU/L of the *N. eutropha* bacterium as described herein (e.g., strain D23). In some embodiments, the composition is at a concentration of at least about 10^9 CFU/L, 10^{10} CFU/L, 10^{11} CFU/L, or 10^{12} CFU/L. In some aspects, this disclosure provides a composition comprising at least about 1, 2, 5, 10, 20, 50, 100, 200, or 500 g of the *N. eutropha* bacterium described herein, e.g., as a dry formulation such as a powder.

[0084] In some aspects, this disclosure provides an article of clothing comprising the *N. eutropha* as described herein (e.g., strain D23). In embodiments, the article of clothing is packaged. In embodiments, the article of clothing is packaged in a material that is resistant to gaseous exchange or resistant to water. The article of clothing may be provided, e.g., at a concentration that provides one or more of a treatment or prevention of a skin disorder, a treatment or prevention of a disease or condition associated with low nitrite levels, a treatment or prevention of body odor, a treatment to supply nitric oxide to a subject, or a treatment to inhibit microbial growth.

[0085] In some aspects, this disclosure provides a cloth comprising the *N. eutropha* as described herein (e.g., strain D23).

[0086] In some aspects, this disclosure provides a yarn comprising the *N. eutropha* as described herein (e.g., strain D23).

[0087] In some aspects, this disclosure provides a thread comprising the *N. eutropha* as described herein (e.g., strain D23).

[0088] In some aspects, this disclosure provides a method of obtaining, e.g., manufacturing, an (optionally axenic) *N. eutropha* bacterium having an optimized growth rate, an optimized NH_4^+ oxidation rate, or an optimized resistance to NH_4^+ , comprising:

[0089] (a) culturing the bacterium under conditions that select for one or more of an optimized growth rate, an optimized NH_4^+ oxidation rate, or an optimized resistance to NH_4^+ , thereby producing a culture;

[0090] (b) testing a sample from the culture for an optimized growth rate, an optimized NH_4^+ oxidation rate, or an optimized resistance to NH_4^+ ; and

[0091] (c) repeating the culturing and testing steps until a bacterium having an optimized growth rate, an optimized NH_4^+ oxidation rate, or an optimized resistance to NH_4^+ is obtained.

[0092] In embodiments, the method comprises a step of obtaining an *N. eutropha* bacterium from a source, such as soil or the skin of an individual. In embodiments, culturing

the bacterium under conditions that select for one or more (e.g., 2 or 3) of an optimized growth rate, an optimized NH_4^+ oxidation rate, or an optimized resistance to NH_4^+ comprises culturing the bacterium in *N. europaea* medium that comprises about 200 mM NH_4^+ . In embodiments, the method comprises a step of creating an axenic culture. In embodiments, the method comprises a step of co-culturing the *N. eutropha* together with at least one other type of ammonia oxidizing bacteria. In embodiments, the *N. eutropha* of step (a) lack an optimized growth rate, an optimized NH_4^+ oxidation rate, and an optimized resistance to NH_4^+ . In embodiments, step (c) comprises repeating the culturing and testing steps until a bacterium having at least two of an optimized growth rate, an optimized NH_4^+ oxidation rate, and an optimized resistance to NH_4^+ is obtained.

[0093] In some aspects, this disclosure provides an *N. eutropha* bacterium as described herein (e.g., strain D23), produced by the methods described above.

[0094] In some aspects, this disclosure provides a method of testing a preparation of (optionally axenic) *N. eutropha*, comprising:

[0095] assaying the *N. eutropha* for one or more of an optimized growth rate, an optimized NH_4^+ oxidation rate, or an optimized resistance to NH_4^+ ; and

[0096] if the *N. eutropha* has one or more of an optimized growth rate, an optimized NH_4^+ oxidation rate, or an optimized resistance to NH_4^+ , classifying the *N. eutropha* as accepted.

[0097] In embodiments, the method further comprises a step of testing the preparation for contaminating organisms. In embodiments, the method further comprises a step of removing a sample from the preparation and conducting testing on the sample. In embodiments, the method further comprises testing medium in which the *N. eutropha* is cultured. In embodiments, the method further comprises packaging *N. eutropha* from the preparation into a package. In embodiments, the method further comprises placing *N. eutropha* from the preparation into commerce.

[0098] In some aspects, this disclosure provides a method of producing, e.g., manufacturing *N. eutropha*, comprising contacting *N. eutropha* with culture medium and culturing the *N. eutropha* until an OD600 of at least about 0.3, 0.4, 0.5, 0.6, 0.7, or 0.8 is reached. In some embodiments, the method comprises culturing the *N. eutropha* until an OD600 of at about 0.3-0.4, 0.4-0.5, 0.5-0.6, 0.6-0.7, or 0.7-0.8 is reached.

[0099] In embodiments, the method further comprises assaying the *N. eutropha* and culture medium for contaminating organisms. In embodiments, the method further comprises assaying the *N. eutropha* for one or more (e.g., 2 or 3) of an optimized growth rate, an optimized NH_4^+ oxidation rate, or an optimized resistance to NH_4^+ . In embodiments, the method comprises producing at least about 10, 20, 50, 100, 200, 500, 1,000, 2,000, 5,000, or 10,000 L per day of *N. eutropha*, e.g., at about 10^{12} CFUs/L. In some embodiments, the *N. eutropha* is at a concentration of about 10^9 , 10^{10} , 10^{11} , 10^{12} , 10^{13} , or 10^{14} CFUs/L. In some embodiments, the *N. eutropha* is at a concentration of least about 10^9 , 10^{10} , 10^{11} , 10^{12} , 10^{13} , or 10^{14} CFUs/L.

[0100] In some aspects, this disclosure provides a method of producing, e.g., manufacturing *N. eutropha*, comprising contacting *N. eutropha* with culture medium and culturing the *N. eutropha* until about at least about 1,000 L at about 10^{12} CFU/L *N. eutropha* are produced.

[0101] In embodiments, the method further comprises a step of assaying the *N. eutropha* for one or more (e.g., 2 or 3) of an optimized growth rate, an optimized NH_4^+ oxidation rate, or an optimized resistance to NH_4^+ .

[0102] In embodiments, the method further comprises a step of testing the *N. eutropha* or culture medium for contaminating organisms. In embodiments, the *N. eutropha* brought into contact with the culture medium is an *N. eutropha* having one or more (e.g., 2 or 3) of an optimized growth rate, an optimized NH_4^+ oxidation rate, or an optimized resistance to NH_4^+ .

[0103] In some aspects, this disclosure provides a method of producing, e.g., manufacturing *N. eutropha*, comprising:

[0104] (a) contacting *N. eutropha* with a culture medium; and

[0105] (b) culturing the *N. eutropha* for 1-2 days, thereby creating a culture, until the culture reaches an OD600 of about 0.5-0.6.

[0106] In embodiments, the method further comprises a step of assaying the *N. eutropha* for one or more of an optimized growth rate, an optimized NH_4^+ oxidation rate, or an optimized resistance to NH_4^+ . In embodiments, the method further comprises a step of testing the culture for contaminating organisms, e.g., bacteria, viruses, fungi, or pathogens, or a combination thereof. In embodiments, the *N. eutropha* of step (a) is an *N. eutropha* having one or more (e.g., 2 or 3) of an optimized growth rate, an optimized NH_4^+ oxidation rate, or an optimized resistance to NH_4^+ . In embodiments, the method comprises producing at least about 1,000 L per day at about 10^{12} CFUs/L of *N. eutropha*.

[0107] In some aspects, this disclosure provides a *N. eutropha* bacterium produced by the methods described above.

[0108] In embodiments, a preparation of *N. eutropha* made by the methods described above. In some aspects, the preparation may comprise about 0.1 milligrams to about 100 milligrams (mg) of *N. eutropha*.

[0109] In some aspects, a reaction mixture may be provided comprising *N. eutropha* at an optical density of about 0.5 to about 0.6. In some aspects, this disclosure provides a method of producing *N. eutropha*-bearing clothing, comprising contacting an article of clothing with of the *N. eutropha* as described herein (e.g., strain D23).

[0110] In embodiments, the method comprises producing at least 10, 100, or 1000 articles of clothing. In embodiments, the method comprises contacting the article of clothing with at least 10^{10} CFUs of *N. eutropha*. In embodiments, the method further comprises packaging the clothing.

[0111] In certain aspects, the present disclosure provides a method of obtaining a formulation of *N. eutropha*, combining contacting *N. eutropha* described herein (e.g., strain D23) with a pharmaceutically or cosmetically acceptable excipient.

[0112] In embodiments, the method further comprises mixing the *N. eutropha* and the excipient. In embodiments, the method is performed under conditions that are substantially free of contaminating organisms, e.g., bacteria, viruses, fungi, or pathogens.

[0113] In certain aspects, the present disclosure provides a method of packaging *N. eutropha*, comprising assembling *N. eutropha* described herein (e.g., strain D23) into a package.

[0114] In embodiments, the package is resistant to gaseous exchange or resistant to water. In embodiments, the package is permeable to gaseous exchange, NH_3 , NH_4^+ , or NO_2^- .

[0115] In certain aspects, the present disclosure provides a method of inhibiting microbial growth on a subject's skin, comprising topically administering to a subject in need thereof an effective dose of the *N. eutropha* bacteria described herein (e.g., strain D23).

[0116] In embodiments, the effective dose is approximately 1×10^9 CFU, 2×10^9 CFU, 5×10^9 CFU, 1×10^{10} CFU, 1.5×10^{10} CFU, 2×10^{10} CFU, 5×10^{10} CFU, or 1×10^{11} CFU. In embodiments, the effective dose is at least about 1×10^9 CFU, 2×10^9 CFU, 5×10^9 CFU, 1×10^{10} CFU, 1.5×10^{10} CFU, 2×10^{10} CFU, 5×10^{10} CFU, or 1×10^{11} CFU. In embodiments, the effective dose is approximately 1×10^9 CFU- 2×10^9 CFU, 2×10^9 CFU- 5×10^9 CFU, 5×10^9 CFU- 1×10^{10} CFU, 1×10^{10} CFU- 1.5×10^{10} CFU, 1×10^{10} CFU- 2×10^{10} CFU, 1.5×10^{10} CFU- 2×10^{10} CFU, 2×10^{10} CFU- 5×10^{10} CFU, or 5×10^{10} CFU- 1×10^{11} CFU. In embodiments, the bacterium is administered at a concentration of about 1×10^8 , 2×10^8 , 5×10^8 , 1×10^9 , 2×10^9 , 5×10^9 , or 1×10^{10} CFU/ml. In embodiments, the bacterium is administered at a concentration of at least about 1×10^8 , 2×10^8 , 5×10^8 , 1×10^9 , 2×10^9 , 5×10^9 , or 1×10^{10} CFU/ml. In embodiments, the bacterium is administered at a concentration of about 1×10^8 - 2×10^8 , 2×10^8 - 5×10^8 , 5×10^8 - 1×10^9 , 1×10^9 - 2×10^9 , 2×10^9 - 5×10^9 , or 5×10^9 - 1×10^{10} CFU/ml. In embodiments, the administration is performed twice per day. In embodiments, the subject is a human. In embodiments, the microbial growth to be inhibited is growth of *Pseudomonas aeruginosa* or *Staphylococcus aureus* (*S. aureus* or SA), *Streptococcus pyogenes* (*S. pyogenes* or SP), or *Acinetobacter baumannii* (*A. baumannii* or AB).

[0117] In certain aspects, the present disclosure provides a method of supplying nitric oxide to a subject, comprising positioning an effective dose of the *N. eutropha* bacteria described herein (e.g., strain D23) in close proximity to the subject.

[0118] In certain aspects, the present disclosure provides a method of reducing body odor, comprising topically administering to a subject in need thereof an effective dose of the *N. eutropha* bacteria described herein (e.g., strain D23).

[0119] In certain aspects, the present disclosure provides a method of treating a disease associated with low nitrite levels, comprising topically administering to a subject in need thereof a therapeutically effective dose of the *N. eutropha* bacteria described herein (e.g., strain D23).

[0120] In embodiments, the disease is HIV dermatitis, infection in a diabetic foot ulcer, atopic dermatitis, acne, e.g., acne vulgaris, eczema, contact dermatitis, allergic reaction, psoriasis, skin infections, vascular disease, vaginal yeast infection, a sexually transmitted disease, heart disease, atherosclerosis, baldness, leg ulcers secondary to diabetes or confinement to bed, angina, particularly chronic, stable angina pectoris, ischemic diseases, congestive heart failure, myocardial infarction, ischemia reperfusion injury, laminitis, hypertension, hypertrophic organ degeneration, Raynaud's phenomenon, fibrosis, fibrotic organ degeneration, allergies, autoimmune sensitization, end stage renal disease, obesity, impotence, or cancer.

[0121] In certain aspects, the present disclosure provides a method of treating a skin disorder, comprising topically administering to a subject in need thereof a therapeutically effective dose of the *N. eutropha* bacteria as described herein (e.g., strain D23). In related aspects, the disclosure provides

an *N. eutropha* bacteria as described herein (e.g., strain D23) for treating a disorder such as a skin disorder. In related aspects, the disclosure provides an *N. eutropha* bacteria as described herein (e.g., strain D23) for the manufacture of a medicament, e.g., a medicament for treating a skin disorder.

[0122] In embodiments, the skin disorder is acne, e.g., acne vulgaris, rosacea, eczema, or psoriasis. In some embodiments, the skin disorder is an ulcer, e.g., venous ulcer, e.g., leg ulcer, e.g., venous leg ulcer, e.g., infection in a diabetic foot ulcer. In some embodiments, topically administering comprises pre-treating the subject with *N. eutropha*, e.g., an *N. eutropha* described herein. In some embodiments, topically administering comprises topically administering prior to occurrence of the skin disorder. In some embodiments, topically administering comprises topically administering subsequent to occurrence of the skin disorder.

[0123] In certain aspects, the present disclosure provides a method of promoting wound healing or closure, comprising administering to a wound an effective dose of the *N. eutropha* bacteria as described herein (e.g., strain D23). In related aspects, the disclosure provides an *N. eutropha* bacteria as described herein (e.g., strain D23) for promoting wound healing. In related aspects, the disclosure provides an *N. eutropha* bacteria as described herein (e.g., strain D23) for the manufacture of a medicament, e.g., a medicament for promoting wound healing.

[0124] In embodiments, the wound comprises one or more undesirable bacteria, e.g., pathogenic bacteria. In embodiments, the wound comprises *S. aureus*, *P. aeruginosa*, *P. aeruginosa*, or *A. baumannii*.

[0125] In embodiments, the *N. eutropha* is administered to the subject prior to occurrence of the wound. In embodiments, administering to the wound comprises administering to the subject prior to occurrence of the wound. In embodiments, the method further comprises administering *N. eutropha* (e.g., an *N. eutropha* described herein, e.g., strain D23) to the wound subsequent to occurrence of the wound. In some aspects, the disclosure provides a method of killing or inhibiting growth of pathogenic bacteria comprising contacting, e.g., applying, *N. eutropha* bacteria (e.g., *N. eutropha* described herein, e.g., strain D23) to the skin.

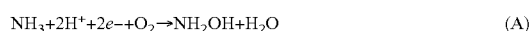
[0126] In embodiments, the pathogenic bacteria contribute to one or more of the following conditions: HIV dermatitis, an ulcer, e.g., venous ulcer, e.g., leg ulcer, e.g., venous leg ulcer, e.g., infection in a diabetic foot ulcer, atopic dermatitis, acne, e.g., acne vulgaris, eczema, contact dermatitis, allergic reaction, psoriasis, urticaria, rosacea, skin infections, vascular disease, vaginal yeast infection, a sexually transmitted disease, heart disease, atherosclerosis, baldness, leg ulcers secondary to diabetes or confinement to bed, angina, particularly chronic, stable angina pectoris, ischemic diseases, congestive heart failure, myocardial infarction, ischemia reperfusion injury, laminitis, hypertension, hypertrophic organ degeneration, Raynaud's phenomenon, fibrosis, fibrotic organ degeneration, allergies, autoimmune sensitization, end stage renal disease, obesity, impotence, pneumonia, primary immunodeficiency, epidermal lysis bulosa, or cancer.

[0127] In embodiments, the condition is an ulcer, e.g., venous ulcer, e.g., leg ulcer, e.g., venous leg ulcer, e.g., infection in a diabetic foot ulcer. In embodiments, the condition is a venous leg ulcer. In embodiments, the condition is acne, e.g., acne vulgaris. In embodiments, the condition is acne vulgaris. In embodiments, the pathogenic

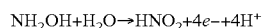
bacteria is one or more of *Propionibacterium acnes*, *Pseudomonas aeruginosa*, *Staphylococcus aureus*, *Streptococcus pyogenes*, or *Acinetobacter baumannii*. In embodiments, the method further comprises determining whether the subject is in need of killing or inhibiting growth of pathogenic bacteria, e.g., determining that the subject is in need of killing or inhibiting growth of pathogenic bacteria. In embodiments, the method further comprises selecting the subject in need of killing or inhibiting growth of pathogenic bacteria.

[0128] In some embodiments, the *N. eutropha* catalyze the following reactions.

[0129] At a neutral pH, ammonia generated from ammonium around neutral pH conditions is the substrate of the initial reaction. The conversion of ammonia to nitrite takes place in two steps catalyzed respectively by ammonia monooxygenase (Amo) and hydroxylamine oxidoreductase (Hao), as follows:



[0130] In some instances, reaction B is reported as follows, to indicate nitrous acid (HNO_2) formation at low pH:



[0131] In certain embodiments, the *N. eutropha* has a doubling time of less than 4, 5, 6, 7, 8, 9, or 10 hours, for instance about 8 hours, e.g., 7-9 hours or 6-10 hours, when grown under batch culture conditions. In some embodiments, the doubling time is at least 3, 4, 5, or 6 hours under batch culture conditions. In some embodiments, the *N. eutropha* has a doubling time of less than 16, 18, 20, 22, 24, or 26 hours, for instance about 20 hours, e.g., 19-21 hours or 18-22 hours, when grown under chemostat (i.e., continuous culture) conditions. In some embodiments, the doubling time is at least 10, 12, 14, 16, or 18 hours under chemostat conditions.

[0132] In certain embodiments, a continuous culture of *N. eutropha* at an OD600 of about 0.15-0.18 is capable of reaching an OD600 of about 0.5-0.6 in about 1-2 days. For instance, in some embodiments, a continuous culture of *N. eutropha* may grow from an OD600 of about 0.15 to at least 0.3, 0.4, 0.5, 0.6, 0.7, or 0.8 over about 1 day; in embodiments the culture may reach an OD in range of 0.4-0.6 or 0.3-0.7 over about 1 day. In embodiments, the continuous culture of *N. eutropha* may grow from an OD600 of about 0.15 to at least 0.3, 0.4, 0.5, 0.6, 0.7, or 0.8 over about 2 days; in embodiments the culture may reach an OD in the range of 0.4-0.6 or 0.3-0.7 over about 2 days. In some embodiments, the continuous culture conditions comprise growth in a bioreactor in *N. europaea* medium, optionally comprising about 200 mM NH_4^+ . In some embodiments, the continuous culture conditions are conditions set out in Example 2.

[0133] In certain embodiments, the *N. eutropha* are capable of converting NH_4^+ (e.g., at about 200 mM) to nitrite (e.g., reaching up to about 180 mM) at a rate of at least about 50, 75, 125, or 150 micromoles NO_2^- per minute, e.g., about 100-150, 75-175, 100-125, 125-150, or 125-175 micromoles/minute, e.g., about 125 micromoles NO_2^- per minute. In some embodiments, the reaction rates are measured in an about 1 L chemostat culture of about 10^9 CFU/ml over the course of 24 hours.

[0134] In certain embodiments, the *N. eutropha* are capable of growing in medium comprising at least 50 mM, 75 mM, 100 mM, 125 mM, 150 mM, 175 mM, 200 mM, 225 mM, 250 mM, 275 mM, or 300 mM NH_4^+ (or NH_3), e.g., about 150-200, 175-225, 200-250, 225-275, 250-300 mM, e.g., about 200 or about 250 mM. In certain embodiments, the *N. eutropha* is grown in a bioreactor under these concentrations of ammonium. In some embodiments, when the *N. eutropha* is grown under these concentrations of ammonium, the concentration of nitrate or nitrite is capable of reaching at least 60, 80, 100, 120, 140, 160, or 180 mM, e.g., about 140-180, 160-200, or 140-200 mM, e.g., about 160 or 180 mM.

[0135] In certain aspects, the present disclosure provides high density cultures of *N. eutropha*, e.g., *N. eutropha* strain D23. For instance, the high density culture composition may comprise a cell suspension of an actively dividing culture of *N. eutropha* bacteria having an OD600 of at least about 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, or 0.7, e.g., about 0.2-0.6, 0.3-0.6, 0.4-0.6, 0.5-0.6, or 0.4-0.7, wherein the composition is substantially free of other organisms

[0136] In some embodiments, the *N. eutropha* are stable for at least 2 weeks, 1 month, 2 months, 3 months, 4 months, 5 months, or 6 months when stored at 4° C. In some embodiments, the method of storage comprises resuspending the cells in a buffer comprising one or more of Na_2HPO_4 and MgCl_2 , for instance 50 mM Na_2HPO_4 and 2 mM MgCl_2 , for instance the storage buffer described in Example 2. For example, the storage conditions may be those specified in Example 2. In some embodiments, the *N. eutropha* are continuously cultured at 200 mM NH_4^+ at a pH of 6-8, e.g., 7, before storage at 4°. Stability can include one or more of 1) retaining viability, 2) retaining a relevant property such as the ability to produce a given level of nitrite.

[0137] In certain embodiments, NH_4^+ and NH_3 may be used interchangeably throughout the disclosure.

[0138] This disclosure provides, inter alia, a method of changing a composition of a skin microbiome of a subject. The method comprises administering, e.g., applying, a preparation comprising ammonia oxidizing bacteria to a surface of the skin, wherein the amount and frequency of administration, e.g., application, is sufficient to reduce the proportion of pathogenic bacteria on the surface of the skin.

[0139] Ammonia oxidizing bacteria are, in some embodiments, ubiquitous Gram-negative obligate chemolithoautotrophic bacteria with a unique capacity to generate energy exclusively from the conversion of ammonia to nitrite.

[0140] In some embodiments, the method may further comprise, selecting the subject on the basis of the subject being in need of a reduction in the proportion of pathogenic bacteria on the surface of the skin.

[0141] In some embodiments, the preparation comprising ammonia oxidizing bacteria comprises at least one of ammonia, ammonium salts, and urea.

[0142] In some embodiments, the preparation comprising ammonia oxidizing bacteria comprises a controlled release material, e.g., slow release material.

[0143] In some embodiments, the preparation of ammonia oxidizing bacteria, comprises an excipient, e.g., one of a pharmaceutically acceptable excipient or a cosmetically acceptable excipient. The excipient, e.g., one of the pharmaceutically acceptable excipient and the cosmetically acceptable excipient, may be suitable for one of topical, nasal, pulmonary, and gastrointestinal administration. The

excipient, e.g., one of the pharmaceutically acceptable excipient and the cosmetically acceptable excipient may be a surfactant. The surfactant may be selected from the group consisting of cocamidopropyl betaine (ColaTeric COAB), polyethylene sorbitol ester (e.g., Tween 80), ethoxylated lauryl alcohol (RhodaSurf 6 NAT), sodium laureth sulfate/lauryl glucoside/cocamidopropyl betaine (Plantapon 611 L UP), sodium laureth sulfate (e.g., RhodaPex ESB 70 NAT), alkyl polyglucoside (e.g., Plantaren 2000 N UP), sodium laureth sulfate (Plantaren 200), Dr. Bronner's Castile soap, lauramine oxide (ColaLux Lo), sodium dodecyl sulfate (SDS), polysulfonate alkyl polyglucoside (PolySufanate 160 P), sodium lauryl sulfate (Stepanol-WA Extra K), and any combination thereof. Dr. Bronner's Castile soap comprises water, organic coconut oil, potassium hydroxide, organic olive oil, organic fair deal hemp oil, organic jojoba oil, citric acid, and tocopherol. In some embodiments, the excipient comprises one or more of, e.g., all of, water, organic coconut oil, potassium hydroxide, organic olive oil, organic fair deal hemp oil, organic jojoba oil, citric acid, and tocopherol.

[0144] In some embodiments, the preparation may be substantially free of other organisms.

[0145] In some embodiments, the preparation may be disposed in a powder, cosmetic, cream, stick, aerosol, salve, wipe, or bandage. The preparation may be provided as a powder, cosmetic, cream, stick, aerosol, salve, wipe, or bandage.

[0146] In some embodiments, the preparation may comprise a moisturizing agent, deodorizing agent, scent, colorant, insect repellent, cleansing agent, or UV-blocking agent.

[0147] In some embodiments, the excipient, e.g., the pharmaceutically acceptable excipient or the cosmetically acceptable excipient may comprise an anti-adherent, binder, coat, disintegrant, filler, flavor, color, lubricant, glidant, sorbent, preservative, or sweetener.

[0148] In some embodiments, the preparation comprising ammonia oxidizing bacteria may comprise between about 10^8 and about 10^{14} CFU/L. In certain aspects, the preparation may comprise between about 1×10^9 CFU/L and about 10×10^9 CFU/L.

[0149] In some embodiments, the preparation comprising ammonia oxidizing bacteria may comprise between about 50 milligrams (mg) and about 1000 mg of ammonia oxidizing bacteria.

[0150] In some embodiments, the mass ratio of ammonia oxidizing bacteria to the excipient, e.g., the pharmaceutically acceptable excipient or the cosmetically acceptable excipient is in a range of about 0.1 grams per liter to about 1 gram per liter.

[0151] In some embodiments, the preparation of ammonia oxidizing bacteria are useful in the treatment or prevention of a disease or condition associated with low nitrite levels, a treatment or prevention of body odor, a treatment to supply nitric oxide to a subject, or a treatment to inhibit microbial growth, e.g., pathogenic bacterial growth.

[0152] In some embodiments, the ammonia oxidizing bacteria is selected from the group consisting of *Nitrosomonas*, *Nitrosococcus*, *Nitrosospria*, *Nitrosocystis*, *Nitrosolobus*, *Nitrosovibrio*, and combinations thereof. The preparation may further comprise an organism selected from the group consisting of *Lactobacillus*, *Streptococcus*, *Bifidobacter*, and combinations thereof. In certain aspects, the preparation is substantially free of organisms other than ammonia oxidizing bacteria.

[0153] In some embodiments, the preparation comprising ammonia oxidizing bacteria may comprise ammonia oxidizing bacteria in a growth state. In some embodiments, the preparation comprising ammonia oxidizing bacteria may comprise ammonia oxidizing bacteria in a storage state.

[0154] In some embodiments, the methods of the present disclosure may be used to deliver a cosmetic product. In some embodiments, the methods of the present disclosure may be used to deliver a therapeutic product. The preparation may be useful for treatment of at least one of HIV dermatitis, infection in a diabetic foot ulcer, atopic dermatitis, acne, e.g., acne vulgaris, eczema, contact dermatitis, allergic reaction, psoriasis, urticaria, rosacea, skin infections, vascular disease, vaginal yeast infection, a sexually transmitted disease, heart disease, atherosclerosis, baldness, leg ulcers secondary to diabetes or confinement to bed, angina, particularly chronic, stable angina pectoris, ischemic diseases, congestive heart failure, myocardial infarction, ischemia reperfusion injury, laminitis, hypertension, hypertrophic organ degeneration, Raynaud's phenomenon, fibrosis, fibrotic organ degeneration, allergies, autoimmune sensitization, end stage renal disease, obesity, impotence, pneumonia, primary immunodeficiency, epidermal lysis bulosa, or cancer.

[0155] In certain aspects, the preparation may be useful for treatment of at least one of acne, e.g., acne vulgaris, eczema, psoriasis, urticaria, rosacea, and skin infections.

[0156] In some embodiments, the preparation may be provided in a container, the preparation and the container having a weight of less than about 50, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1500, or 2000 grams.

[0157] In some embodiments, the preparation has less than about 0.1% to about 10% of surfactant. In certain aspects, the preparation may be substantially free of surfactant.

[0158] In some embodiments, the preparation may comprise a chelator. In some embodiments, the preparation may be substantially free of a chelator.

[0159] In some embodiments, the method may comprise applying the preparation about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, or 24 times per day. In certain aspects, the preparation may be applied one time per day. In certain other aspects, the preparation may be applied two times per day.

[0160] In some embodiments, the preparation may be applied for about 1-3, 3-5, 5-7, 7-9, 5-10, 10-14, 12-18, 12-21, 21-28, 28-35, 35-42, 42-49, 49-56, 46-63, 63-70, 70-77, 77-84, or 84-91 days. In certain aspects, the preparation may be applied for about 16 days.

[0161] In some embodiments, the method may further comprise obtaining a sample from the surface of the skin. In certain aspects, the method may further comprise isolating DNA of bacteria in the sample. In certain aspects, the method may further comprise sequencing DNA of bacteria in the sample.

[0162] In some embodiments, administering the ammonia oxidizing bacteria provides for an increase in the proportion of non-pathogenic bacteria on the surface. In certain aspects, the non-pathogenic bacteria may be commensal non-pathogenic bacteria. In certain aspects, the non-pathogenic bacteria is commensal non-pathogenic bacteria of a genus of *Staphylococcus*. In certain aspects, the non-pathogenic bacteria may be commensal non-pathogenic bacteria *Staphylococcus epidermidis*.

[0163] In some embodiments, the proportion of non-pathogenic bacteria *Staphylococcus* is, or is identified as being, increased after about 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 weeks. In certain aspects, the proportion of non-pathogenic bacteria *Staphylococcus epidermidis* *Staphylococcus* is, or is identified as being, increased after about 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 weeks.

[0164] In some embodiments, potentially pathogenic or disease associated Propionibacteria is, or is identified as being, reduced after about 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 weeks.

[0165] In some embodiments, potentially pathogenic or disease associated *Stenotrophomonas* is, or is identified as being, reduced after about 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 weeks.

[0166] In some embodiments, the surface of the skin comprises a wound.

[0167] In some embodiments, a method of treating acne e.g., acne vulgaris, may be provided by one or more methods of the present disclosure. In some embodiments, a method of treating eczema may be provided by one or more methods of the present disclosure. In some embodiments, a method of treating psoriasis may be provided by one or more methods of the present disclosure. In some embodiments, a method of treating urticaria may be provided by one or more methods of the present disclosure. In some embodiments, a method of treating rosacea may be provided by one or more methods of the present disclosure. In some embodiments, a method of treating skin infection may be provided by one or more methods of the present disclosure. In some embodiments, a method of reducing an amount of undesirable bacteria on a surface of a subject is provided.

[0168] In some embodiments, the method herein (e.g., a method of administering a *N. eutropha* bacterium, e.g., a bacterium of strain D23 to a subject in need thereof), further comprise treating the subject with an antibiotic. In embodiments, the antibiotic is Tetracycline, a Lincosamide such as Clindamycin, a Macrolide such as Erythromycin, an Aminoglycoside such as Gentamicin, a β -lactam such as Piperacillin, β -lactamase inhibitor such as Tazobactam, or any combination thereof (such as a combination of a β -lactam such as Piperacillin and a β -lactamase inhibitor such as Tazobactam). In some embodiments, the antibiotic is an antibiotic to which the bacterium is sensitive. In embodiments, the antibiotic is administered after the bacterium has achieved the desired therapeutic effect. In embodiments, the antibiotic is an antibiotic to which the bacterium is resistant. In embodiments, the antibiotic is administered before or during the period in which the bacterium is producing its therapeutic effect.

[0169] It is understood that compositions and methods herein involving a bacterium can also involve a plurality of bacteria. For instance, a method of administering a *N. eutropha* bacterium can also involve administering a plurality of *N. eutropha* bacteria.

[0170] The present disclosure also provides, in certain aspects, a nucleic acid comprising a sequence of consecutive nucleotides (e.g., 15-100 nucleotides) from within the D23 genome, e.g., a sequence of a gene provided herein, e.g., a gene described in Table 1, FIG. 6-8 or Supplementary Table 1, or SEQ ID NO: 66, or a reverse complement of any of the foregoing. In a related aspect, the present disclosure provides a nucleic acid comprising a sequence of consecutive nucleotides (e.g., 15-100 nucleotides) from within SEQ ID

NO: 1 or a reverse complement thereof. In a related aspect, the present disclosure provides a nucleic acid comprising a sequence of consecutive nucleotides (e.g., 15-100 nucleotides) from within a gene of Table 1 (e.g., a sequence of SEQ ID NO: 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, or 33) or a reverse complement thereof.

[0171] In some embodiments, the nucleic acid has a non-naturally occurring sequence or another modification such as a label, or both. In some embodiments, the sequence of consecutive nucleotides is not a sequence found in *N. Eutropha* strain C91. In some embodiments, the nucleic acid comprises a heterologous sequence 5' to the sequence of 15-100 consecutive nucleotides, or a heterologous sequence 3' to the sequence of 15-100 consecutive nucleotides, or both. In some embodiments, the nucleic acid has a length of 10-15, 15-20, 20-25, 25-30, 30-24, 35-40 nucleotides. In some embodiments, the nucleic acid is bound, e.g., covalently bound, to a detectable label, e.g., a fluorescent label. In some embodiments, the nucleic acid comprises 10-15, 15-20, 20-25, 25-30, 30-24, 35-40, 40-50, 50-60, 60-70, 70-80, 80-90, or 90-100 consecutive nucleotides from within the D23 genome. In some embodiments, the nucleic acid comprises at least about 10, 15, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 150, 200, 250, 300, 350, 400, 450, 500, 600, 700, 800, 900, or 1000 consecutive nucleotides from within the D23 genome. In some embodiments, the nucleic acid is DNA.

[0172] In some aspects, the disclosure provides a composition or a kit comprising a first nucleic acid and a second nucleic acid. In some embodiments, the first nucleic acid comprises consecutive nucleotides (e.g., 15-100) from within SEQ ID NO: 1, SEQ ID NO: 66, a gene of FIGS. 6-8, or a gene of Table 1, or a reverse complement thereof. In some embodiments, the second nucleic acid comprises consecutive nucleotides (e.g., 15-100) from within SEQ ID NO: 1, SEQ ID NO: 66, a gene of FIGS. 6-8, or a gene of Table 1, or a reverse complement thereof. In some embodiments, the nucleic acid has a non-naturally occurring sequence, e.g., a sequence not found in *N. eutropha* strain C91. In some embodiments, the first nucleic acid and the second nucleic acid define an amplicon in a gene of Table 1, e.g., a sequence of SEQ ID NO: 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, or 33) or a reverse complement thereof.

[0173] In some embodiments, the first nucleic acid has a sequence that corresponds to a first region of SEQ ID NO: 1, and the reverse complement of the second nucleic acid has a sequence that corresponds to a second region of SEQ ID NO: 1, and the first and second regions are separated by a distance suitable for PCR. In some embodiments, the reverse complement of the first nucleic acid has a sequence that corresponds to a first region of SEQ ID NO: 1, and the second nucleic acid has a sequence that corresponds to a second region of SEQ ID NO: 1, and the first and second regions are separated by a distance suitable for PCR. In an embodiment, the distance suitable for PCR is no more than 50, 100, 150, 200, 250, 300, 350, 400, 450, 500, 600, 700, 800, 900, or 1000 nucleotides of SEQ ID NO: 1. In some embodiments, the first nucleic acid and second nucleic acid delineate an amplicon in SEQ ID NO: 1. In some embodiments, the first nucleic acid and second nucleic acid each has a melting temperature (T_m) suitable for PCR, e.g., about 55-65° or about 60-65° C. In some embodiments, the T_m of the first nucleic acid is within 10, 9, 8, 7, 6, 5, 4, 3, 2, or 1° C. of the T_m of the second nucleic acid.

[0174] In some embodiments, the first nucleic acid, the second nucleic acid, or each of the first nucleic acid and second nucleic acid further comprises a heterologous sequence 5' to the sequence of consecutive nucleotides. Alternatively or in combination, in some embodiments, the first nucleic acid, the second nucleic acid, or each of the first nucleic acid and second nucleic acid further comprises a heterologous sequence 3' to the sequence of consecutive nucleotides from within SEQ ID NO: 1 or SEQ ID NO: 66. In some embodiments, the first nucleic acid, the second nucleic acid, or each of the first nucleic acid and second nucleic acid has a length of 15-20, 20-25, 25-30, 30-24, or 35-40 nucleotides. In some embodiments, the first nucleic acid, the second nucleic acid, or each of the first nucleic acid and second nucleic acid is bound, e.g., covalently bound, to a detectable label, e.g., a fluorescent label. In some embodiments, the first nucleic acid comprises, or consists of, a sequence of SEQ ID NO: 64. In some embodiments, the second nucleic acid comprises, or consists of, a sequence of SEQ ID NO: 65. In some embodiments, the first nucleic acid, the second nucleic acid, or both, are DNA.

[0175] In some embodiments, the composition or kit comprises at least two (e.g., 3, 4, 5, 6, 7, 8, 9, or 10) pairs of primers, each pair recognizing an amplicon in a gene of Table 1 (e.g., a sequence of SEQ ID NO: 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, or 33) or a reverse complement thereof. In some embodiments, a first pair of primers recognizes an amplicon in an Amo gene (e.g., AmoA1, AmoA2, AmoB1, AmoB2, AmoC1, AmoC2, or AmoC3) and the second pair of primers recognizes an amplicon in an Amo gene (e.g., AmoA1, AmoA2, AmoB1, AmoB2, AmoC1, AmoC2, or AmoC3). In some embodiments, a first pair of primers recognizes an amplicon in an AmoA gene (e.g., AmoA1 or AmoA2). In some embodiments, a second pair of primers recognizes an amplicon in an AmoB gene (e.g., AmoB1 or AmoB2). In some embodiments, a third pair of primers recognizes an amplicon in an AmoC gene (e.g., AmoC1, AmoC2, or AmoC3).

[0176] In some embodiments, the kit comprises a first container in which the first nucleic acid is disposed and a second container in which the second nucleic acid is disposed. The kit may comprise additional containers, e.g., for a third, fourth, fifth, or sixth nucleic acid. In some embodiments, a pair of primers recognizing an amplicon is stored in a single container.

[0177] The present disclosure also provides, in some aspects, a nucleic acid comprising, or consisting of, the sequence of SEQ ID NO: 64. The present disclosure also provides, in some aspects, a nucleic acid comprising, or consisting of, the sequence of SEQ ID NO: 65. The present disclosure also provides, in some aspects, the present disclosure provides a molecule comprising a nucleic acid described herein and a detectable label, e.g., a fluorescent label. The nucleic acid may consist of a sequence of SEQ ID NO: 64 or SEQ ID NO: 65, for example.

[0178] The present disclosure provides, in some aspects, a composition comprising a first molecule and a second molecule. In some embodiments, the first molecule comprises a nucleic acid described herein, e.g., a nucleic acid consisting of the sequence of SEQ ID NO: 64, and optionally comprises a detectable label, e.g., a fluorescent label. In some embodiments, the second molecule comprises a nucleic acid described herein, e.g., a nucleic acid consisting of the

sequence of SEQ ID NO: 65, and optionally comprises a detectable label, e.g., a fluorescent label.

[0179] In some embodiments, the kit comprises a first container in which the first molecule is disposed and a second container in which the second molecule is disposed.

[0180] In some embodiments, a kit described herein further comprises one or more of a buffer, an enzyme (e.g., a polymerase such as a thermostable polymerase such as Taq), nucleotides (e.g., dNTPs), and chain-terminating nucleotides (e.g., dideoxy nucleotides) which are optionally dye-labeled; these components may be provided separately or as part of a single composition.

[0181] In certain aspects, this disclosure provides a method of detecting whether a D23 *N. eutropha* nucleic acid is present in a sample, comprising: performing a polymerase chain reaction (PCR) on the sample using primers specific to D23 *N. eutropha*, and determining whether a PCR product is produced, wherein the presence of a PCR product indicates that the D23 *N. eutropha* nucleic acid was present in the sample. In embodiments, at least two PCR reactions are performed, e.g., 3, 4, 5, 6, 7, 8, 9 or 10 PCR reactions. In embodiments, the PCR reactions are performed in separate reaction volumes. In embodiments, two or more PCR reactions are performed in multiplex.

[0182] In some embodiments, the primers specific to D23 *N. eutropha* are a first nucleic acid and second nucleic acid described herein, e.g., a first and second nucleic acid from a composition or kit described herein. In some embodiments, the first primer comprises or consists of a sequence of SEQ ID NO: 65, and the second primer comprises or consists of a sequence of SEQ ID NO: 66.

[0183] In some embodiments, the PCR reaction is a quantitative or real-time PCR reaction. In some embodiments, the PCR reaction comprises a TaqMan reaction. In some embodiments, the PCR reaction comprises cycling the temperature of a reaction mixture between a denaturing temperature (e.g., about 95° C.), an annealing temperature (e.g., 45-68, 55-65, or 60-65° C.), and an elongation temperature (e.g., about 68° C.) for a number of cycles sufficient to produce a detectable PCR product, e.g., about 10, 15, 20, 25, or 30 cycles. In some embodiments, detecting the PCR product comprises detecting fluorescence from the PCR product. In some embodiments, a positive control is performed, e.g., using a known D23 *N. eutropha* nucleic acid as a template. In some embodiments, a negative control is used, e.g., using no template or using another bacterial nucleic acid as a template.

[0184] In certain aspects, the disclosure provides a method of detecting whether a D23 *N. eutropha* nucleic acid is present in a sample, comprising detecting binding of a nucleic acid described herein to a sample, wherein the presence of binding indicates that the D23 *N. eutropha* nucleic acid was present in the sample. In some embodiments, binding is detected by primer extension or RNase protection.

[0185] In some embodiments of the methods herein, the sample comprises at least 2, 3, 4, 5, 6, 7, 8, 9, or 10 strains of bacteria. In some embodiments, the sample is from the skin of a subject, e.g., a human subject. In some embodiments, the methods herein comprise detecting one or more additional types of bacterium in the sample, e.g., *Pseudomonas aeruginosa*, *Staphylococcus aureus*, *Streptococcus pyogenes*, or *Acinetobacter baumannii*.

[0186] The disclosure contemplates all combinations of any one or more of the foregoing aspects and/or embodiments, as well as combinations with any one or more of the embodiments set forth in the detailed description and examples.

BRIEF DESCRIPTION OF THE DRAWINGS

[0187] FIG. 1 shows the growth of a mixed culture of bacteria comprising *N. eutropha* strain D23. The optical density at a 600 nm wavelength is plotted relative to time.

[0188] FIG. 2A shows the nitrite production of a mixed culture of bacteria comprising *N. eutropha* strain D23. The nitrite concentration is plotted relative to time.

[0189] FIG. 2B-I shows the nitrite production kinetics by *N. eutropha* D23 in batch culture. The nitrite concentration is plotted relative to time.

[0190] FIG. 2B-II shows the nitrite production kinetics by *N. eutropha* D23 in vitro. The nitrite concentration is plotted relative to time.

[0191] FIG. 2C shows *N. eutropha* D23 stability upon storage at 4° C. The nitrite concentration is plotted relative to time.

[0192] FIG. 3A shows the *N. eutropha* D23's ability to inhibit the growth of *P. aeruginosa* (left panel) and *S. aureus* (right panel) in co-culture experiments. The amount of each type of undesirable bacteria (in CFU/ml) is plotted relative to time. In this figure, "AOB" refers to strain D23.

[0193] FIG. 3B shows the *N. eutropha* D23's ability to inhibit the growth of *Streptococcus pyogenes* (left panel) and *Acinetobacter baumannii* (right panel) in co-culture experiments. The amount of each type of undesirable bacteria (in CFU/ml) is plotted relative to time. In this figure, "AOB" refers to strain D23.

[0194] FIG. 3C shows the *N. eutropha* D23's ability to inhibit the growth of *Propionibacterium acnes* in co-culture experiments. The amount of each type of undesirable bacteria (in CFU/ml) is plotted relative to time. In this figure, "AOB" refers to strain D23.

[0195] FIG. 4A (top panel) plots the NO₂⁻ concentration over time in a co-culture experiment. The bottom panel plots pH over time in a co-culture experiment.

[0196] FIG. 4B (top panels) plots the CFU/ml of the indicated bacteria over time in a co-culture experiment. The center panels plot the NO₂⁻ concentration over time in a co-culture experiment.

[0197] The bottom panels plot pH over time in a co-culture experiment.

[0198] FIG. 4C plots the microbicidal activity of D23 against skin pathogens.

[0199] FIG. 4D plots the microbicidal activity of D23 against skin pathogens.

[0200] FIG. 4E shows an alternative plot of microbicidal activity of D23 against skin pathogens.

[0201] FIG. 5A plots the percent wound closure over time in an experiment testing D23's ability to improve wound healing.

[0202] FIG. 5B plots CT₅₀ for various D23 treatments.

[0203] FIG. 5C plots the percent wound closure over time in an experiment testing D23's ability to improve wound healing.

[0204] FIG. 5D plots the percent wound closure over time in an experiment testing D23's ability to improve wound healing.

[0205] FIG. 5E plots CT₅₀ for various D23 treatments.

[0206] FIG. 5F shows images of D23 enhanced wound healing in diabetic mice at Day 1, Day 11, and Day 15.

[0207] FIG. 5G shows blood glucose measurements for various concentrations of D23.

[0208] FIG. 5H shows body weight of test subjects over the course of testing.

[0209] FIG. 5I shows body weight of test subjects over the course of testing.

[0210] FIG. 5J shows PCR scores for a scalp test of subjects. AOB refers to D23 in this Figure.

[0211] FIG. 5K shows a schematic of a human volunteer study for an evaluation of a *Nitrosomonas*-containing topical suspension (AOB-001).

[0212] FIG. 5L (left panel) shows PCR analyses of scalp swabs collected during the study. Percent-positive samples for AOB-specific three-gene signature (*amoA*, *amoB*, *amoC*). The right panel shows PCR analyses of scalp swabs collected during the study. Composite PCR scores for a total of six samples collected from each of 23 volunteers. The scoring scheme used for the positive samples collected at each of six sampling points is indicated.

[0213] FIG. 5M shows genus-level bacterial diversity as determined by 16S rDNA sequencing in skin swab samples collected before and after topical application of AOB-001. The percentage of the total sequence reads representing each of twelve bacterial genera in samples collected at baseline prior to application (Day 0) and immediately after the one week application (Day 8), or one week after stopping topical application (Day 14), are shown. The proportions of *Acinetobacter*, *Burkholderia*, *Enterobacter*, *Escherichia Shigella*, *Klebsiella*, *Nitrosomonas*, *Pantoea*, *Propionibacterium*, *Pseudomonas*, *Serratia*, *Staphylococcus*, and *Stenotrophomonas* are shown.

[0214] FIG. 5N-A shows changes in abundance of *Nitrosomonas* and other species in skin samples collected before and after AOB-001 application. The percentages of the total 16S rDNA sequence reads representing *Nitrosomonas* prior to application (Day 0), immediately after the one-week application (Day 8), or one week after terminating application (Day 14) are shown.

[0215] FIG. 5N-B shows changes in abundance of *Nitrosomonas* and other species in skin samples collected before and after AOB-001 application. Changed patterns in abundance of species were detected by 16S rDNA sequencing in Day 0 versus Day 8 samples collected from AOB users.

[0216] FIG. 5O shows user evaluation of AOB-001. Assessment of AOB-001 cosmetic effects as provided by 23 volunteers upon completion of the one week application to their scalp and face. Subjects were plotted in order of increasing composite PCT scores. (2=agree strongly; 0=no change; -2=disagree strongly).

[0217] FIG. 6 is a table displaying unique D23 genes that have either an assigned open reading frame (ORF) number and a function based on sequence analysis, or a hypothetical gene above 200 base pairs in length. The column headers signify as follows: Feature.ID=a unique identifier for the gene; Type=type of gene, where CDS indicates a protein-coding DNA sequence; Start=starting position of gene in the genome sequence of SEQ ID NO: 1; Stop=end of gene in the genome sequence of SEQ ID NO: 1; Frame=reading frame; Length=length of gene in base pairs; Function=gene or

protein function based on sequence analysis; Subsystem=category of gene function; D23GbklId=a gene identifier.

[0218] FIG. 7 is a table displaying unique D23 genes below 200 base pairs that have an assigned ORF number. Column headers are as described in FIG. 6.

[0219] FIG. 8 is a table displaying unique D23 genes with no assigned ORF number. Column headers are as described in FIG. 6.

[0220] FIG. 9 lists unique C91 genes that do not have a homolog in D23.

[0221] FIG. 10 is a sequence alignment between the *AmoA1* and *AmoA2* proteins in *N. eutropha* strains D23 and C91. The SEQ ID of each protein is listed in Table 1. FIG. 10 discloses SEQ ID NOS 6, 12, 36 and 42, respectively, in order of appearance.

[0222] FIG. 11 is a sequence alignment between the *AmoB1* and *AmoB2* proteins in *N. eutropha* strains D23 and C91. The SEQ ID of each protein is listed in Table 1. FIG. 11 discloses SEQ ID NOS 8, 14, 38 and 44, respectively, in order of appearance.

[0223] FIG. 12 is a sequence alignment between the *AmoC1* and *AmoC2* proteins in *N. eutropha* strains D23 and C91. The SEQ ID of each protein is listed in Table 1. FIG. 12 discloses SEQ ID NOS 34, 40, 10 and 4, respectively, in order of appearance.

[0224] FIG. 13 is a sequence alignment between the *AmoC3* proteins in *N. eutropha* strains D23 and C91. The SEQ ID of each protein is listed in Table 1. FIG. 13 discloses SEQ ID NOS 46 and 16, respectively, in order of appearance.

[0225] FIG. 14 A and FIG. 14 B show a sequence alignment between the *Hao1*, *Hao2*, and *Hao3* proteins in *N. eutropha* strains D23 and C91. The SEQ ID of each protein is listed in Table 1. FIG. 14 discloses SEQ ID NOS 20, 22, 18, 50, 52 and 48, respectively, in order of appearance.

[0226] FIG. 15 is a sequence alignment between the *cycA1*, *cycA2*, and *cycA3* genes in *N. eutropha* strains D23 and C91. The SEQ ID of each protein is listed in Table 1. FIG. 15 discloses SEQ ID NOS 26, 28, 24, 58, 56 and 54, respectively, in order of appearance.

[0227] FIG. 16 is a sequence alignment between the *cycB1* and *cycB2* genes in *N. eutropha* strains D23 and C91. The SEQ ID of each protein is listed in Table 1. FIG. 16 discloses SEQ ID NOS 30, 32, 60 and 62, respectively, in order of appearance.

[0228] FIG. 17 shows a bar graph of proportion of bacteria, by genus versus day.

[0229] FIG. 18 shows a bar graph of proportion of bacteria, by genus versus bacteria genus, for day 0, day 1, day 8, day 14, and day 16.

[0230] Supplementary Table 1 displays the genome annotation of 2,777 genes identified in strain D23 using sequence analysis. Column headers are as described in FIG. 6. "C91 Alias" refers to a homolog in strain C91. Supplementary Table 1 is appended to the end of the Detailed Description and Examples.

[0231] Supplementary Table 2 displays the sequences of selected proteins genes identified in strain D23. Supplementary Table 2 is appended to the end of the Detailed Description and Examples.

DETAILED DESCRIPTION

[0232] Ammonia-oxidizing bacteria (AOB) of the genus *Nitrosomonas* are Gram-negative obligate autotrophic bacteria with a unique capacity to generate nitrite and nitric oxide exclusively from ammonia as an energy source. They are widely present both in soil and water environments and are essential components of environmental nitrification processes. Due to the roles of nitrite and nitric oxide on human skin as important components of several physiological functions, such as vasodilation, skin inflammation and wound healing, these bacteria may have beneficial properties for both healthy and immunopathological skin conditions. These bacteria may be safe for use in humans because they are slow-growing, cannot grow on organic carbon sources, may be sensitive to soaps and antibiotics, and have never been associated with any disease or infection in animals or humans.

1. DEFINITIONS

[0233] An ammonia oxidizing bacterium refers to a bacterium capable of oxidizing ammonia or ammonium to nitrite at a rate, e.g., a substantial rate, e.g., a pre-determined rate, e.g., at least the rate depicted in any one of FIG. 2A, 2B, 2C, 4A, 4B, or 5 or at least 90%, 80%, 70%, 60%, 50%, 40%, 30%, 20%, or 10% of that rate. In some embodiments, the substantial rate refers to the conversion of ammonium ions (NH_4^+) (e.g., at about 200 mM) to nitrite (NO_2^-) at a rate of at least 50, 75, 125, or 150 micromoles NO_2^- per minute, e.g., about 100-150, 75-175, 75-125, 100-125, 125-150, or 125-175 micromoles/minute, e.g., about 125 micromoles NO_2^- per minute. Examples of ammonia oxidizing bacteria include *N. eutropha* strains D23 and C91, and other bacteria in the genera *Nitrosomonas*, *Nitrosococcus*, *Nitrospira*, *Nitrosocystis*, *Nitrosolobus*, and *Nitrosovibrio*. D23 *Nitrosomonas eutropha* strain refers to the strain, designated AOB D23-100, deposited with the American Tissue Culture Collection (ATCC) on Apr. 8, 2014 having accession number PTA-121157. The D23 *Nitrosomonas eutropha* of accession number PTA-121157 has a genome sequence as set out in SEQ ID NO: 1 herein. The nucleic acid sequence(s), e.g., genome sequence, of accession number PTA-121157 are hereby incorporated by reference in their entireties.

[0234] Optimized *Nitrosomonas eutropha* (*N. eutropha*), as that term is used herein, refers to an *N. eutropha* having an optimized growth rate; an optimized NH_4^+ oxidation rate; or optimized resistance to NH_4^+ . In an embodiment it differs from naturally occurring *N. eutropha* by at least one nucleotide, e.g., a nucleotide in a gene selected from ammonia monooxygenase, hydroxylamine oxidoreductase, cytochrome c554, and cytochrome c₅₅₂. The difference can arise, e.g., through selection of spontaneously arising mutation, induced mutation, or directed genetic engineering, of the *N. eutropha*. In an embodiment it differs from a naturally occurring *N. eutropha* in that it has a constellation of alleles, not present together in nature. These differences may provide for one or more of a treatment or prevention of a skin disorder, a treatment or prevention of a disease or condition associated with low nitrite levels, a treatment or prevention of body odor, a treatment to supply nitric oxide to a subject, and a treatment to inhibit microbial growth.

[0235] As used herein, “axenic” refers to a composition comprising an organism that is substantially free of other

organisms. For example, an axenic culture of ammonia oxidizing bacteria is a culture that is substantially free of organisms other than ammonia oxidizing bacteria. For example, an axenic culture of *N. eutropha* is a culture that is substantially free of organisms other than *N. eutropha*. In some embodiments, “substantially free” denotes undetectable by a method used to detect other organisms, e.g., plating the culture and examining colony morphology, or PCR for a conserved gene such as 16S RNA. An axenic composition may comprise elements that are not organisms, e.g., it may comprise nutrients or excipients. Any embodiment, preparation, composition, or formulation of ammonia oxidizing bacteria discussed herein may comprise, consist essentially of, or consist of optionally axenic ammonia oxidizing bacteria.

[0236] Throughout this disclosure, formulation may refer to a composition or preparation.

[0237] As used herein, an “autotroph”, e.g., an autotrophic bacterium, is any organism capable of self-nourishment by using inorganic materials as a source of nutrients and using photosynthesis or chemosynthesis as a source of energy. Autotrophic bacteria may synthesize organic compounds from carbon dioxide and ATP derived from other sources, oxidation of ammonia to nitrite, oxidation of hydrogen sulfide, and oxidation of Fe^{2+} to Fe^{3+} . Autotrophic bacteria of the present disclosure are incapable of causing infection.

[0238] Administered “in combination,” as used herein, means that two (or more) different treatments are delivered to the subject during the course of the subject’s affliction with the disorder, e.g., the two or more treatments are delivered after the subject has been diagnosed with the disorder and before the disorder has been cured or eliminated. In some embodiments, the delivery of one treatment is still occurring when the delivery of the second begins, so that there is overlap. This is sometimes referred to herein as “simultaneous” or “concomitant” or “concurrent delivery”. In other embodiments, the delivery of one treatment ends before the delivery of the other treatment begins. This is sometimes referred to herein as “successive” or “sequential delivery.” In embodiments of either case, the treatment is more effective because of combined administration. For example, the second treatment is a more effective, e.g., an equivalent effect is seen with less of the second treatment, or the second treatment reduces symptoms to a greater extent, than would be seen if the second treatment were administered in the absence of the first treatment, or the analogous situation is seen with the first treatment. In some embodiments, delivery is such that the reduction in a symptom, or other parameter related to the disorder is greater than what would be observed with one treatment delivered in the absence of the other. The effect of the two treatments can be partially additive, wholly additive, or greater than additive (i.e., synergistic). The delivery can be such that an effect of the first treatment delivered is still detectable when the second is delivered.

[0239] Complete *N. europaea* medium refers to the *N. europaea* growth medium described in Ensign et al., “In vitro activation of ammonia monooxygenase from *Nitrosomonas europaea* by copper.” J Bacteriol. 1993 April; 175(7):1971-80.

[0240] To “culture” refers to a process of placing an amount of a desired bacterium under conditions that promote its growth, i.e., promoting cell division. The conditions can involve a specified culture medium, a set temperature

range, and/or an agitation rate. Bacteria can be cultured in a liquid culture or on plates, e.g., agar plates.

[0241] The term “isolated,” as used herein, refers to material that is removed from its original or native environment (e.g., the natural environment if it is naturally occurring). For example, a naturally-occurring polynucleotide or polypeptide present in a living animal is not isolated, but the same polynucleotide or polypeptide, separated by human intervention from some or all of the co-existing materials in the natural system, is isolated. Such polynucleotides could be part of a vector and/or such polynucleotides or polypeptides could be part of a composition, and still be isolated in that such vector or composition is not part of the environment in which it is found in nature.

[0242] The terms “nucleic acid,” “nucleic acid sequence,” “nucleotide sequence,” or “polynucleotide sequence,” and “polynucleotide” are used interchangeably. They refer to a polymeric form of nucleotides of any length, e.g., deoxyribonucleotides or ribonucleotides, or analogs thereof. The polynucleotide may be either single-stranded or double-stranded, and if single-stranded may be the coding strand or non-coding (antisense) strand. A polynucleotide may comprise modified nucleotides, such as methylated nucleotides and nucleotide analogs. The sequence of nucleotides may be interrupted by non-nucleotide components. A polynucleotide may be further modified after polymerization, such as by conjugation with a labeling component.

[0243] The nucleic acid may be a recombinant polynucleotide, or a polynucleotide of genomic, cDNA, semisynthetic, or synthetic origin which either does not occur in nature or is linked to another polynucleotide in a nonnatural arrangement.

[0244] As used herein, the term “optimized growth rate” refers to one or more of: a doubling time of less than about 4, 5, 6, 7, 8, 9, or 10 hours when cultured under batch conditions as described herein in Example 2; a doubling time of less than about 16, 18, 20, 22, 24, or 26 hours, when grown under chemostat conditions as described herein in Example 2; or growing from an OD₆₀₀ of about 0.15 to at least about 0.3, 0.4, 0.5, 0.6, 0.7, or 0.8 over about 1 or 2 days. In an embodiment, optimized growth rate is one having a doubling time that it is at least 10, 20, 30, 40, or 50% shorter than that of a naturally occurring *N. eutropha*.

[0245] As used herein, “optimized NH₄⁺ oxidation rate” refers to a rate of at least about 50, 75, 125, or 150 micromoles per minute of converting NH₃ or NH₄⁺ into NO₂⁻. For instance, the rate may be at least about 50, 75, 125, or 150 micromoles per minute of converting NH₄⁺ (e.g., at about 200 mM) to NO₂⁻. In an embodiment, an optimized NH₄⁺ oxidation rate is one in which NH₃ or NH₄⁺ is converted into NO₂⁻ at least 10, 20, 30, 40, or 50% more rapidly than is seen with a naturally occurring *N. eutropha*.

[0246] Percent (%) amino acid sequence identity, with respect to the amino acid sequences here (e.g., proteins expressed by *N. eutropha* D23) is defined as the percentage of amino acid residues in a candidate sequence that are identical with the amino acid residues in the reference sequence, which may be a naturally-occurring *N. eutropha* sequence or an *N. eutropha* D23 sequence, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity, and not considering any conservative substitutions as part of the sequence identity. Alignment for purposes of determining percent amino acid sequence identity can be achieved in various ways that

are within the means of those skilled in the art, for instance, using publicly available computer software such as BLAST, ALIGN or Megalign (DNASTAR) software. Those skilled in the art can determine appropriate parameters for measuring alignment, including any algorithms needed to achieve maximal alignment over the full length of the sequences being compared. For instance, the WU-BLAST-2 software may be used to determine amino acid sequence identity (Altschul et al, Methods in Enzymology 266, 460-480 (1996); <http://blast.wustl.edu/blast/README.html>). WU-BLAST-2 uses several search parameters, most of which are set to the default values. The adjustable parameters are set with the following values: overlap span=1, overlap fraction=0.125, word threshold (T)=1. HSP score (S) and HSP S2 parameters are dynamic values and are established by the program itself, depending upon the composition of the particular sequence, however, the minimum values may be adjusted as appropriate.

[0247] Amino acid substitutions can be the result of replacing one amino acid with another amino acid having similar structural and/or chemical properties, such as the replacement of a leucine with a serine, i.e., conservative amino acid replacements. Typical but not limiting conservative substitutions are the replacements, for one another, among the aliphatic amino acids Ala, Val, Leu and Ile; interchange of Ser and Thr containing hydroxy residues, interchange of the acidic residues Asp and Glu, interchange between the amide-containing residues Asn and Gln, interchange of the basic residues Lys and Arg, interchange of the aromatic residues Phe and Tyr, and interchange of the small-sized amino acids Ala, Ser, Thr, Met and Gly. Additional conservative substitutions include the replacement of an amino acid by another of similar spatial or steric configuration, for example the interchange of Asn for Asp, or Gln for Glu. Amino acid substitutions can also be the result of replacing one amino acid with another amino acid having dissimilar structural and/or chemical properties, i.e., non-conservative amino acid replacements. Insertions or deletions may optionally be in the range of 1 to 5 amino acids. The variation allowed may be determined by systematically making insertions, deletions or substitutions of amino acids in the sequence and testing the resulting variants for activity in the in vivo or in vitro assays for, e.g., metabolizing urea or ammonia.

[0248] Percent (%) sequence identity with respect to the nucleic acid sequences here (e.g., the *N. eutropha* D23 genome and portions thereof) is defined as the percentage of nucleotides in a candidate sequence that are identical with the nucleotides in the reference sequence, which may be a naturally-occurring *N. eutropha* sequence or an *N. eutropha* D23 sequence, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity. Alignment for purposes of determining percent nucleotide sequence identity can be achieved in various ways that are within the means of those skilled in the art, for instance, using publicly available computer software such as BLAST. Those skilled in the art can determine appropriate parameters for measuring alignment, including any algorithms needed to achieve maximal alignment over the full length of the sequences being compared.

[0249] The terms “polypeptide”, “peptide” and “protein” (if single chain) are used interchangeably herein to refer to amino acid polymers. The polymer may be linear or branched, it may comprise modified amino acids, and it may

be interrupted by non-amino acids. The terms also encompass an amino acid polymer that has been modified; for example, disulfide bond formation, glycosylation, lipidation, acetylation, phosphorylation, or any other manipulation, such as conjugation with a labeling component. The polypeptide can be isolated from natural sources, can be a produced by recombinant techniques from a eukaryotic or prokaryotic host, or can be a product of synthetic procedures.

[0250] As used herein, “optimized resistance to NH_4^{++} ” refers to an ability to grow in conditions of greater than 50, 75, 100, 125, 150, 175, 200, 225, 250, 275, or 300 mM NH_3 or NH_4^+ for at least about 24 or 48 hours. In an embodiment, an optimized resistance to NH_4^+ refers to the ability to grow at least 10, 20, 30, 40, or 50% more rapidly, or at least 10, 20, 30, 40, or 50% longer, in the presence of a selected concentration of NH_3 or NH_4^+ than can a naturally occurring *N. eutropha*.

[0251] As used herein with respect to a comparison between nucleic acid or protein sequences, “similar” means having homology. A similar gene or protein may comprise, e.g., substitutions (such as conservative or non-conservative substitutions), insertions (e.g., of at least 1, 2, 3, 4, 5, 10, 15, 20, 25, 30 amino acids, and for example up to 2, 3, 4, 5, 10, 15, 20, 25, 30, or 50 amino acids, or any positive combination thereof, or the number of nucleotides necessary to encode said amino acids), or deletions (e.g., of at least 1, 2, 3, 4, 5, 10, 15, 20, 25, 30 amino acids, and for example up to 2, 3, 4, 5, 10, 15, 20, 25, 30, or 50 amino acids, or any positive combination thereof, or the number of nucleotides necessary to encode said amino acids), or any combination thereof. Each of substitutions, insertions, and deletions may be positioned at the N-terminus, C-terminus, or a central region of the protein or gene. In embodiments, a conservative substitution is one that does not alter the charge and/or polarity and/or approximate size and/or geometry at the substituted position.

[0252] As used herein, “transgenic” means comprising one or more exogenous portions of DNA. The exogenous DNA is derived from another organism, e.g., another bacterium, a bacteriophage, an animal, or a plant.

[0253] As used herein, treatment of a disease or condition refers to reducing the severity or frequency of at least one symptom of that disease or condition, compared to a similar but untreated patient. Treatment can also refer to halting, slowing, or reversing the progression of a disease or condition, compared to a similar but untreated patient. Treatment may comprise addressing the root cause of the disease and/or one or more symptoms.

[0254] As used herein a therapeutically effective amount refers to a dose sufficient to prevent advancement, or to cause regression of a disease or condition, or which is capable of relieving a symptom of a disease or condition, or which is capable of achieving a desired result. A therapeutically effective dose can be measured, for example, as a number of bacteria or number of viable bacteria (e.g., in CFUs) or a mass of bacteria (e.g., in milligrams, grams, or kilograms), or a volume of bacteria (e.g., in mm^3).

[0255] As used herein, the term “viability” refers to the autotrophic bacteria’s, e.g., ammonia oxidizing bacteria’s, ability to oxidize ammonia, ammonium, or urea to nitrite at a pre-determined rate. In some embodiments, the rate refers to the conversion of ammonium ions (NH_4^+) (e.g., at about 200 mM) to nitrite (NO_2^-) at a rate of at least 50, 75, 125,

or 150 micromoles NO_2^- per minute, e.g., about 100-150, 75-175, 75-125, 100-125, 125-150, or 125-175 micromoles/minute, e.g., about 125 micromoles NO_2^- per minute.

[0256] “Growth media” or “AOB media,” as referred to herein comprises the following components of Table 3 or Table 4 herein.

[0257] In some embodiments, the states most relevant to the present disclosure are the state of growth, e.g., maximal growth, characterized by a pH of at least about 7.6, ammonia, trace minerals, oxygen and carbon dioxide. Another state may be characterized by a pH of about 7.4 or less and characterized by an absence of carbon dioxide. Under low carbon dioxide conditions, ammonia oxidizing bacteria, e.g., *Nitrosomonas*, continues to oxidize ammonia into nitrite and generates ATP, but lacking carbon dioxide, e.g., lacking sufficient carbon dioxide, to fix and generate protein, it instead generates polyphosphate, which it uses as an energy storage medium. This may allow the ammonia oxidizing bacteria to remain in a “storage state” for a period of time, e.g., a pre-determined period of time, for example, at least 1, 2, 3, 4, 5, 6, 7, days, 1, 2, 3, 4 weeks, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12 months, 1, 2, 3, 4, or 5 years. In some embodiments, the ammonia oxidizing bacteria may remain in a storage state for at least about 6 months to about 1 year.

[0258] As used herein, “growth state” refers to autotrophic bacteria, e.g., ammonia oxidizing bacteria, in a state or in an environment, e.g., a media, e.g., a culture media, e.g., a growth media, that may have a pH of at least about 7.6. Levels of at least one of ammonia, ammonium ions, and urea may be between about 1 micromolar and 1000 millimolar. Levels of trace materials are between about 0.01 micromolar iron and 200 micromolar iron. Levels of oxygen are between about 5% and 100% oxygen saturation (e.g., of media). Levels of carbon dioxide are between about 20 ppm and 10% saturation (e.g., of media). In certain aspects, levels of at least one of ammonia, ammonium ions, and urea may be between about 10 micromolar and 100 millimolar. Levels of trace materials are between about 0.1 micromolar iron and 20 micromolar iron. Levels of oxygen are between about 5% and 100% oxygen saturation. Levels of carbon dioxide are between about 200 ppm and 5% saturation (e.g., of media).

[0259] As used herein, “polyphosphate loading state” refers to autotrophic bacteria, e.g., ammonia oxidizing bacteria, in a state or in an environment, e.g., a media, e.g., a culture media, e.g., a growth media, that may have a pH of about 7.4, or less. Levels of at least one of ammonia, ammonium ions, and urea are between about 1 micromolar and 2000 millimolar. Levels of trace materials are between 0.01 micromolar iron and 200 micromolar iron. Levels of oxygen are between about 0% and 100% O_2 saturation (e.g., of media). Levels of carbon dioxide are between/less than about zero and 400 ppm, and phosphate levels greater than about 1 micromolar. In certain aspects, levels of at least one of ammonia, ammonium ions, and urea are between about 10 micromolar and 200 millimolar. Levels of trace materials are between 0.1 micromolar iron and 20 micromolar iron. Levels of oxygen are between about 5% and 100% O_2 saturation. Levels of carbon dioxide are between/less than about zero and 200 ppm, and phosphate levels greater than about 10 micromolar.

[0260] The polyphosphate loading state may be induced for a period of time, e.g., a pre-determined period of time. The pre-determined period of time may be the time period that allows sufficient polyphosphate accumulation in the ammo-

nia oxidizing bacteria. This pre-determined period of time is the period of time suitable to provide for sufficient polyphosphate loading to allow for the ammonia oxidizing bacteria to be stored for an extended period of time. The pre-determined period of time may be at least partially based on a period of time of about 0.2-10 times, 0.3-5 times, 0.5-3 times, 0.5-1.5 times, or 0.5 to 1 times the doubling time for the ammonia oxidizing bacteria. The pre-determined period of time may be at least partially based on a period of time of about one doubling time for the ammonia oxidizing bacteria. In some embodiments, the pre-determined period of time is between about 8 hours and 12 hours. In some embodiments, the pre-determined period of time is about 10 hours. In some embodiments, the pre-determined period of time is about 24 hours.

[0261] A purpose of the polyphosphate loading state may be to provide AOB with sufficient ammonia, ammonium ions, and/or urea, and O_2 such that ATP can be produced, but to deny them CO_2 and carbonate such that they are unable to use that ATP to fix CO_2 and instead use that ATP to generate polyphosphate which may be stored by the bacteria.

[0262] As used herein, the term "storage state" refers to autotrophic bacteria, e.g., ammonia oxidizing bacteria, in a state or in an environment, e.g., a media, e.g., a culture media, e.g., a growth media, having a pH of about 7.4 or less (in some embodiments, the pH may be 7.6 or less). Levels of at least one of ammonia, ammonium ions, and urea are between about 10^{-1} and 1000 micromolar. Levels of trace materials are between about 0.1 and 100 micromolar. Levels of oxygen are between about 0 and 100% saturation (e.g., of media). Levels of carbon dioxide are between about 0 and 800 ppm. In certain aspects, levels of at least one of ammonia, ammonium ions, and urea are between about 10^{-1} and 100 micromolar. Levels of trace materials are between about 1 and 10 micromolar. Levels of oxygen are between about 0 and 100% saturation (e.g., of media). Levels of carbon dioxide are between about 0 and 400 ppm.

[0263] AOB are produced according to some embodiments of the present disclosure by generating AOB biomass during a growth state, then exposing the AOB to a polyphosphate loading state and then removing the media and resuspending the AOB in a buffer, e.g., a storage buffer (i.e., the storage state).

[0264] The ammonia oxidizing bacteria may remain in a "storage state" for a period of time, e.g., a pre-determined period of time, for example, at least 1, 2, 3, 4, 5, 6, 7, days, 1, 2, 3, 4 weeks, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12 months, 1, 2, 3, 4, or 5 years. In some embodiments, the ammonia oxidizing bacteria may remain in a storage state for at least about 6 months to about 1 year. Upon revival, the viability of the ammonia oxidizing bacteria is at least about 50%, 60%, 70%, 80%, 90%, or 100% of the viability as of the ammonia oxidizing bacteria prior to storage e.g., in a growth state). In some embodiments, the preparation of ammonia oxidizing bacteria may be prepared, such that no more than 10%, 20%, 30%, 40%, 50%, 60%, or 70% of the ability to oxidize NH_4^+ is lost upon storage at selected conditions.

[0265] The time that it takes to revive the ammonia oxidizing bacteria from a storage state (or a polyphosphate loading state) may be a pre-determined period of time. For example, the pre-determined period of time may be less than about 75 hours, or less than about 72 hours. The pre-determined period of time may at least partially based on a

period time of about 0.2-10 times, 0.3-5 times, 0.5-3 times, 0.5-1.5 times, or 0.5 to 1 times the doubling time for the ammonia oxidizing bacteria. The pre-determined period of time may be at least partially based on a period of time of about one doubling time for the ammonia oxidizing bacteria. The pre-determined period of time may be between about 8 hours and 12 hours. The pre-determined period of time may be about 10 hours. The pre-determined time may be less than about 75 hours, 72 hours, 70 hours, 68 hours, 65 hours, 60 hours, 55 hours, 50 hours, 45 hours, 40 hours, 35 hours, 30 hours, 25 hours, 20 hours, 15 hours, 10 hours, 5 hours, 4 hours, 3, hours, 2 hours, or 1 hour. The pre-determined period of time may be between about 5 minutes and 5 hours. The pre-determined period of time may be about 5-10 minutes, 10-15 minutes, 15-20 minutes, 20-25 minutes, 25-30 minutes, 30-45 minutes, 45-60 minutes, 60 minutes-1.5 hours, 1.5 hours-2 hours, 2 hours-2.5 hours, 2.5 hours-3 hours, 3 hours-3.5 hours, 3.5 hours-4 hours, 4 hours-4.5 hours, 4.5 hours-5 hours. In some embodiments, the pre-determined period of time may be about 2 hours. The pre-determined period of time, e.g., may be the time it may take to achieve revival of the ammonia oxidizing bacteria, e.g., achieve viability of the ammonia oxidizing bacteria as compared to the viability of the bacteria prior to storage (e.g., in a growth state), e.g., at least about 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 97%, 98%, 99%, or 100% viability.

2. AMMONIA OXIDIZING BACTERIA (AOBS), *N. EUTROPHA* STRAIN D23 AND SIMILAR BACTERIA

[0266] Autotrophic ammonia oxidizing bacteria, which may be referred to herein as AOBs or AOB, are obligate autotrophic bacteria as noted by Alan B. Hooper and A. Krummel at al. Alan B. Hooper, Biochemical Basis of Obligate Autotrophy in *Nitrosomonas europaea*, Journal of Bacteriology, February 1969, p. 776-779. Antje Krummel et al., Effect of Organic Matter on Growth and Cell Yield of Ammonia-Oxidizing Bacteria, Arch Microbiol (1982) 133: 50-54. These bacteria derive all metabolic energy only from the oxidation of ammonia to nitrite with nitric oxide (NO) as an intermediate product in their respiration chain and derive virtually all carbon by fixing carbon dioxide. They are incapable of utilizing carbon sources other than a few simple molecules.

[0267] Ammonia oxidizing bacteria (AOB) are widely found in the environment, and in the presence of ammonia, oxygen and trace metals will fix carbon dioxide and proliferate. AOB may be slow growing and toxic levels of ammonia may kill fish and other organisms before AOB can proliferate and reduce ammonia to non-toxic levels. Slow growth of AOB also may delay the health benefits of the NO and nitrite the AOB produce when applied to the skin.

[0268] Supplementing the aquarium, skin, or process with sufficient viable AOB grown and stored for that purpose is desired. AOB do not form spores, so storage in the dry state with high viability is difficult, and storage in the wet state leaves them metabolically active. Decay of nitrifying capacity during storage of AOB for wastewater treatment has been studied, as for example (Munz G, Lubello C, Oleszkiewicz J A. Modeling the decay of ammonium oxidizing bacteria. Water Res. 2011 January; 45(2): 557-64. OI: 10.1016/j.watres.2010.09.022.)

[0269] Growth, prolonged storage, and restoration of activity of *Nitrosomonas* is discussed by Cassidy et al. (U.S. Pat. No. 5,314,542) where they disclose growing *Nitrosomonas*, removing toxic waste products, storing in sterile water of appropriate salinity for periods of time up to one year, and then reviving by adding buffer (CaCO_3) and 200 ppm, of ammonium, which reviving takes 72 hours.

[0270] As obligate autotrophs, AOB synthesize protein via the fixing of CO_2 using the energy and reducing equivalents generated by the oxidation of ammonia to nitrite. Growth requires ammonia, oxygen, minerals and carbon dioxide.

[0271] *Nitrosomonas* may exist in several metabolic states, according to "Polyposphate and Orthophosphate Content of *Nitrosomonas europaea* as a Function of Growth" by K. R. Terry and A. B. Hooper, Journal of Bacteriology, July 1970, p. 199-206, Vol. 103, No. 1.

[0272] In certain embodiments of the disclosure, the ammonia oxidizing bacteria may be axenic. The preparation (formulation or composition) of ammonia oxidizing bacteria may comprise, consist essentially of, or consist of axenic ammonia oxidizing bacteria. The ammonia oxidizing bacteria may be from a genus selected from the group consisting of *Nitrosomonas*, *Nitrosococcus*, *Nitrosospria*, *Nitrosocystis*, *Nitrosolobus*, *Nitrosovibrio*, and combinations thereof.

[0273] This disclosure provides, inter alia, *N. eutropha* strain D23, a unique, e.g., optimized strain of ammonia oxidizing bacteria that can increase production of nitric oxide and nitric oxide precursors on the surface of a subject, e.g., a human subject. This disclosure also provides methods of using the bacteria and articles comprising the bacteria.

[0274] In embodiments, the *N. eutropha* is non-naturally occurring. For instance, it may have accumulated desirable mutations during a period of selection. In other embodiments, desirable mutations may be introduced by an experimenter. In some embodiments, the *N. eutropha* may be a purified preparation, and may be an optimized *N. eutropha*.

[0275] In preferred embodiments, the *N. eutropha* strain is autotrophic and so incapable of causing infection. A preferred strain utilizes urea as well as ammonia, so that hydrolysis of the urea in sweat would not be necessary prior to absorption and utilization by the bacteria. Also, in order to grow at low pH, the bacteria may either absorb NH_4^+ ions or urea. The selected strain should also be capable of living on the external skin of a subject, e.g., a human, and be tolerant of conditions there.

[0276] Although this disclosure refers to *N. eutropha* strain D23 in detail, the preparations, methods, compositions, treatments, wearable articles, and articles of clothing may be used with one or more of: one or more other strains of *N. eutropha*, one or more other species of *Nitrosomonas*, and one or more other ammonia oxidizing bacteria. Autotrophic AOBs are obligate autotrophic bacteria as noted by Alan B. Hooper and A. Krummel at al. Alan B. Hooper, Biochemical Basis of Obligate Autotrophy in *Nitrosomonas europaea*, Journal of Bacteriology, February 1969, p. 776-779. Antje Krummel et al., Effect of Organic Matter on Growth and Cell Yield of Ammonia-Oxidizing Bacteria, Arch Microbiol (1982) 133: 50-54. These bacteria derive all metabolic energy only from the oxidation of ammonia to nitrite with nitric oxide (NO) as an intermediate product in their respiration chain and derive virtually all carbon by fixing carbon dioxide. They are incapable of utilizing carbon sources other than a few simple molecules.

[0277] In certain embodiments, the *N. eutropha* is the strain deposited with the American Tissue Culture Collection (ATCC) on Apr. 8, 2014, designated AOB D23-100 (25 vials) under accession number PTA-121157.

[0278] In certain embodiments, the *N. eutropha* comprises a chromosome having a sequence at least 70%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 99.5% identical to SEQ ID NO: 1 (the strain D23 whole-genome sequence).

[0279] In certain embodiments, a bacterium with the above-mentioned sequence characteristics has one or more of (1) an optimized growth rate as measured by doubling time, (2) an optimized growth rate as measured by OD600, (3) an optimized NH_4^+ oxidation rate, (4) an optimized resistance to NH_4^+ , and (4) an optimized resistance to NO_2^- . Particular sub-combinations of these properties are specified in the following paragraph.

[0280] In some embodiments, the *N. eutropha* described herein has one or more of: (1) an optimized growth rate as measured by doubling time, (2) an optimized growth rate as measured by OD600, (3) an optimized NH_4^+ oxidation rate, (4) an optimized resistance to NH_4^+ , and (4) an optimized resistance to NO_2^- . For instance, the bacterium may have properties (1) and (2); (2) and (3); (3) and (4); or (4) and (5) from the list at the beginning of this paragraph. As another example, the bacterium may have properties (1), (2), and (3); (1), (2), and (4); (1), (2), and (5); (1), (3), and (4); (1), (3), and (5); (1), (4), and (5); (2), (3), and (4); (2), (3), and (5), or (3), (4), and (5) from the list at the beginning of this paragraph. As a further example, the bacterium may have properties (1), (2), (3), and (4); (1), (2), (3), and (5); (1), (2), (4), and (5); (1), (3), (4), and (5); or (2), (3), (4), and (5) from the list at the beginning of this paragraph. In some embodiments, the bacterium has properties (1), (2), (3), (4), and (5) from the list at the beginning of this paragraph.

[0281] This disclosure also provides an axenic composition of *N. eutropha* having one or more of: (1) an optimized growth rate as measured by doubling time, (2) an optimized growth rate as measured by OD600, (3) an optimized NH_4^+ oxidation rate, (4) an optimized resistance to NH_4^+ , and (4) an optimized resistance to NO_2^- . For instance, the axenic *N. eutropha* composition may have properties (1) and (2); (2) and (3); (3) and (4); or (4) and (5) from the list at the beginning of this paragraph. As another example, the axenic *N. eutropha* composition may have properties (1), (2), and (3); (1), (2), and (4); (1), (2), and (5); (1), (3), and (4); (1), (3), and (5); (1), (4), and (5); (2), (3), and (4); (2), (3), and (5), or (3), (4), and (5) from the list at the beginning of this paragraph. As a further example, the axenic *N. eutropha* composition may have properties (1), (2), (3), and (4); (1), (2), (3), and (5); (1), (2), (4), and (5); (1), (3), (4), and (5); or (2), (3), (4), and (5) from the list at the beginning of this paragraph. In some embodiments, the axenic *N. eutropha* composition has properties (1), (2), (3), (4), and (5) from the list at the beginning of this paragraph.

[0282] *N. eutropha* strain D23, as deposited in the form of 25 vials on Apr. 8, 2014, in the ATCC patent depository, designated AOB D23-100, under accession number PTA-121157, comprises a circular genome having SEQ ID NO: 1 or its complement. Accordingly, in some embodiments, an *N. eutropha* strain described herein comprises a nucleic acid sequence, e.g., a genome, that is similar to SEQ ID NO: 1 or its complement.

[0283] For instance, the *N. eutropha* may comprise a nucleic acid sequence having a 1,000 base pair portion

having at least about 80%, 85%, 90%, 95%, 97%, 98%, 99%, or 99.5% identity to a 1,000 base pair portion of SEQ ID NO: 1 or its complement. The 1,000 base pair portion may span, e.g., nucleotides $(n*1,000)+1$ to $(n+1)*1,000$, where $n=0, 1, 2, 3 \dots 2538$, e.g., nucleotides 1-1,000, 1,001-2,000, and so on through the end of SEQ ID NO: 1.

[0284] In embodiments, the *N. eutropha* comprises a nucleic acid sequence having a 2,000 base pair portion having at least about 80%, 85%, 90%, 95%, 97%, 98%, 99%, or 99.5% identity to a 2,000 base pair portion of SEQ ID NO: 1 or its complement. The 2,000 base pair portion may span, e.g., nucleotides $(n*2,000)+1$ to $(n+1)*2,000$, where $n=0, 1, 2, 3 \dots 1269$, e.g., nucleotides 1-2,000, 2,001-4,000, and so on through the end of SEQ ID NO: 1.

[0285] In embodiments, the *N. eutropha* comprises a nucleic acid sequence having a 5,000 base pair portion having at least about 80%, 85%, 90%, 95%, 97%, 98%, 99%, or 99.5% identity to a 5,000 base pair portion of SEQ ID NO: 1 or its complement. The 5,000 base pair portion may span, e.g., nucleotides $(n*5,000)+1$ to $(n+1)*5,000$, where $n=0, 1, 2, 3 \dots 508$, e.g., nucleotides 1-5,000, 5,001-10,000, and so on through the end of SEQ ID NO: 1.

[0286] In embodiments, the *N. eutropha* comprises a nucleic acid sequence having a 10,000 base pair portion having at least about 80%, 85%, 90%, 95%, 97%, 98%, 99%, or 99.5% identity to a 10,000 base pair portion of SEQ ID NO: 1 or its complement. The 10,000 base pair portion may span, e.g., nucleotides $(n*10,000)+1$ to $(n+1)*10,000$, where $n=0, 1, 2, 3 \dots 254$, e.g., nucleotides 1-10,000, 10,001-20,000, and so on through the end of SEQ ID NO: 1.

[0287] In embodiments, the *N. eutropha* comprises a nucleic acid sequence having a 20,000 base pair portion having at least about 80%, 85%, 90%, 95%, 97%, 98%, 99%, or 99.5% identity to a 20,000 base pair portion of SEQ ID NO: 1 or its complement. The 20,000 base pair portion may span, e.g., nucleotides $(n*20,000)+1$ to $(n+1)*20,000$, where $n=0, 1, 2, 3 \dots 127$, e.g., nucleotides 1-20,000, 20,001-40,000, and so on through the end of SEQ ID NO: 1.

[0288] In embodiments, the *N. eutropha* comprises a nucleic acid sequence having a 50,000 base pair portion having at least about 80%, 85%, 90%, 95%, 97%, 98%, 99%, or 99.5% identity to a 50,000 base pair portion of SEQ ID NO: 1 or its complement. The 50,000 base pair portion may span, e.g., nucleotides $(n*50,000)+1$ to $(n+1)*50,000$, where $n=0, 1, 2, 3 \dots 51$, e.g., nucleotides 1-50,000, 50,001-100,000, and so on through the end of SEQ ID NO: 1.

[0289] In embodiments, the *N. eutropha* comprises a nucleic acid sequence having a 100,000 base pair portion having at least about 80%, 85%, 90%, 95%, 97%, 98%, 99%, or 99.5% identity to a 100,000 base pair portion of SEQ ID NO: 1 or its complement. The 100,000 base pair portion may span, e.g., nucleotides $(n*100,000)+1$ to $(n+1)*100,000$, where $n=0, 1, 2, 3 \dots 26$, e.g., nucleotides 1-100,000, 100,001-20,000, and so on through the end of SEQ ID NO: 1.

[0290] In some aspects, the present disclosure provides a composition of *N. eutropha* comprising a chromosome at least about 70%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 99.5% identical to SEQ ID NO: 1. In some aspects, the present disclosure provides an axenic composition of *N.*

eutropha comprising a chromosome at least about 70%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 99.5% identical to SEQ ID NO: 1.

[0291] In certain embodiments, the *N. eutropha* strain comprises a nucleic acid sequence, e.g., a genome, that hybridizes to SEQ ID NO: 1, or to the genome of the D23 strain deposited in the form of 25 vials with the ATCC patent depository on Apr. 8, 2014, designated AOB D23-100, under accession number PTA-121157, or their complements, under low stringency, medium stringency, high stringency, or very high stringency, or other hybridization condition described herein. As used herein, the term “hybridizes under low stringency, medium stringency, high stringency, or very high stringency conditions” describes conditions for hybridization and washing. Guidance for performing hybridization reactions can be found in *Current Protocols in Molecular Biology*, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6, which is incorporated by reference. Aqueous and nonaqueous methods are described in that reference and either can be used. Specific hybridization conditions referred to herein are as follows: 1) low stringency hybridization conditions in 6× sodium chloride/sodium citrate (SSC) at about 45° C., followed by two washes in 0.2×SSC, 0.1% SDS at least at 50° C. (the temperature of the washes can be increased to 55° C. for low stringency conditions); 2) medium stringency hybridization conditions in 6×SSC at about 45° C., followed by one or more washes in 0.2×SSC, 0.1% SDS at 60° C.; 3) high stringency hybridization conditions in 6×SSC at about 45° C., followed by one or more washes in 0.2×SSC, 0.1% SDS at 65° C.; 4) very high stringency hybridization conditions are 0.5M sodium phosphate, 7% SDS at 65° C., followed by one or more washes at 0.2×SSC, 1% SDS at 65° C. Very high stringency conditions (4) are suitable conditions and the ones that should be used unless otherwise specified.

[0292] The genome of strain D23 (SEQ ID NO: 1) was compared with the genome of *N. eutropha* C91. An annotation of the D23 genome is shown in Supplementary Table 1, which lists the positions of 2,777 genes in SEQ ID NO: 1 as identified by sequence analysis. In certain embodiments, the *N. eutropha* described herein comprises one or more genes or proteins listed in Supplementary Table 1, or a gene or protein similar to one of said genes or proteins.

[0293] Accordingly, in some embodiments, the *N. eutropha* comprises a gene of Supplementary Table 1, or a protein encoded by said gene. In certain embodiments, the *N. eutropha* comprises a gene that is similar (e.g., at least 70%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, or 99.5% identical) to a gene of Supplementary Table 1, or a protein encoded by said gene. In embodiments, the *N. eutropha* comprises genes or proteins that are identical or similar to at least 2, 3, 4, 5, 10, 20, 30, 40, 50, 100, 150, 200, 250, 300, 350, 400, 450, 500, 1000, 1500, 2000, 2500, or all the genes of Supplementary Table 1, or a protein encoded by said genes.

[0294] In some embodiments, the *N. eutropha* described herein (e.g., strain D23) comprises one or more genes or proteins that are absent from strain C91, or a gene or protein similar to one of said genes or proteins. Examples of these genes are set out in FIG. 6-8 and are described in more detail in Example 4 herein.

[0295] Accordingly, with respect to FIG. 6, in some embodiments, the *N. eutropha* comprises genes that are identical or similar (e.g., at least 70%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, or 99.5% identical) to 1, 2, 3, 4, 5,

10, 15, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, or all of the genes in FIG. 6. In some embodiments, the *N. eutropha* comprises proteins that are identical or similar (e.g., at least 70%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, or 99.5% identical) to 1, 2, 3, 4, 5, 10, 15, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, or all of the proteins encoded by the genes listed in FIG. 6.

[0296] With respect to FIG. 7, in some embodiments, the *N. eutropha* comprises genes that are identical or similar (e.g., at least 70%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, or 99.5% identical) to 1, 2, 3, 4, 5, 10, 15, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, or all of the genes in FIG. 7. In some embodiments, the *N. eutropha* comprises proteins that are identical or similar (e.g., at least 70%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, or 99.5% identical) to 1, 2, 3, 4, 5, 10, 15, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, or all of the proteins encoded by the genes listed in FIG. 7.

[0297] With respect to FIG. 8, in some embodiments, the *N. eutropha* comprises genes that are identical or similar (e.g., at least 70%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, or 99.5% identical) to 1, 2, 3, 4, 5, 10, 15, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 200, or all of the genes in FIG. 8. In some embodiments, the *N. eutropha* comprises proteins that are identical or similar (e.g., at least 70%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, or 99.5% identical) to 1, 2, 3, 4, 5, 10, 15, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 200, or all of the proteins encoded by the genes listed in FIG. 8.

[0298] With respect to FIGS. 6-8 collectively, in some embodiments, the *N. eutropha* comprises genes that are identical or similar (e.g., at least 70%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, or 99.5% identical) to 1, 2, 3, 4, 5, 10, 20, 40, 60, 80, 100, 150, 200, 250, 300, 350, 400, 450, 500, or all of the genes in FIGS. 6-8. In some embodiments, the *N. eutropha* comprises proteins that are identical or similar (e.g., at least 70%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, or 99.5% identical) to 1, 2, 3, 4, 5, 10, 20, 40, 60, 80, 100, 150, 200, 250, 300, 350, 400, 450, 500, or all of the proteins encoded by genes listed in FIGS. 6-8.

[0299] In some embodiments, the *N. eutropha* described herein (e.g., strain D23) lacks one or more genes or proteins that are unique to strain C91, or a gene or protein similar to one of said genes or proteins. Examples of these genes are set out in FIG. 9 and are described in more detail in Example 4 herein. Accordingly, in some embodiments, the *N. eutropha* described herein lacks at least 1, 2, 3, 4, 5, 10, 20, 50, 100, 150, 200, 250, or all of the genes of FIG. 9. In some embodiments, the *N. eutropha* described herein lacks up to 2, 3, 4, 5, 10, 20, 50, 100, 150, 200, 250, or all of the genes of FIG. 9. In embodiments, the *N. eutropha* described herein lacks about 1-5, 5-10, 10-20, 20-50, 50-100, 100-150, 150-200, 200-250, or 250-all of the genes of FIG. 9.

[0300] Sequencing of the D23 genome revealed several genes of potential interest, including genes involved in ammonia metabolism (e.g., ammonia monooxygenase, hydroxylamine oxidoreductase, cytochrome c554, and cytochrome c_M552). All of these genes are present in multiple copies, and in general the copies are not identical to each other. One set of genes of interest is the ammonia monooxygenase synthesis operon amoCAB, which is present in two copies, along with a third copy of amoC. The operons have homologs in C91, i.e., Neut_2078/7/6 and

Neut_2319/8/7. Another set of genes of interest is hydroxylamine oxidoreductase (hao), which is present in three copies. The hao homologs in C91 are designated Neut_1672, 1793, and 2335. A third set of genes of interest is the cytochrome c554 gene encoded by cycA, which is present in three copies. The corresponding C91 genes are designated Neut_1670, 1791, and 2333. A fourth set of genes of interest is the cytochrome c_M552 genes encoded by cycB, which are present in two copies. The homologous C91 genes are designated Neut_1790 and 2332. Each group of genes is summarized in Table 1 and is discussed in more detail below.

TABLE 1

Sequences of ammonia metabolism genes in <i>N. eutropha</i> strain D23.			
SEQ ID in strain D23	SEQ ID in strain C91	Type	Gene name
1. ammonia monooxygenase			
4	34	Protein	amoC1
5	35	DNA	amoC1
6	36	Protein	amoA1
7	37	DNA	amoA1
8	38	Protein	amoB1
9	39	DNA	amoB1
10	40	Protein	amoC2
11	41	DNA	amoC2
12	42	Protein	amoA2
13	43	DNA	amoA2
14	44	Protein	amoB2
15	45	DNA	amoB2
16	46	Protein	amoC3
17	47	DNA	amoC3
2. hydroxylamine oxidoreductase			
18	48	Protein	hao1
19	49	DNA	hao1
20	50	Protein	hao2
21	51	DNA	hao2
22	52	Protein	hao3
23	53	DNA	hao3
3. cytochrome c554			
24	54	Protein	c554 cycA1
25	55	DNA	c554 cycA1
26	56	Protein	c554 cycA2
27	57	DNA	c554 cycA2
28	58	Protein	c554 cycA3
29	59	DNA	c554 cycA3
4. cytochrome c _M 552			
30	60	Protein	c _M 552 cycB1
31	61	DNA	c _M 552 cycB1
32	62	Protein	c _M 552 cycB2
33	63	DNA	c _M 552 cycB2

[0301] In some aspects, the *N. eutropha* described herein comprises genes identical to or similar to the genes and proteins of Table 1.

[0302] More particularly, in certain aspects, this disclosure provides a composition of *N. eutropha*, e.g., a purified preparation of *N. eutropha* comprising a nucleic acid sequence at least about 70%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 99.5% identical to an ammonia monooxygenase sequence of Table 1. In certain aspects, this disclosure provides a composition of *N. eutropha* comprising a nucleic acid sequence at least about 70%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 99.5% identical to a hydroxylamine oxidoreductase sequence of Table 1. In certain aspects, this disclosure provides a composition of *N.*

eutropha comprising a nucleic acid sequence at least about 70%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 99.5% identical to a cytochrome c554 sequence of Table 1. In certain aspects, this disclosure provides a composition of *N. eutropha* comprising nucleic acid sequences at least about 70%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 99.5% identical to a cytochrome c_M552 sequence of Table 1.

[0303] In certain aspects, this disclosure provides a composition of *N. eutropha* comprising an amino acid sequence at least about 70%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 99.2%, 99.3%, 99.4%, 99.5%, or 99.6% identical to an ammonia monooxygenase sequence of Table 1. In certain aspects, this disclosure provides a composition of *N. eutropha* comprising an amino acid sequence at least about 70%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 99.4%, 99.5%, 99.6%, or 99.7% identical to hydroxylamine oxidoreductase sequence of Table 1. In certain aspects, this disclosure provides a composition of *N. eutropha* comprising an amino acid sequence at least about 70%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 99.1%, 99.2%, 99.3%, 99.5%, 99.6%, or 99.7% identical to a cytochrome c554 sequence of Table 1. In certain aspects, this disclosure provides a composition of *N. eutropha* comprising amino acid sequences at least about 70%, 80%, 85%, 90%, 95%, 96%, 97%, 97.1%, 97.2%, 97.5%, 98%, 98.5%, 98.6%, 98.7%, 98.8%, 99%, or 99.5% identical to a cytochrome c_M552 sequence of Table 1.

[0304] In some embodiments, the *N. eutropha* are present in an axenic composition, and e.g., in the form of a purified preparation of optimized *N. eutropha*.

[0305] More particularly, in certain aspects, this disclosure provides an axenic composition of *N. eutropha* comprising a nucleic acid sequence at least about 70%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 98.5%, 98.8%, 98.9%, 99%, 99.2%, 99.3%, 99.4%, 99.5%, or 99.6% identical to an ammonia monooxygenase sequence of Table 1. In certain aspects, this disclosure provides an axenic composition of *N. eutropha* comprising a nucleic acid sequence at least about 70%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 99.5% identical to a cytochrome c554 sequence of Table 1. In certain aspects, this disclosure provides an axenic composition of *N. eutropha* comprising nucleic acid sequences at least about 70%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 99.5% identical to a cytochrome c_M552 sequence of Table 1.

[0306] In certain aspects, this disclosure provides an axenic composition of *N. eutropha* comprising an amino acid sequence at least about 70%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 98.5%, 98.8%, 98.9%, 99%, 99.2%, 99.3%, 99.4%, 99.5%, or 99.6% identical to an ammonia monooxygenase sequence of Table 1. In certain aspects, this disclosure provides an axenic composition of *N. eutropha* comprising an amino acid sequence at least about 70%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 99.4%, 99.5%, 99.6%, or 99.7% identical to hydroxylamine oxidoreductase sequence of Table 1. In certain aspects, this disclosure provides an axenic composition of *N. eutropha* comprising an amino acid sequence at least about 70%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 99.1%, 99.2%, 99.3%, 99.5%, 99.6%, or 99.7% identical to a

cytochrome c554 sequence of Table 1. In certain aspects, this disclosure provides an axenic composition of *N. eutropha* comprising amino acid sequences at least about 70%, 80%, 85%, 90%, 95%, 96%, 97%, 97.1%, 97.2%, 97.5%, 98%, 98.5%, 98.6%, 98.7%, 98.8%, 99%, or 99.5% identical to a cytochrome c_M552 sequence of Table 1.

[0307] In some embodiments, the *N. eutropha* comprises a gene or protein comprising a sequence at least about 70%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 99.5% identical to a strain D23 sequence of Table 1, e.g., any of SEQ IDs 4-33. Substitutions may be conservative or non-conservative; also, insertions and deletions are contemplated. In some embodiments, the *N. eutropha* comprises a gene or protein comprising a sequence of Table 1, e.g., any of SEQ IDs 4-33. In some embodiments, the protein has an N-terminal and/or C-terminal extension or deletion of up to about 1, 2, 3, 4, 5, 6, 8, 10, 15, 20, 25, 50, or 100 amino acids.

[0308] Alignment of the nucleic acid sequences of Table 1 shows the percent identity between homologs in C91 and D23. The following paragraphs discuss this percent identity and describe various genes having homology to the D23 genes of Table 1.

[0309] More specifically, the amoA1 genes are about 98.8% identical (i.e., at 821/831 positions). Accordingly, in some embodiments, the *N. eutropha* described herein comprise D23 nucleotides at at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or all of the positions that differ in this gene between strains C91 and D23. In embodiments, the *N. eutropha* described herein comprise D23 nucleotides at at most 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or all of the positions that differ in this gene between strains C91 and D23. In embodiments, the *N. eutropha* described herein comprise a gene at least about 98.8%, 98.9%, 99.0%, 99.2%, 99.4%, 99.6%, 99.8%, or 100% identical to the D23 amoA1 gene.

[0310] The amoA2 genes are about 98.8% identical (i.e., at 821/831 positions). Accordingly, in some embodiments, the *N. eutropha* described herein comprise D23 nucleotides at at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or all of the positions that differ in this gene between strains C91 and D23. In embodiments, the *N. eutropha* described herein comprise D23 nucleotides at at most 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or all of the positions that differ in this gene between strains C91 and D23. In embodiments, the *N. eutropha* described herein comprise a gene at least about 98.8%, 98.9%, 99.0%, 99.2%, 99.4%, 99.6%, 99.8%, or 100% identical to the D23 amoA2 gene.

[0311] The amoB1 genes are about 99.1% identical (i.e., at 1255/1266 positions). Accordingly, in some embodiments, the *N. eutropha* described herein comprise D23 nucleotides at at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, or all of the positions that differ in this gene between strains C91 and D23. In embodiments, the *N. eutropha* described herein comprise D23 nucleotides at at most 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, or all of the positions that differ in this gene between strains C91 and D23. In embodiments, the *N. eutropha* described herein comprise a gene at least about 99.1%, 99.2%, 99.4%, 99.6%, 99.8%, or 100% identical to the D23 amoB1 gene.

[0312] The amoB2 genes are about 99.1% identical (i.e., at 1254/1266 positions). Accordingly, in some embodiments, the *N. eutropha* described herein comprise D23 nucleotides at at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12 or all of the positions that differ in this gene between strains C91 and D23. In embodiments, the *N. eutropha* described herein

[0314] The amoC2 genes are about 99.8% identical (i.e., at 814/816 positions). Accordingly, in some embodiments, the *N. eutropha* described herein comprise D23 nucleotides at at least 1, 2, or all of the positions that differ in this gene between strains C91 and D23. In embodiments, the *N. eutropha* described herein comprise D23 nucleotides at at most 1, 2, or all of the positions that differ in this gene between strains C91 and D23. In embodiments, the *N. eutropha* described herein comprise a gene at least about 99.8%, 99.9%, or 100% identical to the D23 amoC2 gene.

[0315] The amoC3 genes are about 98.9% identical (i.e., at 816/825 positions). Accordingly, in some embodiments, the *N. eutropha* described herein comprise D23 nucleotides at at least 1, 2, 3, 4, 5, 6, 7, 8, 9, or all of the positions that differ in this gene between strains C91 and D23. In embodiments, the *N. eutropha* described herein comprise D23 nucleotides at at most 1, 2, 3, 4, 5, 6, 7, 8, 9, or all of the positions that differ in this gene between strains C91 and D23. In embodiments, the *N. eutropha* described herein comprise a gene at least about 98.9%, 99.0%, 99.2%, 99.4%, 99.6%, 99.8%, or 100% identical to the D23 amoC3 gene.

[0316] The *hao1* genes are about 99.0% identical (i.e., at 1696/1713 positions). Accordingly, in some embodiments, the *N. eutropha* described herein comprise D23 nucleotides at at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, or all of the positions that differ in this gene between strains C91 and D23. In embodiments, the *N. eutropha* described herein comprise D23 nucleotides at at most 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, or all of the positions that differ in this gene between strains C91 and D23. In embodiments, the *N. eutropha* described herein comprise a gene at least about 99.0%, 99.2%, 99.4%, 99.6%, 99.8%, or 100% identical to the D23 *hao1* gene.

[0317] The hao2 genes are about 99.4% identical (i.e., at 1702/1713 positions). Accordingly, in some embodiments, the *N. eutropha* described herein comprise D23 nucleotides at at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, or all of the positions that differ in this gene between strains C91 and D23. In embodiments, the *N. eutropha* described herein comprise D23 nucleotides at at most 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, or all of the positions that differ in this gene between strains C91 and D23. In embodiments, the *N. eutropha* described herein comprise a gene at least about 99.4%, 99.6%, 99.8%, or 100% identical to the D23 hao2 gene.

[0318] The hao3 genes are about 99.2% identical (i.e., at 1700/1713 positions). Accordingly, in some embodiments, the *N. eutropha* described herein comprise D23 nucleotides at at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, or all of the

positions that differ in this gene between strains C91 and D23. In embodiments, the *N. eutropha* described herein comprise D23 nucleotides at at most 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, or all of the positions that differ in this gene between strains C91 and D23. In embodiments, the *N. eutropha* described herein comprise a gene at least about 99.2%, 99.4%, 99.6%, 99.8%, or 100% identical to the D23 hao3 gene.

[0319] The *cycA1* genes are about 98.0% identical (i.e., at 694/708 positions). Accordingly, in some embodiments, the *N. eutropha* described herein comprise D23 nucleotides at at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, or all of the positions that differ in this gene between strains C91 and D23. In embodiments, the *N. eutropha* described herein comprise D23 nucleotides at at most 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, or all of the positions that differ in this gene between strains C91 and D23. In embodiments, the *N. eutropha* described herein comprise a gene at least about 98.0%, 98.2%, 98.4%, 98.6%, 98.8%, 99.0%, 99.2%, 99.4%, 99.6%, 99.8%, or 100% identical to the D23 *cycA1* gene.

[0320] The *cycA2* genes are about 98.7% identical (i.e., at 699/708 positions). Accordingly, in some embodiments, the *N. eutropha* described herein comprise D23 nucleotides at at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, or all of the positions that differ in this gene between strains C91 and D23. In embodiments, the *N. eutropha* described herein comprise D23 nucleotides at at most 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, or all of the positions that differ in this gene between strains C91 and D23. In embodiments, the *N. eutropha* described herein comprise a gene at least about 98.7%, 98.8%, 99.0%, 99.2%, 99.4%, 99.6%, 99.8%, or 100% identical to the D23 *cycA2* gene.

[0321] The *cycA3* genes are about 99.3% identical (i.e., at 703/708 positions). Accordingly, in some embodiments, the *N. eutropha* described herein comprise D23 nucleotides at at least 1, 2, 3, 4, 5, or all of the positions that differ in this gene between strains C91 and D23. In embodiments, the *N. eutropha* described herein comprise D23 nucleotides at at most 1, 2, 3, 4, 5, or all of the positions that differ in this gene between strains C91 and D23. In embodiments, the *N. eutropha* described herein comprise a gene at least about 99.3%, 99.4%, 99.6%, 99.8%, or 100% identical to the D23 *cycA3* gene.

[0322] The *cycB1* genes are about 96.7% identical (i.e., at 696/720 positions). Accordingly, in some embodiments, the *N. eutropha* described herein comprise D23 nucleotides at at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, or all of the positions that differ in this gene between strains C91 and D23. In embodiments, the *N. eutropha* described herein comprise D23 nucleotides at at most 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, or all of the positions that differ in this gene between strains C91 and D23. In embodiments, the *N. eutropha* described herein comprise a gene at least about 96.7%, 96.8%, 97.0%, 97.2%, 97.4%, 97.6%, 97.8%, 98.0%, 98.2%, 98.4%, 98.4%, 98.6%, 98.8%, 99.0%, 99.2%, 99.4%, 99.6%, 99.8%, or 100% identical to the D23 *cycB1* gene.

[0323] The cycB2 genes are about 97.1% identical (i.e., at 702/723 positions). Accordingly, in some embodiments, the *N. eutropha* described herein comprise D23 nucleotides at at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, or all of the positions that differ in this gene

between strains C91 and D23. In embodiments, the *N. eutropha* described herein comprise D23 nucleotides at at most 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, or all of the positions that differ in this gene between strains C91 and D23. In embodiments, the *N. eutropha* described herein comprise a gene at least about 97.1%, 97.2%, 97.4%, 97.6%, 97.8%, 98.0%, 98.2%, 98.4%, 98.4%, 98.6%, 98.8%, 99.0%, 99.2%, 99.4%, 99.6%, 99.8%, or 100% identical to the D23 cycB2 gene.

[0324] The following four paragraphs describe genes and proteins of Table 1 in more detail.

[0325] Ammonia monooxygenase is an enzyme involved in ammonia oxidation, that catalyzes the reaction $\text{NH}_3 + \text{O}_2 + 2\text{e}^- + 2\text{H}^+ \rightleftharpoons \text{NH}_2\text{OH} + \text{H}_2\text{O}$ (Ensign et al., 1993). In *N. eutropha* strain D23, the ammonia monooxygenase operon comprises three genes designated amoA, amoB, and amoC. Strain D23 comprises two copies of the entire operon, and a third copy of amoC. These genes and the corresponding proteins are listed in Table 1 above. In certain embodiments, the *N. eutropha* described herein comprise 1 or 2 ammonia monooxygenase subunit A genes and/or protein of Table 1 (e.g., the D23 sequences of Table 1), or genes and/or proteins similar thereto. In some embodiments, the *N. eutropha* described herein comprise 1 or 2 ammonia monooxygenase subunit B genes and/or proteins of Table 1 (e.g., the D23 sequences of Table 1), or genes and/or proteins similar thereto. In certain embodiments, the *N. eutropha* described herein comprise 1, 2, or 3 ammonia monooxygenase subunit C genes and/or proteins of Table 1 (e.g., the D23 sequences of Table 1), or genes and/or proteins similar thereto. In some embodiments, the *N. eutropha* described herein comprise at least one or two each of (a) an ammonia monooxygenase subunit A gene and/or protein of Table 1 (e.g., the D23 sequences of Table 1), (b) an ammonia monooxygenase subunit B gene and/or protein of Table 1 (e.g., the D23 sequences of Table 1), and (c) an ammonia monooxygenase subunit C gene and/or protein of Table 1 (e.g., the D23 sequences of Table 1). For instance, the *N. eutropha* may comprise all of the ammonia monooxygenase genes and/or proteins of Table 1 (e.g., the D23 sequences of Table 1), or genes and/or proteins similar thereto. Even more specifically, in some embodiments, the *N. eutropha* comprises all of the D23 ammonia monooxygenase genes of Table 1. In some embodiments, the *N. eutropha* comprises all of the D23 ammonia monooxygenase proteins of Table 1. Hydroxylamine oxidoreductases catalyze the general reaction $\text{NH}_2\text{OH} + \text{O}_2 \rightarrow \text{NO}_2^- + \text{H}_2\text{O}$. They typically use heme as a cofactor. *N. eutropha* strain D23 comprises three hydroxylamine oxidoreductases, designated hao1, hao2, and hao3. These genes and the corresponding proteins are listed in Table 1 above. In some embodiments, the *N. eutropha* described herein comprise 1, 2, or 3 hydroxylamine oxidoreductase genes and/or proteins of Table 1 (e.g., the D23 sequences of Table 1), or genes and/or proteins similar thereto. For instance, the *N. eutropha* may comprise all of the hydroxylamine oxidoreductase genes and/or proteins of Table 1 (e.g., the D23 sequences of Table 1), or genes and/or proteins similar thereto. Even more specifically, in some embodiments, the *N. eutropha* comprises all of the D23 hydroxylamine oxidoreductase genes of Table 1. In some embodiments, the *N. eutropha* comprises all of the D23 hydroxylamine oxidoreductase proteins of Table 1.

[0326] The capacity of D23 to aerobically catabolize ammonia as the sole source of energy and reductant requires

two specialized protein complexes, Amo and Hao as well as the cytochromes c554 and c_M552 , which relay the electrons to the quinone pool. The NO reductase activity of c554 is important during ammonia oxidation at low oxygen concentrations. *N. eutropha* strain D23 comprises three cytochrome c554 genes, designated cycA1, cycA2, and cycA3. These genes and the corresponding proteins are listed in Table 1 above. In some embodiments, the *N. eutropha* described herein comprise 1, 2, or 3 cytochrome c554 genes and/or proteins of Table 1 (e.g., the D23 sequences of Table 1), or genes and/or proteins similar thereto. For instance, the *N. eutropha* may comprise all of the cytochrome c554 genes and/or proteins of Table 1 (e.g., the D23 sequences of Table 1), or genes and/or proteins similar thereto. Even more specifically, in some embodiments, the *N. eutropha* comprises all of the D23 cytochrome c554 genes of Table 1. In some embodiments, the *N. eutropha* comprises all of the D23 cytochrome c554 proteins of Table 1.

[0327] The capacity of D23 to aerobically catabolize ammonia as the sole source of energy and reductant requires two specialized protein complexes, Amo and Hao as well as the Cytochromes c554 and c_M552 , which relay the electrons to the quinone pool. Cytochrome c_M552 reduces quinones, with electrons originating from Hao. *N. eutropha* strain D23 comprises two cytochrome c_M552 genes, designated cycB1 and cycB2. These genes and the corresponding proteins are listed in Table 1 above. In some embodiments, the *N. eutropha* described herein comprise 1 or 2 cytochrome c_M552 genes and/or proteins of Table 1 (e.g., the D23 sequences of Table 1), or genes and/or proteins similar thereto. For instance, the *N. eutropha* may comprise both of the cytochrome c_M552 genes and/or proteins of Table 1 (e.g., the D23 sequences of Table 1), or genes and/or proteins similar thereto. Even more specifically, in some embodiments, the *N. eutropha* comprises both of the D23 cytochrome c_M552 genes of Table 1. In some embodiments, the *N. eutropha* comprises both of the D23 Cytochrome c_M552 proteins of Table 1.

[0328] In some embodiments, the *N. eutropha* described herein comprises a combination of genes and/or proteins selected from Table 1. This combination may comprise, for instance, genes and/or proteins listed in the preceding four paragraphs. For instance, the combination may comprise genes and/or proteins from two classes within Table 1. Accordingly, in some embodiments, the *N. eutropha* comprises one or more ammonia monooxygenase genes and/or proteins and one or more hydroxylamine oxidoreductase genes and/or proteins as described in Table 1, or as described in the preceding four paragraphs. In embodiments, the *N. eutropha* comprises one or more ammonia monooxygenase genes and/or proteins and one or more cytochrome c554 genes and/or proteins as described in Table 1, or as described in the preceding four paragraphs. In embodiments, the *N. eutropha* comprises one or more ammonia monooxygenase genes and/or proteins and one or more cytochrome c_M552 genes and/or proteins as described in Table 1, or as described in the preceding four paragraphs. In embodiments, the *N. eutropha* comprises one or more hydroxylamine oxidoreductase genes and/or proteins and one or more cytochrome c554 genes and/or proteins as described in Table 1, or as described in the preceding four paragraphs. In embodiments, the *N. eutropha* comprises one or more hydroxylamine oxidoreductase genes and/or proteins and

one or more cytochrome c_M552 genes and/or proteins as described in Table 1, or as described in the preceding four paragraphs.

[0329] The combination may also comprise genes and/or proteins from three classes within Table 1. Accordingly, in some embodiments, the *N. eutropha* comprises one or more ammonia monooxygenase genes and/or proteins and one or more hydroxylamine oxidoreductase genes and/or proteins and one or more cytochrome $c554$ genes and/or proteins as described in Table 1, or as described in the aforementioned four paragraphs. In embodiments, the *N. eutropha* comprises one or more ammonia monooxygenase genes and/or proteins and one or more hydroxylamine oxidoreductase genes and/or proteins and one or more cytochrome c_M552 genes and/or proteins as described in Table 1, or as described in the aforementioned four paragraphs. In embodiments, the *N. eutropha* comprises one or more ammonia monooxygenase genes and/or proteins and one or more cytochrome $c554$ genes and/or proteins and/or one or more cytochrome c_M552 genes and/or proteins as described in Table 1, or as described in the aforementioned four paragraphs. In embodiments, the *N. eutropha* comprises one or more one or more hydroxylamine oxidoreductase genes and/or proteins and one or more cytochrome $c554$ genes and/or proteins and/or one or more cytochrome c_M552 genes and/or proteins as described in Table 1, or as described in the aforementioned four paragraphs.

[0330] The combination may comprise genes and/or proteins from all four classes within Table 1. Accordingly, in some embodiments, the *N. eutropha* comprises one or more ammonia monooxygenase genes and/or proteins and one or more hydroxylamine oxidoreductase genes and/or proteins and one or more cytochrome $c554$ genes and/or proteins and/or one or more cytochrome c_M552 genes as described in Table 1, or as described in the aforementioned four paragraphs.

[0331] Table 2 (below) lists sequence differences between the D23 and C91 proteins of Table 1. For example, AmoA1 has M at position 1 in C91 but V at position 1 in D23, and this difference is abbreviated as M1V in Table 2. As another example, the D23 CycB1 has an insertion of DDD between residues 194 and 195 of the C91 protein, so that the added residues are residues number 195, 196, and 197 of the D23 protein and this difference is abbreviated as 195insD, 196insD, and 197insD respectively in Table 2. The sequence alignments that form the basis for Table 2 are shown in FIGS. 10-16.

TABLE 2

Amino acid sequence differences between <i>N. eutropha</i> strains D23 and C91	
Protein	Sequence characteristics of D23 compared to C91
1. ammonia monooxygenase	
AmoA1	M1V, M160L, P167A
AmoA2	M1V, M160L, P167A
AmoB1	I33V, V165I
AmoB2	I33V, V165I
AmoC1	N/A
AmoC2	N/A
AmoC3	V79A, I271V

TABLE 2-continued

Amino acid sequence differences between <i>N. eutropha</i> strains D23 and C91	
Protein	Sequence characteristics of D23 compared to C91
2. hydroxylamine oxidoreductase	
Hao1	N85S, V163A, G312E
Hao2	N85S, G312E
Hao3	N85S, G312E
3. cytochrome $c554$	
$c554$ CycA1	A65T, A186T
$c554$ CycA2	A65T
$c554$ CycA3	A65T
4. cytochrome c_M552	
c_M552 CycB1	I63V, S189P, D194G, 195insD, 196insD, 197insD, 206insE, 207insE
c_M552 CycB2	I63V, S189P, 206insE, 207insE

[0332] Accordingly, the *N. eutropha* described herein may comprise one or more of the sequence characteristics listed in Table 2. For instance, the *N. eutropha* may comprise at least 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, or all of the sequence characteristics of Table 2. In some embodiments, the *N. eutropha* comprises no more than 2, 3, 4, 5, 10, 15, 20, 25, 30, or all of the sequence characteristics of Table 2. In embodiments, the *N. eutropha* comprises 1-5, 5-10, 10-15, 15-20, 20-25, 25-30, or all of the sequence characteristics of Table 2. The *N. eutropha* may also comprise fragments of said proteins.

[0333] As to individual categories of genes or proteins, in some embodiments, the *N. eutropha* comprises at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or all of the sequence characteristics of Table 2, Section 1 (which describes ammonia monooxygenases). In embodiments, the *N. eutropha* comprises 1-5, 3-7, 4-8, or 5-10 of the sequence characteristics of Table 2, Section 1. For instance, in some embodiments, the *N. eutropha* comprises at least 1, 2, or 3 sequence characteristics of an amoA gene or protein as listed in Table 2, and/or no more than 2 or 3 of these characteristics. The *N. eutropha* may also comprise at least 1 or 2 sequence characteristics of an amoB gene or protein as listed in Table 2. In addition, the *N. eutropha* may comprise at least 1 or 2 sequence characteristics of the amoC3 gene as listed in Table 2. The *N. eutropha* may also comprise fragments of said proteins.

[0334] With respect to hao genes and proteins, the *N. eutropha* may comprise at least 1, 2, 3, 4, 5, 6, 7, 8, or all of the sequence characteristics of Table 2, Section 2 (which describes hydroxylamine oxidoreductases). In embodiments, the *N. eutropha* comprises 1-4, 2-5, 3-6, or 4-8 of the sequence characteristics of Table 2, Section 2. The *N. eutropha* may also comprise at least 1, 2 or 3 sequence characteristics of Hao1 as listed in Table 1, and/or no more than 2 or 3 of these characteristics. The *N. eutropha* may also comprise at least 1 or 2 sequence characteristics of Hao2 or Hao3 as listed in Table 2. The *N. eutropha* may also comprise fragments of said proteins.

[0335] Turning now to cytochrome $c554$, the *N. eutropha* may comprise at least 1, 2, 3, 4, or all of the sequence characteristics of Table 2, Section 3 (which describes cytochrome $c554$). In embodiments, the *N. eutropha* comprises at most 2, 3, 4, or all of the sequence characteristics of Table 2 Section 3. In embodiments, the *N. eutropha* comprises at least 1 or 2 sequence characteristics of

cytochrome c554 CycA1 as listed in Table 2. The *N. eutropha* may also comprise at least 1 sequence characteristic of c554 CycA2 or c554 CycA3 as listed in Table 2. The *N. eutropha* may also comprise fragments of said proteins.

[0336] With respect to the c_M552 genes and proteins, the *N. eutropha* may comprise at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or all of the sequence characteristics of Table 2, Section 4 (which describes cytochrome c_M552). In embodiments, the *N. eutropha* comprises at most 2, 3, 4, 5, 6, 7, 8, 9, 10, or all the sequence characteristics of Table 2, Section 4. For instance, in embodiments the *N. eutropha* comprises 1-5, 2-7, 3-8, or 5-10 sequence characteristics of Table 2, Section 4. In embodiments, at least 1, 2, 3, 4, 5, 6, or 7 sequence characteristics of c_M552 CycB1 as listed in Table 2, and/or no more than 2, 3, 4, 5, 6, or 7 of these characteristics. The *N. eutropha* may also comprise at least 1, 2, or 3 sequence characteristics of c_M552 CycB2 as listed in Table 2, and/or no more than 2 or 3 of these characteristics. The *N. eutropha* may also comprise fragments of said proteins.

[0337] It is understood that the paragraphs above, which refer to sequence characteristics of various *N. eutropha* proteins, also describe the sequences of nucleic acids that encode these proteins.

[0338] The sequencing analysis described herein revealed that strain D23 lacks plasmids. Consequently, in some embodiments, the *N. eutropha* bacterium lacks plasmids, i.e., all of its DNA is contained in the chromosome. In some embodiments, the *N. eutropha* bacterium lacks endogenous plasmids, but carries one or more transgenic plasmids.

[0339] This D23 strain is not believed to be a product of nature, but rather has acquired certain mutations and characteristics during an extended period of culture and selection in the laboratory. For instance, D23 has an ability to grow in conditions of greater than about 200 or 250 mM NH₄⁺ for more than 24 hours.

[0340] In some embodiments, the *N. eutropha* disclosed herein differ from naturally occurring bacteria in the abundance of siderophores. For instance, the *N. eutropha* may have elevated or reduced levels of siderophores compared to *N. eutropha* C91. Generally, siderophores are secreted iron-chelating compounds that help bacteria scavenge iron from their environment. Some siderophores are peptides, and others are small organic molecules.

[0341] The AOBs, for example, *N. eutropha* contemplated in this disclosure may comprise mutations relative to wild-type *N. eutropha* and/or the *N. eutropha* sequences disclosed herein. These mutations may, e.g., occur spontaneously, be introduced by random mutagenesis, or be introduced by targeted mutagenesis. For instance, the *N. eutropha* may lack one or more genes or regulatory DNA sequences that wild-type *N. eutropha* typically comprises. The *N. eutropha* may also comprise point mutations, substitutions, insertions, deletions, and/or rearrangements relative to the sequenced strain or a wild-type strain. The *N. eutropha* may be a purified preparation of optimized *N. eutropha*.

[0342] In certain embodiments, the *N. eutropha* is transgenic. For instance, it may comprise one or more genes or regulatory DNA sequences that wild-type *N. eutropha* D23 lacks. More particularly, the *N. eutropha* may comprise, for instance, a reporter gene, a selective marker, a gene encoding an enzyme, or a promoter (including an inducible or repressible promoter). In some embodiments the additional gene or regulatory DNA sequence is integrated into the bacterial chromosome; in some embodiments the additional

gene or regulatory DNA sequence is situated on a plasmid, for instance a plasmid related to a plasmid found in *N. eutropha* N91.

[0343] In some preferred embodiments, the *N. eutropha* differs by at least one nucleotide from naturally occurring bacteria. For instance, the *N. eutropha* may differ from naturally occurring bacteria in a gene or protein that is part of a relevant pathway, e.g., an ammonia metabolism pathway, a urea metabolism pathway, or a pathway for producing nitric oxide or nitric oxide precursors. More particularly, the *N. eutropha* may comprise a mutation that elevates activity of the pathway, e.g., by increasing levels or activity of an element of that pathway.

[0344] The above-mentioned mutations can be introduced using any suitable technique. Numerous methods are known for introducing mutations into a given position. For instance, one could use site-directed mutagenesis, oligonucleotide-directed mutagenesis, or site-specific mutagenesis. Non-limiting examples of specific mutagenesis protocols are described in, e.g., Mutagenesis, pp. 13.1-13.105 (Sambrook and Russell, eds., Molecular Cloning A Laboratory Manual, Vol. 3, 3.sup.rd ed. 2001). In addition, non-limiting examples of well-characterized mutagenesis protocols available from commercial vendors include, without limitation, Altered Sites® II in vitro Mutagenesis Systems (Promega Corp., Madison, Wis.); Erase-a-Base® System (Promega, Madison, Wis.); GeneTailor™ Site-Directed Mutagenesis System (Invitrogen, Inc., Carlsbad, Calif.); QuikChange® II Site-Directed Mutagenesis Kits (Stratagene, La Jolla, Calif.); and Transformer™ Site-Directed Mutagenesis Kit (BD-Clontech, Mountain View, Calif.).

[0345] In some embodiments, the preparation of ammonia oxidizing bacteria may comprise a concentration or amount of ammonia oxidizing bacteria in order to at least partially treat a condition or disease. The preparation of ammonia oxidizing bacteria may comprise a concentration or amount of ammonia oxidizing bacteria in order to alter, e.g., reduce or increase, an amount, concentration or proportion of a bacterium, or genus of bacteria, on a surface, e.g., a skin surface. The bacteria may be non-pathogenic or pathogenic, or potentially pathogenic.

[0346] In some embodiments, the preparation of ammonia oxidizing bacteria may comprise between about 10⁸ to about 10¹⁴ CFU/L. The preparation may comprise at least 10⁸, 10⁹, 10¹⁰, 10¹¹, 2×10¹¹, 5×10¹¹, 10¹², 2×10¹², 5×10¹², 10¹³, 2×10¹³, 5×10¹³, or 10¹⁴; or about 10⁸-10⁹, 10⁹-10¹⁰, 10¹¹-10¹¹, 10¹¹-10¹², 10¹²-10¹³, or 10¹³-10¹⁴ CFU/L. In certain aspects, the preparation may comprise between about 1×10⁹ CFU/L to about 10×10⁹ CFU/L. In certain aspects, the preparation may comprise between about 1×10⁹ CFU to about 10×10⁹ CFU.

[0347] In some embodiments, the preparation of ammonia oxidizing bacteria may comprise between about 0.1 milligrams (mg) and about 1000 mg of ammonia oxidizing bacteria. In certain aspects, the preparation may comprise between about 50 mg and about 1000 mg of ammonia oxidizing bacteria. The preparation may comprise between about 0.1-0.5 mg, 0.2-0.7 mg, 0.5-1.0 mg, 0.5-2 mg, 0.5-5 mg, 2.5-5 mg, 2.5-7.0 mg, 5.0-10 mg, 7.5-15 mg, 10-15 mg, 15-20 mg, 15-25 mg, 20-30 mg, 25-50 mg, 25-75 mg, 50-75 mg, 50-100 mg, 75-100 mg, 100-200 mg, 200-300 mg, 300-400 mg, 400-500 mg, 500-600 mg, 600-700 mg, 700-

800 mg, 800-900 mg, 900-1000 mg, 100-250 mg, 250-500 mg, 100-500 mg, 500-750 mg, 750-1000 mg, or 500-1000 mg.

[0348] In some embodiments, the preparation of ammonia oxidizing bacteria may comprise a mass ratio of ammonia oxidizing bacteria to an excipient, e.g., a pharmaceutically acceptable excipient or a cosmetically acceptable excipient in a range of about 0.1 grams per liter to about 1 gram per liter. The preparation may comprise a mass ratio of ammonia oxidizing bacteria to an excipient in a range of about 0.1-0.2, 0.2-0.3, 0.1-0.5, 0.2-0.7, 0.5-1.0, or 0.7-1.0 grams per liter.

[0349] In some embodiments, the preparation of ammonia oxidizing bacteria may be in a growth state. A growth state may be provided by exposing ammonia oxidizing bacteria to an environment that may promote growth. The growth state may be a state, e.g., ammonia oxidizing bacteria in an environment that allows immediate availability of ammonia oxidizing bacteria to convert ammonium ions (NH_4^+) to nitrite (NO_2^-). The growth state may comprise providing ammonia oxidizing bacteria in an environment having a pH of greater than about 7.6. The growth state may also comprise providing ammonia oxidizing bacteria in an environment having ammonia, ammonium salts, and/or urea, trace minerals and sufficient oxygen and carbon dioxide, as described above in Section 1.

[0350] In some embodiments, the preparation of ammonia oxidizing bacteria may be in a polyphosphate loading state, wherein the state or the environment, e.g., a media, e.g., a culture media, e.g., a growth media, may have a pH of less than about 7.4. Levels of at least one of ammonia, ammonium ions, and urea may be between about 10 micromolar and 200 millimolar. Levels of trace materials may be between 0.1 micromolar iron and 20 micromolar iron. Levels of oxygen may be between about 5% and 100% oxygen saturation. Levels of carbon dioxide may be between/less than about zero and 200 ppm, and phosphate levels greater than about 10 micromolar. The purpose of the polyphosphate loading state is to provide AOB with ammonia and oxygen such that ATP can be produced, but to deny them carbon dioxide and carbonate such that they are unable to use that ATP to fix carbon dioxide and instead use that ATP to generate polyphosphate which may be stored.

[0351] In some embodiments, the preparation of ammonia oxidizing bacteria may be in a storage state. A storage state may be defined as ammonia oxidizing bacteria in an environment in which they may be stored to be later revived. The storage state may be a state, e.g., ammonia oxidizing bacteria in an environment that allows availability of ammonia oxidizing bacteria after being revived, e.g., after being placed in an environment promoting a growth state for a predetermined period of time.

[0352] The storage state may comprise providing ammonia oxidizing bacteria in an environment having a pH of less than about 7.4. The storage state may also comprise providing ammonia oxidizing bacteria in an environment having ammonia, ammonia salts, and/or urea, trace minerals, oxygen, and low concentrations of carbon dioxide, as described above in Section 1.

[0353] Storage may also be accomplished by storing at 4° C. for up to several months. The storage buffer in some embodiments may comprise 50 mM Na_2HPO_4 -2 mM MgCl_2 (pH 7.6).

[0354] In some embodiments, ammonia oxidizing bacteria may be cryopreserved. A 1.25 ml of ammonia oxidizing

bacteria mid-log culture may be added to a 2 ml cryotube and 0.75 ml of sterile 80% glycerol. Tubes may be shaken gently, and incubate at room temperature for 15 min to enable uptake of the cryoprotective agents by the cells. The tubes may be directly stored in a -80° C. freezer for freezing and storage.

[0355] For resuscitation of cultures, frozen stocks may be thawed on ice for 10-20 minutes, and then centrifuged at 8,000×g for 3 minutes at 4° C. The pellet may be washed by suspending it in 2 ml AOB medium followed by another centrifugation at 8,000×g for 3 minutes at 4° C. to reduce potential toxicity of the cryoprotective agents. The pellet may be resuspended in 2 ml of AOB medium, inoculated into 50 ml of AOB medium containing 50 mM NH_4^+ , and incubated in dark at 30° C. by shaking at 200 rpm.

[0356] In some embodiments, the preparation of ammonia oxidizing bacteria may comprise ammonia oxidizing bacteria in a storage state and/or ammonia oxidizing bacteria in a polyphosphate loading state and/or ammonia oxidizing bacteria in a growth state.

[0357] Without wishing to be bound by theory, by maintaining ammonia oxidizing bacteria under conditions or in an environment of low carbon dioxide, with sufficient oxygen and ammonia, they may accumulate polyphosphate for a pre-determined period, e.g., for a period of about one doubling time, e.g., for about 8-12 hours, e.g., for about 10 hours. The ammonia oxidizing bacteria may accumulate sufficient polyphosphate to extend their storage viability, storage time, and accelerate their revival. This may occur with or without the addition of buffer and ammonia.

[0358] The presence of sufficient stored polyphosphate may allow the ammonia oxidizing bacteria the ATP resources to maintain metabolic activity even in the absence of ammonia and oxygen, and to survive insults that would otherwise be fatal.

[0359] The process of oxidation of ammonia to generate ATP has two steps. The first step is the oxidation of ammonia to hydroxylamine by ammonia monooxygenase (Amo), followed by the conversion of hydroxylamine to nitrite by hydroxylamine oxidoreductase (Hao). Electrons from the second step (conversion of hydroxylamine to nitrite) are used to power the first step (oxidation of ammonia to hydroxylamine).

[0360] If an ammonia oxidizing bacteria does not have hydroxylamine to generate electrons for Amo, then hydroxylamine is not available for Hao. For example, acetylene irreversibly inhibits the enzyme crucial for the first step in the oxidation of ammonia to nitrite, the oxidation of ammonia to hydroxylamine. Once AOB are exposed to acetylene, Amo is irreversibly inhibited and new enzyme must be synthesized before hydroxylamine can be generated. In a normal consortium biofilm habitat, AOB may share and receive hydroxylamine from other AOB (even different strains with different susceptibilities to inhibitors) and so the biofilm tends to be more resistant to inhibitors such as acetylene than an individual organism. AOB can use stored polyphosphate to synthesize new Amo, even in the absence of hydroxylamine.

[0361] Any embodiment, preparation, composition, or formulation of ammonia oxidizing bacteria discussed herein may comprise, consist essentially of, or consist of optionally axenic ammonia oxidizing bacteria.

3. METHODS OF PRODUCING *N. EUTROPHA*

[0362] Methods of culturing various *Nitrosomonas* species are known in the art. *N. eutropha* may be cultured, for example, using *N. europaea* medium as described in Example 2 below. Ammonia oxidizing bacteria may be cultured, for example, using the media described in Table 3 or Table 4, above.

[0363] *N. eutropha* may be grown, for example, in a liquid culture or on plates. Suitable plates include 1.2% R2A agar, 1.2% agar, 1.2% agarose, and 1.2% agarose with 0.3 g/L pyruvate.

[0364] In some embodiments, ammonia oxidizing bacteria, such as *N. eutropha* is cultured in organic free media. One advantage of using organic free media is that it lacks substrate for heterotrophic bacteria to metabolize except for that produced by the autotrophic bacteria. Another advantage of using the as-grown culture is that substantial nitrite accumulates in the culture media, and this nitrite is also inhibitory of heterotrophic bacteria and so acts as a preservative during storage.

[0365] In some embodiments, ammonia oxidizing bacteria such as an *N. eutropha* strain with improved, e.g., optimized, properties is produced by an iterative process of propagation and selecting for desired properties. In some embodiments, the selection and propagation are carried out simultaneously. In some embodiments, the selection is carried out in a reaction medium (e.g., complete *N. europaea* medium) comprising 50 mM, 75 mM, 100 mM, 125 mM, 150 mM, 175 mM, 200 mM, 225 mM, 250 mM, 275 mM, or 300 mM NH_4^+ , e.g., at least 200 mM NH_4^+ . In some embodiments, the period of propagation and/or selection is at least 1, 2, 3, or 6 months. In embodiments, the period of propagation and/or selection is at least 1, 2, 4, 6, 8, or 10 years.

[0366] In some aspects, the ammonia oxidizing bacteria, such as the *N. eutropha* are manufactured on a commercial scale. In some embodiments, commercial scale refers to a liquid culturing method with a culture medium volume of at least 10,000, 20,000, 30,000, 50,000, or 100,000 liters (L). In some embodiments, the bacteria are produced in a bioreactor. The bioreactor may maintain the bacteria at a constant temperature, e.g., about 26-30 degrees Celsius using, for example a thermal jacket for insulation, a temperature sensor, and a heating or cooling element. The bioreactor may have an apparatus for stirring the culture to improve distribution of nutrients like ammonia, urea, oxygen, carbon dioxide, and various minerals. The bioreactor may also have an inlet tube for addition of new medium, and an outlet tube for collection of cells. The bioreactor may also have an aerator for distributing oxygen and/or carbon dioxide to the culture. The bioreactor may be, e.g., a batch reactor, a fed batch reactor, or a continuous reactor. In some embodiments, commercial scale production of *N. eutropha* yields a batch of 1,000 to 100,000 L per day at about 10^{12} CFU/liter and 1,000 to 100,000. The commercial scale production may yield e.g., a batch of 1,000-5,000, 5,000-10,000, 10,000-50,000, or 50,000-100,000 L/day. The commercial scale production may yield e.g., a batch of 1,000-5,000, 5,000-10,000, 10,000-50,000, or 50,000-100,000 L per batch. In some embodiments, the yield is at a concentration of at least 10^{10} , 10^{11} , 2×10^{11} , 5×10^{11} , or 10^{12} , or about 10^{10} - 10^{11} , 10^{11} - 10^{12} , 10^{12} - 10^{13} , or 10^{13} - 10^{14} CFU/L

[0367] In some embodiments, typically including commercial scale production, quality control (QC) testing steps are carried out. The general steps of QC typically comprise,

1) culturing *N. eutropha*, 2) performing a testing step on the culture or an aliquot thereof, and 3) obtaining a value from the testing step, and optionally: 4) comparing the obtained value to a reference value or range of acceptable values, and 5) if the obtained value meets the acceptable reference value or range, then classifying the culture as acceptable, and if the obtained value does not meet the acceptable reference value or range, then classifying the culture as unacceptable. If the culture is classified as acceptable, the culture may, e.g., be allowed to continue growing and/or may be harvested and added to a commercial product. If the culture is classified as unacceptable, the culture may, e.g., be safely disposed of or the defect may be remedied.

[0368] The testing step may comprise measuring the optical density (OD) of the culture. OD is measured in a spectrophotometer, and provides information on the amount of light transmitted through the sample as distinguished from light absorbed or scattered. In some embodiments, the OD600 (e.g., optical density of light with a wavelength of 600 nm) may be determined. This measurement typically indicates the concentration of cells in the medium, where a higher optical density corresponds to a higher cell density.

[0369] The testing step may comprise measuring the pH of the culture. The pH of an *N. eutropha* culture indicates the rate of nitrogen oxidation, and can also indicate whether the culture comprises a contaminating organism. pH may be measured using, e.g., a pH-sensing device comprising a electrode (such as a hydrogen electrode, quinhydrone-Electrode, antimony electrode, glass electrode), a pH-sensing device comprising a semiconductor, or a color indicator reagent such as pH paper.

[0370] In certain embodiments, producing the ammonia oxidizing bacteria such as *N. eutropha* comprises carrying out various quality control steps. For instance, one may test the medium in which the *N. eutropha* is grown, e.g., to determine whether it has an appropriate pH, whether it has a sufficiently low level of waste products, and/or whether it has a sufficiently high level of nutrients. One may also test for the presence of contaminating organisms. A contaminating organism is typically an organism other than an ammonia oxidizing bacteria such as *N. eutropha*, for instance an organism selected *Microbacterium* sp., *Alcaligenaceae* bacterium, *Caulobacter* sp., *Burkholderia multivorans*, *Escherichia coli*, *Klebsiella pneumoniae*, and *Staphylococcus aureus*. One may test for contaminants by, e.g., extracting DNA, amplifying it, and sequencing a conserved gene such as 16S rRNA. One may also test for contaminants by plating culture on agar plates and observing colony morphology. *N. eutropha* typically forms red colonies, so non-red colonies are often indicative of contaminating organisms.

4. COMPOSITIONS COMPRISING AMMONIA OXIDIZING BACTERIA; COMPOSITIONS COMPRISING *N. EUTROPHA*

[0371] The present disclosure provides, inter alia, compositions comprising ammonia oxidizing bacteria, e.g., a preparation of ammonia oxidizing bacteria, or a purified preparation of ammonia oxidizing bacteria e.g., a natural product, or a fortified natural product. The compositions comprising ammonia oxidizing bacteria, e.g., a preparation of ammonia oxidizing bacteria, or a purified preparation of ammonia oxidizing bacteria may be provided in a cosmetic

product or a therapeutic product. The preparation may comprise, inter alia, at least one of ammonia, ammonium salts, and urea.

[0372] The present disclosure provides, inter alia, compositions comprising *N. eutropha*, e.g., a purified preparation of an optimized *N. eutropha*. In some embodiments, the *N. eutropha* in the compositions has at least one property selected from an optimized growth rate, an optimized NH_4^+ oxidation rate, and an optimized resistance to NH_4^+ .

[0373] In some aspects, the present disclosure provides compositions with a defined number of species. For instance, this disclosure provides a composition having *N. eutropha* and one other type of organism, and no other types of organism. In other examples, the composition has *N. eutropha* and 2, 3, 4, 5, 6, 7, 8, 9, or 10 other types of organism, and no other types of organism. The other type of organism in this composition may be, for instance, a bacterium, such as an ammonia-oxidizing bacterium. Suitable ammonia-oxidizing bacteria for this purpose include those in the genera *Nitrosomonas*, *Nitrosococcus*, *Nitrospira*, *Nitrosocystis*, *Nitrosolobus*, or *Nitrosovibrio*.

[0374] In some embodiments, the composition comprising *N. eutropha* provides conditions that support *N. eutropha* viability. For instance, the composition may promote *N. eutropha* growth and metabolism or may promote a dormant state (e.g., freezing) from which viable *N. eutropha* can be recovered. When the composition promotes growth or metabolism, it may contain water and/or nutrients that *N. eutropha* consumes, e.g., as ammonium, ammonia, urea, oxygen, carbon dioxide, or trace minerals. In some embodiments, the composition comprising ammonia oxidizing bacteria provides conditions that support ammonia oxidizing bacteria viability. For instance, the composition may promote ammonia oxidizing bacteria growth and metabolism or may promote a dormant state (e.g., freezing) or storage state as described herein, from which viable ammonia oxidizing bacteria can be recovered. When the composition promotes growth or metabolism, it may contain water and/or nutrients that ammonia oxidizing bacteria consumes, e.g., as ammonium ions, ammonia, urea, oxygen, carbon dioxide, or trace minerals.

[0375] In some embodiments, one or more other organisms besides ammonia oxidizing bacteria may be included in the preparation of ammonia oxidizing bacteria. For example, an organism of the genus selected from the group consisting of *Lactobacillus*, *Streptococcus*, *Bifidobacter*, and combinations thereof, may be provided in the preparation of ammonia oxidizing bacteria. In some embodiments, the preparation may be substantially free of other organisms.

[0376] Preparations of ammonia oxidizing bacteria may comprise between about 10^8 to about 10^{14} CFU/L. The preparation may comprise at least about 10^8 , 10^9 , 10^{10} , 10^{11} , 2×10^{11} , 5×10^{11} , 10^{12} , 2×10^{12} , 5×10^{12} , 10^{13} , 2×10^{13} , 5×10^{13} , or 10^{14} ; or about 10^8 - 10^9 , 10^9 - 10^{10} , 10^{10} - 10^{11} , 10^{11} - 10^{12} , 10^{12} - 10^{13} , or 10^{13} - 10^{14} CFU/L.

[0377] In some embodiments, the preparation may comprise at least 10^8 , 10^9 , 10^{10} , 10^{11} , 2×10^{11} , 5×10^{11} , 10^{12} , 2×10^{12} , 5×10^{12} , 10^{13} , 2×10^{13} , 5×10^{13} , or 10^{14} ; or about 10^8 - 10^9 , 10^9 - 10^{10} , 10^{10} - 10^{11} , 10^{11} - 10^{12} , 10^{12} - 10^{13} , or 10^{13} - 10^{14} CFU/ml.

[0378] In some embodiments, the preparation may comprise between about 1×10^9 to about 10×10^9 CFU/L. In some embodiments, the preparation may comprise about 3×10^{10} CFU, e.g., 3×10^{10} CFU per day. In some embodiments, the

preparation may comprise about 1×10^9 to about 10×10^9 CFU, e.g., about 1×10^9 to about 10×10^9 CFU per day.

[0379] In some embodiments, the preparation of ammonia oxidizing bacteria may comprise between about 0.1 milligrams (mg) and about 1000 mg of ammonia oxidizing bacteria. In certain aspects, the preparation may comprise between about 50 mg and about 1000 mg of ammonia oxidizing bacteria. The preparation may comprise between about 0.1-0.5 mg, 0.2-0.7 mg, 0.5-1.0 mg, 0.5-2 mg, 0.5-5 mg, 2.5-5 mg, 2.5-7.0 mg, 5.0-10 mg, 7.5-15 mg, 10-15 mg, 15-20 mg, 15-25 mg, 20-30 mg, 25-50 mg, 25-75 mg, 50-75 mg, 50-100 mg, 75-100 mg, 100-200 mg, 200-300 mg, 300-400 mg, 400-500 mg, 500-600 mg, 600-700 mg, 700-800 mg, 800-900 mg, 900-1000 mg, 100-250 mg, 250-500 mg, 100-500 mg, 500-750 mg, 750-1000 mg, or 500-1000 mg.

[0380] In some embodiments, the preparation of ammonia oxidizing bacteria may comprise a mass ratio of ammonia oxidizing bacteria to an excipient, e.g., a pharmaceutically acceptable excipient or a cosmetically acceptable excipient in a range of about 0.1 grams per liter to about 1 gram per liter. The preparation may comprise a mass ratio of ammonia oxidizing bacteria to an excipient in a range of about 0.1-0.2, 0.2-0.3, 0.1-0.5, 0.2-0.7, 0.5-1.0, or 0.7-1.0 grams per liter.

[0381] Advantageously, a formulation may have a pH that promotes AOB, e.g., *N. eutropha* viability, e.g., metabolic activity. Urea would hydrolyze to ammonia and would raise the pH to 7 to 8. AOB are very active at this pH range and would lower the pH to about 6 where the NH_3 converts to ammonium and is unavailable. Lower pH levels, e.g. about pH 4, are also acceptable. The ammonia oxidizing bacteria, e.g., *N. eutropha* may be combined with one or more pharmaceutically or cosmetically acceptable excipients. In some embodiments, "pharmaceutically acceptable excipient" refers to a pharmaceutically-acceptable material, composition, or vehicle, such as a liquid or solid filler, diluent, solvent, or encapsulating material. In some embodiments, each excipient is "pharmaceutically acceptable" in the sense of being compatible with the other ingredients of a pharmaceutical formulation, and suitable for use in contact with the tissue or organ of humans and animals without excessive toxicity, irritation, allergic response, immunogenicity, or other problems or complications, commensurate with a reasonable benefit/risk ratio. See, Remington: The Science and Practice of Pharmacy, 21st ed.; Lippincott Williams & Wilkins: Philadelphia, Pa., 2005; Handbook of Pharmaceutical Excipients, 6th ed.; Rowe et al., Eds.; The Pharmaceutical Press and the American Pharmaceutical Association: 2009; Handbook of Pharmaceutical Additives, 3rd ed.; Ash and Ash Eds.; Gower Publishing Company: 2007; Pharmaceutical Preformulation and Formulation, 2nd ed.; Gibson Ed.; CRC Press LLC: Boca Raton, Fla., 2009.

[0382] In some embodiments, a cosmetically acceptable excipient refers to a cosmetically acceptable material, composition, or vehicle, such as a liquid or solid filler, diluent, solvent, or encapsulating material. In some embodiments, each excipient is cosmetically acceptable in the sense of being compatible with the other ingredients of a cosmetic formulation, and suitable for use in contact with the tissue or organ of humans and animals without excessive toxicity, irritation, allergic response, immunogenicity, or other problems or complications, commensurate with a reasonable benefit/risk ratio.

[0383] While it is possible for the active ingredient, e.g., ammonia oxidizing bacteria, e.g., *N. eutropha*, to be administered alone, in many embodiments it present in a pharmaceutical formulation or composition. Accordingly, this disclosure provides a pharmaceutical formulation comprising ammonia oxidizing bacteria, for example, *N. eutropha* and a pharmaceutically acceptable excipient. Pharmaceutical compositions may take the form of a pharmaceutical formulation as described below.

[0384] The pharmaceutical formulations described herein include those suitable for oral, parenteral (including subcutaneous, intradermal, intramuscular, intravenous, and intraarticular), inhalation (including fine particle dusts or mists which may be generated by means of various types of metered doses, pressurized aerosols, nebulizers or insufflators, and including intranasally or via the lungs), rectal and topical (including dermal, transdermal, transmucosal, buccal, sublingual, and intraocular) administration, although the most suitable route may depend upon, for example, the condition and disorder of the recipient.

[0385] The formulations may conveniently be presented in unit dosage form and may be prepared by any of the methods known in the art of pharmacy. Typically, methods include the step of bringing the active ingredient (e.g., ammonia oxidizing bacteria, e.g., *N. eutropha*) into association with a pharmaceutical carrier which constitutes one or more accessory ingredients. In general the formulations are prepared by uniformly and intimately bringing into association the active ingredient with liquid carriers or finely divided solid carriers or both and then, if necessary, shaping the product into the desired formulation.

[0386] Formulations may be presented as discrete units such as capsules, cachets or tablets, each containing a predetermined amount of, e.g., *N. eutropha*; as a powder or granules; as a solution or a suspension in an aqueous liquid or a non-aqueous liquid; or as an oil-in-water liquid emulsion or a water-in-oil liquid emulsion. The active ingredient may also be presented as a bolus, electuary or paste. Various pharmaceutically acceptable carriers and their formulation are described in standard formulation treatises, e.g., Remington's Pharmaceutical Sciences by E. W. Martin. See also Wang, Y. J. and Hanson, M. A., Journal of Parenteral Science and Technology, Technical Report No. 10, Supp. 42:2 S, 1988.

[0387] The ammonia oxidizing bacteria, e.g., *N. eutropha* compositions can, for example, be administered in a form suitable for immediate release or extended release. Suitable examples of sustained-release systems include suitable polymeric materials, for example semi-permeable polymer matrices in the form of shaped articles, e.g., films, or microcapsules; suitable hydrophobic materials, for example as an emulsion in an acceptable oil; or ion exchange resins. Sustained-release systems may be administered orally; rectally; parenterally; intracisternally; intravaginally; intraperitoneally; topically, for example as a powder, ointment, gel, drop or transdermal patch; buccally; or as a spray.

[0388] Preparations for administration can be suitably formulated to give controlled release of ammonia oxidizing bacteria, e.g., *N. eutropha*. For example, the pharmaceutical compositions may be in the form of particles comprising one or more of biodegradable polymers, polysaccharide jellifying and/or bioadhesive polymers, or amphiphilic polymers.

These compositions exhibit certain biocompatibility features which allow a controlled release of an active substance. See U.S. Pat. No. 5,700,486.

[0389] Exemplary compositions include suspensions which can contain, for example, microcrystalline cellulose for imparting bulk, alginic acid or sodium alginate as a suspending agent, methylcellulose as a viscosity enhancer, dicalcium phosphate, starch, magnesium stearate and/or lactose and/or other excipients, binders, extenders, disintegrants, diluents and lubricants, mannitol, lactose, sucrose and/or cyclodextrins. Also included in such formulations may be high molecular weight excipients such as celluloses (avicel) or polyethylene glycols (PEG). Such formulations can also include an excipient to aid mucosal adhesion such as hydroxy propyl cellulose (HPC), hydroxy propyl methyl cellulose (HPMC), sodium carboxy methyl cellulose (SCMC), maleic anhydride copolymer (e.g., Gantrez), and agents to control release such as polyacrylic copolymer (e.g. Carbopol 934). Lubricants, glidants, flavors, coloring agents and stabilizers may also be added for ease of fabrication and use. The surfactant may be a zwitterionic surfactant, a non-ionic surfactant, or an anionic surfactant.

[0390] Excipients, such as surfactants that may be used with embodiments of the present disclosure may include one or more of cocamidopropyl betaine (ColaTeric COAB), polyethylene sorbitol ester (e.g., Tween 80), ethoxylated lauryl alcohol (RhodaSurf 6 NAT), sodium laureth sulfate/lauryl glucoside/cocamidopropyl betaine (Plantapon 611 L UP), sodium laureth sulfate (e.g., RhodaPex ESB 70 NAT), alkyl polyglucoside (e.g., Plantaren 2000 N UP), sodium laureth sulfate (Plantaren 200), Dr. Bronner's Castile soap, Dr. Bronner's Castile baby soap, Lauramine oxide (ColaLux Lo), sodium dodecyl sulfate (SDS), polysulfonate alkyl polyglucoside (PolySufanate 160 P), sodium lauryl sulfate (Stepanol-WA Extra K), and combinations thereof. Dr. Bronner's Castile soap and Dr. Bronner's baby soap comprises water, organic coconut oil, potassium hydroxide, organic olive oil, organic fair deal hemp oil, organic jojoba oil, citric acid, and tocopherol.

In some embodiments, surfactants may be used with ammonia oxidizing bacteria in amounts that allow nitrite production to occur. In some embodiments, the preparation may have less than about 0.0001% to about 10% of surfactant. In some embodiments, the preparation may have between about 0.1% and about 10% surfactant. In some embodiments, the concentration of surfactant used may be between about 0.0001% and about 10%. In some embodiments, the preparation may be substantially free of surfactant.

[0391] In some embodiments, the formulation, e.g., preparation, may include other components that may enhance effectiveness of ammonia oxidizing bacteria, or enhance a treatment or indication.

[0392] In some embodiments, a chelator may be included in the preparation. A chelator may be a compound that may bind with another compound, e.g., a metal. The chelator may provide assistance in removing an unwanted compound from an environment, or may act in a protective manner to reduce or eliminate contact of a particular compound with an environment, e.g., ammonia oxidizing bacteria, e.g. a preparation of ammonia oxidizing bacteria, e.g., an excipient. In some embodiments, the preparation may be substantially free of chelator.

[0393] Formulations may also contain anti-oxidants, buffers, bacteriostats that prevent the growth of undesired

bacteria, solutes, and aqueous and non-aqueous sterile suspensions which may include suspending agents and thickening agents. The formulations may be presented in unit-dose or multi-dose containers, for example sealed ampoules and vials, and may be stored in a freeze-dried (lyophilised) condition requiring only the addition of a sterile liquid carrier, for example saline or water-for-injection, immediately prior to use. Extemporaneous solutions and suspensions may be prepared from powders, granules and tablets of the kind previously described. Exemplary compositions include solutions or suspensions which can contain, for example, suitable non-toxic, pharmaceutically acceptable diluents or solvents, such as mannitol, 1,3-butanediol, water, Ringer's solution, an isotonic sodium chloride solution, or other suitable dispersing or wetting and suspending agents, including synthetic mono- or diglycerides, and fatty acids, including oleic acid, or Cremaphor. An aqueous carrier may be, for example, an isotonic buffer solution at a pH of from about 3.0 to about 8.0, a pH of from about 3.5 to about 7.4, for example from 3.5 to 6.0, for example from 3.5 to about 5.0. Useful buffers include sodium citrate-citric acid and sodium phosphate-phosphoric acid, and sodium acetate/acetic acid buffers. The composition in some embodiments does not include oxidizing agents.

Excipients that can be included are, for instance, proteins, such as human serum albumin or plasma preparations. If desired, the pharmaceutical composition may also contain minor amounts of non-toxic auxiliary substances, such as wetting or emulsifying agents, preservatives, and pH buffering agents and the like, for example sodium acetate or sorbitan monolaurate. In some embodiments, excipients, e.g., a pharmaceutically acceptable excipient or a cosmetically acceptable excipient, may comprise an anti-adherent, binder, coat, disintegrant, filler, flavor, color, lubricant, glidant, sorbent, preservative, or sweetener. In some embodiments, the preparation may be substantially free of excipients.

[0394] In some embodiments, the preparation may be substantially free of one or more of the compounds or substances listed in the disclosure.

[0395] Exemplary compositions for aerosol administration include solutions in saline, which can contain, for example, benzyl alcohol or other suitable preservatives, absorption promoters to enhance bioavailability, and/or other solubilizing or dispersing agents. Conveniently in compositions for aerosol administration the ammonia oxidizing bacteria, e.g., *N. eutropha* is delivered in the form of an aerosol spray presentation from a pressurized pack or a nebulizer, with the use of a suitable propellant, e.g., dichlorodifluoro-methane, trichlorofluoromethane, dichlorotetrafluoroethane, carbon dioxide or other suitable gas. In the case of a pressurized aerosol the dosage unit can be determined by providing a valve to deliver a metered amount. Capsules and cartridges of e.g., gelatin can be formulated to contain a powder mix of the *N. eutropha* and a suitable powder base, for example lactose or starch. In certain embodiments, *N. eutropha* is administered as an aerosol from a metered dose valve, through an aerosol adapter also known as an actuator. Optionally, a stabilizer is also included, and/or porous particles for deep lung delivery are included (e.g., see U.S. Pat. No. 6,447,743).

[0396] Formulations may be presented with carriers such as cocoa butter, synthetic glyceride esters or polyethylene glycol. Such carriers are typically solid at ordinary tempera-

tures, but liquefy and/or dissolve at body temperature to release the ammonia oxidizing bacteria, e.g., *N. eutropha*.

[0397] Exemplary compositions for topical administration include a topical carrier such as Plastibase (mineral oil gelled with polyethylene). In some aspects, the composition and/or excipient may be in the form of one or more of a liquid, a solid, or a gel. For example, liquid suspensions may include, but are not limited to, water, saline, phosphate-buffered saline, or an ammonia oxidizing storage buffer. Gel formulations may include, but are not limited to agar, silica, polyacrylic acid (for example Carbopol®), carboxymethyl cellulose, starch, guar gum, alginate or chitosan. In some embodiments, the formulation may be supplemented with an ammonia source including, but not limited to ammonium chloride or ammonium sulfate.

[0398] In some embodiments, an ammonia oxidizing bacteria, e.g., *N. eutropha* composition is formulated to improve NO penetration into the skin. A gel-forming material such as KY jelly or various hair gels would present a diffusion barrier to NO loss to ambient air, and so improve the skin's absorption of NO. The NO level in the skin will generally not greatly exceed 20 nM/L because that level activates GC and would cause local vasodilatation and oxidative destruction of excess NO.

[0399] It should be understood that in addition to the ingredients particularly mentioned above, the formulations as described herein may include other agents conventional in the art having regard to the type of formulation in question.

[0400] The formulation, e.g., preparation, e.g., composition may be provided in a container, delivery system, or delivery device, having a weight, including or not including the contents of the container, that may be less than about 50, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1500, or 2000 grams.

[0401] Suitable unit dosage formulations are those containing an effective dose, as hereinbefore recited, or an appropriate fraction thereof, of ammonia oxidizing bacteria, e.g., *N. eutropha*.

[0402] A therapeutically effective amount of ammonia oxidizing bacteria, e.g., *N. eutropha* may be administered as a single pulse dose, as a bolus dose, or as pulse doses administered over time. Thus, in pulse doses, a bolus administration of ammonia oxidizing bacteria, e.g., *N. eutropha* is provided, followed by a time period wherein ammonia oxidizing bacteria, e.g., *N. eutropha* is administered to the subject, followed by a second bolus administration. In specific, non-limiting examples, pulse doses are administered during the course of a day, during the course of a week, or during the course of a month.

[0403] In some embodiments, a preparation of ammonia oxidizing bacteria, e.g., a formulation, e.g., a composition, may be applied for a pre-determined number of days. This may be based, for example, at least in part, on the severity of the condition or disease, the response to the treatment, the dosage applied and the frequency of the dose. For example, the preparation may be applied for about 1-3, 3-5, 5-7, 7-9, 5-10, 10-14, 12-18, 12-21, 21-28, 28-35, 35-42, 42-49, 49-56, 46-63, 63-70, 70-77, 77-84, 84-91 days, for about 1 month, for about 2 months, for about 3 months. In some embodiments, the ammonia oxidizing bacteria is administered for an indefinite period of time, e.g., greater than one year, greater than 5 years, greater than 10 years, greater than

15 years, greater than 30 years, greater than 50 years, greater than 75 years. In certain aspects, the preparation may be applied for about 16 days.

[0404] In some embodiments, a preparation of ammonia oxidizing bacteria, e.g., a formulation, e.g., a composition, may be applied a pre-determined number of times per day. This may be based, for example, at least in part, on the severity of the condition or disease, the response to the treatment, the dosage applied and the frequency of the dose. For example, the preparation may be applied 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24 times per day.

[0405] In some embodiments, the preparation may be applied one time per day. In other embodiments, the preparation may be applied two times per day. In some embodiments, the preparation may be applied a first pre-determined amount for a certain number of days, and a second pre-determined amount for a certain subsequent number of days. In some embodiments, the preparation may be applied for about 16 days.

[0406] Consumer Products

[0407] Ammonia oxidizing bacteria, e.g., *N. eutropha* may be associated with a variety of consumer products, and examples of such products are set out below. In some embodiments, the ammonia oxidizing bacteria, e.g., *N. eutropha* associated with a product is admixed with the product, for example, spread evenly throughout the product, and in some embodiments, ammonia oxidizing bacteria, e.g., the *N. eutropha* associated with a product is layered on the product.

[0408] In some embodiments, ammonia oxidizing bacteria, e.g., *N. eutropha* is associated with a powder. Powders are typically small particulate solids that are not attached to each other and that can flow freely when tilted. Exemplary powders for consumer use include talcum powder and some cosmetics (e.g., powder foundation).

[0409] In some embodiments, the ammonia oxidizing bacteria is associated with a cosmetic. The cosmetic may be a substance for topical application intended to alter a person's appearance, e.g., a liquid foundation, a powder foundation, blush, or lipstick. The cosmetic may be any substance recited in the Food and Drug Administration regulations, e.g., under 21 C.F.R. §720.4.

[0410] The cosmetic may be at least one of a baby product, e.g., a baby shampoo, a baby lotion, a baby oil, a baby powder, a baby cream; a bath preparation, e.g., a bath oil, a tablet, a salt, a bubble bath, a bath capsule; an eye makeup preparation, e.g., an eyebrow pencil, an eyeliner, an eye shadow, an eye lotion, an eye makeup remover, a mascara; a fragrance preparation, e.g., a cologne, a toilet water, a perfume, a powder (dusting and talcum), a sachet; hair preparations, e.g., hair conditioners, hair sprays, hair straighteners, permanent waves, rinses, shampoos, tonics, dressings, hair grooming aids, wave sets; hair coloring preparations, e.g., hair dyes and colors, hair tints, coloring hair rinses, coloring hair shampoos, hair lighteners with color, hair bleaches; makeup preparations, e.g., face powders, foundations, leg and body paints, lipstick, makeup bases, rouges, makeup fixatives; manicuring preparations, e.g., basecoats and undercoats, cuticle softeners, nail creams and lotions, nail extenders, nail polish and enamel, nail polish and enamel removers; oral hygiene products, e.g., dentrifices, mouthwashes and breath fresheners; bath soaps and detergents, deodorants, douches, feminine hygiene

deodorants; shaving preparations, e.g., aftershave lotions, beard softeners, talcum, preshave lotions, shaving cream, shaving soap; skin care preparations, e.g., cleansing, depilatories, face and neck, body and hand, foot powders and sprays, moisturizing, night preparations, paste masks, skin fresheners; and suntan preparations, e.g., gels, creams, and liquids, and indoor tanning preparations.

[0411] In some embodiments, the formulations, compositions, or preparations described herein, may comprise, be provided as, or disposed in at least one of a baby product, e.g., a baby shampoo, a baby lotion, a baby oil, a baby powder, a baby cream; a bath preparation, e.g., a bath oil, a tablet, a salt, a bubble bath, a bath capsule; a powder (dusting and talcum), a sachet; hair preparations, e.g., hair conditioners, rinses, shampoos, tonics, face powders, cuticle softeners, nail creams and lotions, oral hygiene products, mouthwashes, bath soaps, douches, feminine hygiene deodorants; shaving preparations, e.g., aftershave lotions, skin care preparations, e.g., cleansing, face and neck, body and hand, foot powders and sprays, moisturizing, night preparations, paste masks, skin fresheners; and suntan preparations, e.g., gels, creams, and liquids.

In some embodiments, ammonia oxidizing bacteria, e.g., *N. eutropha* is associated with a cosmetic. The cosmetic may be a substance for topical application intended to alter a person's appearance, e.g., a liquid foundation, a powder foundation, blush, or lipstick. Other components may be added to these cosmetic preparations as selected by one skilled in the art of cosmetic formulation such as, for example, water, mineral oil, coloring agent, perfume, aloe, glycerin, sodium chloride, sodium bicarbonate, pH buffers, UV blocking agents, silicone oil, natural oils, vitamin E, herbal concentrates, lactic acid, citric acid, talc, clay, calcium carbonate, magnesium carbonate, zinc oxide, starch, urea, and erythorbic acid, or any other excipient known by one of skill in the art, including those disclosed herein.

[0412] In some embodiments, the preparation may be disposed in, or provided as, a powder, cosmetic, cream, stick, aerosol, salve, wipe, or bandage.

[0413] In some embodiments, ammonia oxidizing bacteria, e.g., *N. eutropha* is associated with a cream. The cream may be a fluid comprising a thickening agent, and generally has a consistency that allows it to be spread evenly on the skin. Exemplary creams include moisturizing lotion, face cream, and body lotion.

[0414] In some embodiments, ammonia oxidizing bacteria, e.g., the *N. eutropha* is associated with a stick. A stick is typically a solid that, when placed in contact with a surface, transfers some of the stick contents to the surface. Exemplary sticks include deodorant stick, lipstick, lip balm in stick form, and sunscreen applicator sticks.

[0415] In some embodiments, ammonia oxidizing bacteria, e.g., the *N. eutropha* is associated with an aerosol. An aerosol is typically a colloid of fine solid particles or fine liquid droplets, in a gas such as air. Aerosols may be created by placing the *N. eutropha* (and optionally carriers) in a vessel under pressure, and then opening a valve to release the contents. The container may be designed to only exert levels of pressure that are compatible with *N. eutropha* viability. For instance, the high pressure may be exerted for only a short time, and/or the pressure may be low enough not to impair viability. Examples of consumer uses of aerosols include for sunscreen, deodorant, perfume, hairspray, and insect repellent.

[0416] In some embodiments, ammonia oxidizing bacteria, e.g., the *N. eutropha* is associated with a salve. A salve may be a topically applied agent with a liquid or cream-like consistency, intended to protect the skin or promote healing. Examples of salves include burn ointments and skin moisturizers.

[0417] In some embodiments, ammonia oxidizing bacteria, e.g., the *N. eutropha* is associated with a wipe. A wipe may be a flexible material suitable for topically applying a liquid or cream onto skin. The wipe may be, e.g., paper-based or cloth based. Exemplary wipes include tissues and wet wipes.

[0418] The compositions comprising ammonia oxidizing bacteria, e.g., *N. eutropha* may also comprise one or more of a moisturizing agent, deodorizing agent, scent, colorant, insect repellent, cleansing agent, or UV-blocking agent.

[0419] For instance, the moisturizing agent may be an agent that reduces or prevents skin dryness. Exemplary moisturizing agents include humectants (e.g., urea, glycerin, alpha hydroxy acids and dimethicone) and emollients (e.g., lanolin, mineral oil and petrolatum). Moisturizing agents may be included, e.g., in ammonia oxidizing bacteria, e.g., *N. eutropha*-containing creams, balms, lotions, or sunscreen.

[0420] A deodorizing agent may be an agent that reduces unwanted odors. A deodorizing agent may work by directly neutralizing odors, preventing perspiration, or preventing the growth of odor-producing bacteria. Exemplary deodorizing agents include aluminum salts (e.g., aluminum chloride or aluminum chlorohydrate), cyclomethicone, talc, baking soda, essential oils, mineral salts, hops, and witch hazel. Deodorizing agents are typically present in spray or stick deodorants, and can also be found in some soaps and clothing.

[0421] An insect repellent may be an agent that can be applied to surfaces (e.g., skin) that discourage insects and other arthropods from lighting on the surface. Insect repellents include DEET (N,N-diethyl-m-toluamide), p-menthane-3,8-diol (PMD), icaridin, nepetalactone, citronella oil, neem oil, bog myrtle, dimethyl carbate, Tricyclodecenylyl allyl ether, and IR3535 (3-[N-Butyl-N-acetyl]-aminopropionic acid, ethyl ester).

[0422] A cleansing agent may be an agent that removes dirt or unwanted bacteria from a surface like skin. Exemplary cleansing agents include bar soaps, liquid soaps, and shampoos.

[0423] A UV-blocking agent may be an agent that can be applied to a surface to reduce the amount of ultraviolet light the surface receives. A UV-blocking agent may block UV-A and/or UV-B rays. A UV blocking agent can function by absorbing, reflecting, or scattering UV. Exemplary UV-blocking agents include absorbers, e.g., homosalate, octisalate (also called octyl salicylate), octinoxate (also called octyl methoxycinnamate or OMC), octocrylene, oxybenzone, and avobenzone, and reflectors (e.g., titanium dioxide and zinc oxide). UV-blocking agents are typically present in sunscreens, and can also be found in skin creams and some cosmetics.

[0424] In some embodiments, ammonia oxidizing bacteria, e.g., *N. eutropha* is associated with a conditioner. Conditioner generally refers to a substance with cream-like consistency that can be applied to hair to improve its appearance, strength, or manageability.

[0425] In some embodiments, ammonia oxidizing bacteria, e.g., *N. eutropha* is associated with cloth. Cloth gener-

ally refers to a flexible material suitable to be made into clothing, e.g., having enough material strength to withstand everyday motion by a wearer. Cloth can be fibrous, woven, or knit; it can be made of a naturally occurring material or a synthetic material. Exemplary cloth materials include cotton, flax, wool, ramie, silk, denim, leather, nylon, polyester, and spandex, and blends thereof.

[0426] In some embodiments, ammonia oxidizing bacteria, e.g., *N. eutropha* is associated with yarn. Yarn generally refers to a long, thin spun flexible material that is suitable for knitting or weaving. Yarn can be made of, e.g., wool, cotton, polyester, and blends thereof.

[0427] In some embodiments, ammonia oxidizing bacteria, e.g., *N. eutropha* is associated with thread. Thread generally refers to a long, thin spun flexible material that is suitable for sewing. Thread generally has a thinner diameter than yarn. Thread can be made of, e.g., cotton, polyester, nylon, silk, and blends thereof.

[0428] Articles of clothing such as, for example, shoes, shoe inserts, pajamas, sneakers, belts, hats, shirts, underwear, athletic garments, helmets, towels, gloves, socks, bandages, and the like, may also be treated with ammonia oxidizing bacteria, e.g., *N. eutropha*. Bedding, including sheets, pillows, pillow cases, and blankets may also be treated with ammonia oxidizing bacteria, e.g., *N. eutropha*. In some embodiments, areas of skin that cannot be washed for a period of time may also be contacted with ammonia oxidizing bacteria, e.g., *N. eutropha*. For example, skin enclosed in orthopedic casts which immobilize injured limbs during the healing process, and areas in proximity to injuries that must be kept dry for proper healing such as stitched wounds may benefit from contact with the ammonia oxidizing bacteria, e.g., *N. eutropha*.

[0429] In some aspects, the present disclosure provides a wearable article comprising an *N. eutropha* strain as described herein. A wearable article may be a light article that can be closely associated with a user's body, in a way that does not impede ambulation. Examples of wearable articles include a wristwatch, wristband, headband, hair elastic, hair nets, shower caps, hats, hairpieces, and jewelry. The wearable article comprising an ammonia oxidizing bacteria, e.g., *N. eutropha* strain described herein may provide, e.g., at a concentration that provides one or more of a treatment or prevention of a skin disorder, a treatment or prevention of a disease or condition associated with low nitrite levels, a treatment or prevention of body odor, a treatment to supply nitric oxide to a subject, or a treatment to inhibit microbial growth.

[0430] In some embodiments, the ammonia oxidizing bacteria, e.g., *N. eutropha* is associated with a product intended to contact the hair, for example, a brush, comb, shampoo, conditioner, headband, hair elastic, hair nets, shower caps, hats, and hairpieces. Nitric oxide formed on the hair, away from the skin surface, may be captured in a hat, scarf or face mask and directed into inhaled air.

[0431] Articles contacting the surface of a human subject, such as a diaper, may be associated with ammonia oxidizing bacteria, e.g., *N. eutropha*. Because diapers are designed to hold and contain urine and feces produced by incontinent individuals, the urea in urine and feces can be hydrolyzed by skin and fecal bacteria to form free ammonia which is irritating and may cause diaper rash. Incorporation of bacteria that metabolize urea into nitrite or nitrate, such as ammonia oxidizing bacteria, e.g., *N. eutropha*, may avoid

the release of free ammonia and may release nitrite and ultimately NO which may aid in the maintenance of healthy skin for both children and incontinent adults. The release of nitric oxide in diapers may also have anti-microbial effects on disease causing organisms present in human feces. This effect may continue even after disposable diapers are disposed of as waste and may reduce the incidence of transmission of disease through contact with soiled disposable diapers

[0432] In some embodiments, the product comprising ammonia oxidizing bacteria, e.g., *N. eutropha* is packaged. The packaging may serve to compact the product or protect it from damage, dirt, or degradation. The packaging may comprise, e.g., plastic, paper, cardboard, or wood. In some embodiments the packaging is impermeable to bacteria. In some embodiments the packaging is permeable to oxygen and/or carbon dioxide.

5. METHODS OF TREATMENT WITH *N. EUTROPHA*

[0433] The present disclosure provides various methods of treating diseases and conditions using ammonia oxidizing bacteria, e.g., *N. eutropha*. The ammonia oxidizing bacteria, e.g., *N. eutropha* that may be used to treat diseases and conditions include all the ammonia oxidizing bacteria, e.g., *N. eutropha* compositions described in this application, e.g., a purified preparation of optimized ammonia oxidizing bacteria, e.g., *N. eutropha*, e.g. those in Section 2 above, for instance strain D23.

[0434] For instance, the disclosure provides uses, for treating a condition or disease (e.g., inhibiting microbial growth on a subject's skin), an optionally axenic composition of *N. eutropha* comprising a nucleic acid sequence at least about 70%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 99.5% identical to SEQ ID NO: 1; an optionally axenic composition of *N. eutropha* comprising a nucleic acid sequence at least about 70%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 99.5% identical to any one of the strain D23 nucleic acids of Table 1. In embodiments, the *N. eutropha* comprises at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, or all of the strain D23 nucleic acids of Table 1. In embodiments, the *N. eutropha* comprises one or more nucleic acids of FIGS. 6-8. As a further example, this disclosure provides uses, for treating a condition or disease, an optionally axenic composition of *N. eutropha* comprising an amino acid sequence at least about 70%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 99.5% identical to any one of the strain D23 protein sequences of Table 1. In embodiments, the *N. eutropha* comprises at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, or all of the strain D23 protein sequences of Table 1. In embodiments, the *N. eutropha* comprises one or more proteins encoded by the nucleic acids of FIGS. 6-8. The *N. eutropha* of this paragraph may be used to treat, e.g., diabetic ulcers, e.g., diabetic foot ulcers, chronic wounds, acne, rosacea, eczema, or psoriasis.

[0435] In certain embodiments, the disclosure provides uses, for treating a condition or disease (e.g., inhibiting microbial growth on a subject's skin), an optionally axenic composition of *N. eutropha* having one or more of: (1) an optimized growth rate, (2) an optimized NH_4^+ oxidation rate, (3) an optimized resistance to NH_3 , (4) an optimized resistance to, NH_4^+ , and (5) an optimized resistance to, NO_2^- . For instance, the axenic *N. eutropha* composition may have properties (1) and (2); (2) and (3); (3) and (4); or

(4) and (5) from the list at the beginning of this paragraph. As another example, the axenic *N. eutropha* composition may have properties (1), (2), and (3); (1), (2), and (4); (1), (2), and (5); (1), (3), and (4); (1), (3), and (5); (1), (4), and (5); (2), (3), and (4); (2), (3), and (5), or (3), (4), and (5) from the list at the beginning of this paragraph. As a further example, the optionally axenic *N. eutropha* composition may have properties (1), (2), (3), and (4); (1), (2), (3), and (5); (1), (2), (4), and (5); (1), (3), (4), and (5); or (2), (3), (4), and (5) from the list at the beginning of this paragraph. In some embodiments, the axenic *N. eutropha* composition has properties (1), (2), (3), (4), and (5) from the list at the beginning of this paragraph. The *N. eutropha* of this paragraph may be used to treat, e.g., diabetic ulcers, e.g., diabetic foot ulcers, chronic wounds, acne, rosacea, eczema, or psoriasis.

[0436] In some embodiments, optionally axenic *N. eutropha* (e.g., strain D23) are used to treat a subject. Subjects may include an animal, a mammal, a human, a non-human animal, a livestock animal, or a companion animal.

[0437] In some embodiments, optionally axenic *N. eutropha* described herein (e.g., the *N. eutropha* described in this Section and in Section 2 above, e.g., strain D23) are used to inhibit the growth of other organisms. For instance, *N. eutropha* D23 is well-adapted for long-term colonization of human skin, and in some embodiments it out-competes other bacteria that are undesirable on the skin. Undesirable skin bacteria include, e.g., those that can infect wounds, raise the risk or severity of a disease, or produce odors. Certain undesirable skin bacteria include *S. aureus*, *P. aeruginosa*, *S. pyogenes*, and *A. baumannii*. The *N. eutropha* described herein may out-compete other organisms by, e.g., consuming scarce nutrients, or generating byproducts that are harmful to other organisms, e.g., changing the pH of the skin to a level that is not conducive to the undesirable organism's growth.

[0438] Accordingly, the present disclosure provides, inter alia, a method of inhibiting microbial growth on a subject's skin, comprising topically administering to a human in need thereof an effective dose of optionally axenic *N. eutropha* bacteria as described herein (e.g., strain D23). Similarly, the present disclosure provides optionally axenic *N. eutropha* as described herein (e.g., strain D23) for use in inhibiting microbial growth on a subject's skin. Likewise, the present disclosure provides a use of optionally axenic *N. eutropha* (e.g., strain D23) in the manufacture of a medicament for inhibiting microbial growth on a subject's skin.

[0439] The present disclosure also provides a method of supplying nitric oxide to a subject, comprising positioning an effective dose of optionally axenic *N. eutropha* bacteria described herein (e.g., strain D23) in close proximity to the subject. Similarly, the present disclosure provides optionally axenic *N. eutropha* (e.g., strain D23) as described herein for use in supplying nitric oxide to a subject. Likewise, the present disclosure provides a use of optionally axenic *N. eutropha* (e.g., strain D23) in the manufacture of a medicament or composition suitable for position in close proximity to a subject.

[0440] The present disclosure also provides a method of reducing body odor, comprising topically administering to a subject in need thereof an effective dose of optionally axenic *N. eutropha* bacteria described herein (e.g., strain D23). Similarly, the present disclosure provides optionally axenic *N. eutropha* as described herein (e.g., strain D23) for use in

reducing body odor in a subject. Likewise, the present disclosure provides a use of optionally axenic *N. eutropha* as described herein (e.g., strain D23) in the manufacture of a medicament or composition for reducing body odor.

[0441] The present disclosure also provides a method of treating or preventing a disease associated with low nitrite levels, comprising topically administering to a subject in need thereof a therapeutically effective dose of optionally axenic *N. eutropha* bacteria described herein (e.g., strain D23). Similarly, the present disclosure provides a topical formulation of optionally axenic *N. eutropha* as described herein (e.g., strain D23) for use in treating a disease associated with low nitrite levels. Likewise, the present disclosure provides a use of optionally axenic *N. eutropha* as described herein (e.g., strain D23) in the manufacture of a topical medicament for treating a disease associated with low nitrite levels.

[0442] The present disclosure also provides a method of treating or preventing a skin disorder or skin infection, comprising topically administering to a subject in need thereof a therapeutically effective dose of optionally axenic *N. eutropha* bacteria as described herein (e.g., strain D23). Similarly, the present disclosure provides optionally axenic *N. eutropha* as described herein (e.g., strain D23) for use in treating a skin disorder in a subject. Likewise, the present disclosure provides a use of optionally axenic *N. eutropha* as described herein (e.g., strain D23) in the manufacture of a medicament for treating skin disorder. In embodiments, the skin disorder is acne, rosacea, eczema, psoriasis, or urticaria; the skin infection is impetigo.

[0443] While not wishing to be bound by theory, it is proposed that treatment of acne with a therapeutically effective dose of optionally axenic *N. eutropha* bacteria as described herein (e.g., strain D23) may involve the down-regulation of inflammation due to NO generation; and/or limiting and/or inhibiting the spread and proliferation of *Propionibacterium acnes* associated with acne vulgaris through acidified nitrite and NO production.

[0444] For instance, the disclosure provides uses, for treating a condition or disease (e.g., inhibiting microbial growth on a subject's skin), a composition of ammonia oxidizing bacteria. In embodiments, the ammonia oxidizing bacteria may be used to treat, e.g., chronic wounds, acne, rosacea, eczema, psoriasis, urticaria, skin infections, or diabetic ulcers, e.g., diabetic foot ulcers.

[0445] The systems and methods of the present disclosure may provide for, or contain contents, to be useful for treating or preventing a skin disorder, treating or preventing a disease or condition associated with low nitrite levels, a treating or preventing body odor, treating to supply nitric oxide to a subject, or treating to inhibit microbial growth.

[0446] The systems and methods of the present disclosure may provide for reducing an amount of undesirable bacteria from an environment, e.g., a surface of a subject.

[0447] The systems and methods of the present disclosure may provide for, or contain contents, to be useful in a treatment of at least one of HIV dermatitis, infection in a diabetic foot ulcer, atopic dermatitis, acne, eczema, contact dermatitis, allergic reaction, psoriasis, urticaria, rosacea, skin infections, vascular disease, vaginal yeast infection, a sexually transmitted disease, heart disease, atherosclerosis, baldness, leg ulcers secondary to diabetes or confinement to bed, angina, particularly chronic, stable angina pectoris, ischemic diseases, congestive heart failure, myocardial infarction,

ischemia reperfusion injury, laminitis, hypertension, hypertrophic organ degeneration, Raynaud's phenomenon, fibrosis, fibrotic organ degeneration, allergies, autoimmune sensitization, end stage renal disease, obesity, impotence, pneumonia, primary immunodeficiency, epidermal lysis bulosa, or cancer.

[0448] The systems and methods of the present disclosure may provide for, or contain contents, to be useful in a treatment of at least one of acne, eczema, psoriasis, urticaria, rosacea, skin infections and wounds, e.g., an infected wound.

[0449] In some embodiments, ammonia oxidizing bacteria may be used to treat a subject. Subjects may include an animal, a mammal, a human, a non-human animal, a livestock animal, or a companion animal.

[0450] In some embodiments, ammonia oxidizing bacteria described herein are used to inhibit the growth of other organisms. For instance, ammonia oxidizing bacteria may be well-adapted for long-term colonization of human skin, and in some embodiments it out-competes other bacteria that are undesirable on the skin. Undesirable skin bacteria include, e.g., those that can infect wounds, raise the risk or severity of a disease, or produce odors. Undesirable bacteria may be referred to as pathogenic bacteria. Certain undesirable skin bacteria include *Staphylococcus aureus* (*S. aureus*), e.g., methicillin resistant *Staphylococcus aureus* *Pseudomonas aeruginosa* (*P. aeruginosa*), *Streptococcus pyogenes* (*S. pyogenes*), *Acinetobacter baumannii* (*A. baumannii*), *Propionibacteria*, and *Stenotrophomonas*. The ammonia oxidizing bacteria described herein may out-compete other organisms by, e.g., consuming scarce nutrients, or generating byproducts that are harmful to other organisms, e.g., changing the pH of the skin to a level that is not conducive to the undesirable organism's growth.

[0451] Accordingly, the present disclosure provides, inter alia, a method of inhibiting microbial growth on a subject's skin, comprising topically administering to a human in need thereof an effective dose of ammonia oxidizing bacteria as described herein. Similarly, the present disclosure provides ammonia oxidizing bacteria as described herein for use in inhibiting microbial growth on a subject's skin. Likewise, the present disclosure provides a use of ammonia oxidizing bacteria in the manufacture of a medicament for inhibiting microbial growth on a subject's skin.

[0452] The present disclosure provides, inter alia, a method of changing a composition of a skin microbiome, e.g., modulating a composition of a skin microbiome, e.g., modulating or changing the proportions of the skin microbiome, in an environment, e.g., a surface, e.g., a surface of a subject. The method may comprise administering, e.g., applying, a preparation comprising ammonia oxidizing bacteria to an environment, e.g., a surface, e.g., a surface of a subject. In some embodiments, the amount and frequency of administration, e.g., application, may be sufficient to reduce the proportion of pathogenic bacteria on the surface of the skin. In some embodiments, the subject may be selected on the basis of the subject being in need of a reduction in the proportion of pathogenic bacteria on the surface of the skin.

[0453] The present disclosure may further provide obtaining a sample from the surface of the skin, and isolating DNA of bacteria in the sample. Sequencing of the DNA of bacteria in the sample may also be performed to determine or monitor the amount or proportion of bacteria in a sample of a subject.

[0454] The present disclosure may also provide for increasing the proportion of non-pathogenic bacteria on the surface. In some embodiments, the non-pathogenic bacteria may be commensal non-pathogenic bacteria. In some embodiments, the non-pathogenic bacteria may be of the *Staphylococcus* genus. In some embodiments, the non-pathogenic bacteria may be *Staphylococcus epidermidis*. In some embodiments, the non-pathogenic bacteria that is increased in proportion may be of the *Staphylococcus* genus, comprising at least about 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% *Staphylococcus epidermidis*.

[0455] The increase in the proportion of non-pathogenic bacteria may occur with a pre-determined period of time, e.g., in less than 1 day, 2 days, 3 days, 4 days, 5 days, 1 week, 2 weeks, 3 weeks, or 4 weeks, or in less than 1-3, 3-5, 5-7, 7-9, 5-10, 10-14, 12-18, 12-21, 21-28, 28-35, 35-42, 42-49, 49-56, 46-63, 63-70, 70-77, 77-84, 84-91 days.

[0456] The increase in the proportion of *Staphylococcus* bacteria, e.g., *Staphylococcus epidermidis*, may be observed in less than about 3 weeks, e.g., about 16 days, e.g., about 2 weeks.

[0457] The present disclosure may provide for decreasing the proportion of pathogenic bacteria, e.g., potentially pathogenic bacteria, e.g., disease-associated bacteria on the surface. In some embodiments, the pathogenic bacteria may be Propionibacteria. In some embodiments, the pathogenic bacteria may be *Stenotrophomonas*.

[0458] The decrease in the proportion of pathogenic bacteria may occur with a pre-determined period of time, e.g., in less than 1 day, 2 days, 3 days, 4 days, 5 days, 1 week, 2 weeks, 3 weeks, or 4 weeks, or in less than 1-3, 3-5, 5-7, 7-9, 5-10, 10-14, 12-18, 12-21, 21-28, 28-35, 35-42, 42-49, 49-56, 46-63, 63-70, 70-77, 77-84, 84-91 days.

[0459] The decrease in the proportion of Propionibacteria bacteria and/or *Stenotrophomonas* may be observed in less than about 3 weeks, e.g., about 16 days, e.g., about 2 weeks.

[0460] The present disclosure also provides a method of supplying nitric oxide to a subject, comprising positioning an effective dose of ammonia oxidizing bacteria described herein in close proximity to the subject. Similarly, the present disclosure provides ammonia oxidizing bacteria as described herein for use in supplying nitric oxide to a subject. Likewise, the present disclosure provides a use of in the manufacture of a medicament or composition suitable for position in close proximity to a subject.

[0461] The present disclosure also provides a method of reducing body odor, comprising topically administering to a subject in need thereof an effective dose of ammonia oxidizing bacteria described herein. Similarly, the present disclosure provides ammonia oxidizing bacteria as described herein for use in reducing body odor in a subject. Likewise, the present disclosure provides a use of ammonia oxidizing bacteria as described herein in the manufacture of a medicament or composition for reducing body odor.

[0462] The present disclosure also provides a method of treating or preventing a disease associated with low nitrite levels, comprising topically administering to a subject in need thereof a therapeutically effective dose of ammonia oxidizing bacteria described herein. Similarly, the present disclosure provides a topical formulation of ammonia oxidizing bacteria as described herein for use in treating a disease associated with low nitrite levels. Likewise, the

present disclosure provides a use of ammonia oxidizing bacteria as described herein in the manufacture of a topical medicament for treating a disease associated with low nitrite levels.

[0463] The present disclosure also provides a method of treating or preventing a skin disorder or skin infection, comprising topically administering to a subject in need thereof a therapeutically effective dose of ammonia oxidizing bacteria as described herein. Similarly, the present disclosure provides ammonia oxidizing bacteria as described herein for use in treating a skin disorder in a subject. Likewise, the present disclosure provides a use of ammonia oxidizing bacteria as described herein in the manufacture of a medicament for treating skin disorder. In embodiments, the skin disorder is acne, rosacea, eczema, psoriasis, or urticaria; the skin infection is impetigo.

[0464] While not wishing to be bound by theory, it is proposed that treatment of rosacea with a therapeutically effective dose of optionally axenic *N. eutropha* bacteria as described herein (e.g., strain D23) may involve downregulation due to NO generation. This may be due to expression of Kazal-type KLK5/KLK7 inhibitor(s) that may reduce formation of the human cathelicidin peptide LL-37 from its precursor propeptide hCAP18.

[0465] While not wishing to be bound by theory, it is proposed that treatment of eczema and/or atopic dermatitis with a therapeutically effective dose of optionally axenic *N. eutropha* bacteria as described herein (e.g., strain D23) may involve downregulation of inflammation due to NO generation; and/or limiting and/or inhibiting the spread and proliferation of *S. aureus* and other skin pathogens often associated with very high colonization rates and skin loads in atopic dermatitis through acidified nitrite and NO production.

[0466] While not wishing to be bound by theory, it is proposed that treatment of psoriasis with a therapeutically effective dose of optionally axenic *N. eutropha* bacteria as described herein (e.g., strain D23) may involve downregulation of inflammation due to NO generation and reduction in formation of human cathelicidin peptide LL-37.

[0467] While not wishing to be bound by theory, it is proposed that treatment of psoriasis with a therapeutically effective dose of optionally axenic *N. eutropha* bacteria as described herein (e.g., strain D23) may involve downregulation of inflammation due to NO generation.

[0468] While not wishing to be bound by theory, it is proposed that treatment of impetigo or other skin and soft tissue infections with a therapeutically effective dose of optionally axenic *N. eutropha* bacteria as described herein (e.g., strain D23) may involve limiting and/or inhibiting the spread and proliferation of *S. aureus* and *S. pyogenes*.

[0469] The present disclosure also provides a method of promoting wound healing, comprising administering to a wound an effective dose of optionally axenic *N. eutropha* bacteria as described herein (e.g., strain D23). Similarly, the present disclosure provides optionally axenic *N. eutropha* as described herein (e.g., strain D23) for use in treating a wound. Likewise, the present disclosure provides a use of optionally axenic *N. eutropha* as described herein (e.g., strain D23) in the manufacture of a medicament or a composition for treating a wound.

[0470] Optionally axenic *N. eutropha* as described herein (e.g., strain D23) may be used to promote wound healing in a patient that has an impaired healing ability, e.g., a diabetic patient.

[0471] In some embodiments, this disclosure provides methods of using optionally axenic *N. eutropha* as described herein (e.g., strain D23) to prevent a disease or disorder, e.g., a skin disorder. Prevention, in certain embodiments, means reducing the risk of a subject developing a disease, compared to a similar untreated subject. The risk need not be reduced to zero.

[0472] Individuals having a reduced bathing frequency, such as astronauts, submarine crew members, military personnel during a campaign, civilian workers in remote locations, refugees, bedridden individuals and many others may maintain healthier skin by maintaining *N. eutropha* on the skin. With regard to bedridden individuals, the *N. eutropha* in some embodiments reduces the frequency or severity of bed sores by augmenting inadequate circulation.

[0473] It is appreciated that many modern degenerative diseases may be caused by a lack of NO species, and that AOB on the external skin can supply those species by diffusion, and that application of AOB to the skin resolves long standing medical conditions. In certain embodiments, AOB are applied to a subject to offset modern bathing practices, especially with anionic detergents remove AOB from the external skin.

[0474] One suitable method of topical application to apply sufficient *N. eutropha* and then wear sufficient clothing so as to induce sweating. However, many people will want to derive the benefits of AOB while maintaining their current bathing habits, in which case, a culture of the bacteria can be applied along with sufficient substrate for them to produce NO. A nutrient solution approximating the inorganic composition of human sweat can be used for this purpose. Using bacteria adapted to media approximating human sweat minimizes the time for them to adapt when applied. Since sweat evaporates once excreted onto the skin surface, using a culture media that has a higher ionic strength is desirable. A concentration approximately twice that of human sweat is suitable, but other conditions are also contemplated. AOB's nutritional needs are typically met with NH_3 or urea, O_2 , CO_2 , and minerals. In some embodiments, the substrate comprises trace minerals including iron, copper, zinc, cobalt, molybdenum, manganese, sodium, potassium, calcium, magnesium, chloride, phosphate, sulfate, or any combination thereof.

[0475] In some embodiments, the present disclosure provides a method of treating a wound by applying a bandage comprising *N. eutropha* to the wound. Also provided are methods of producing such a bandage. The bandage may comprise, for example, an adhesive portion to affix the bandage to undamaged skin near the wound and a soft, flexible portion to cover or overlay the wound. In some embodiments, the bandage contains no other organisms but *N. eutropha*. The bandage may be made of a permeable material that allows gasses like oxygen and carbon dioxide to reach the *N. eutropha* when the bandage is applied to the wound. In certain embodiments, the bandage comprises nutrients for *N. eutropha* such as ammonium, ammonia, urea, or trace minerals. In certain embodiments, the bandage comprises an antibiotic to which the *N. eutropha* is resistant. The antibiotic resistance may arise from one or more endogenous resistance gene or from one or more transgenic.

[0476] In some embodiments, the *N. eutropha* is administered at a dose of about 10^8 - 10^9 CFU, 10^9 - 10^{10} CFU, 10^{10} - 10^{11} CFU, or 10^{11} - 10^{12} CFU per application. In some embodiments, the *N. eutropha* is administered topically at a dose of about 10^{10} - 10^{11} CFU, e.g., about 1×10^{10} - 5×10^{10} , 1×10^{10} - 3×10^{10} , or 1×10^{10} - 2×10^{10} CFU.

[0477] In some embodiments, the *N. eutropha* is administered in a volume of about 1-2, 2-5, 5-10, 10-15, 12-18, 15-20, 20-25, or 25-50 ml per dose. In some embodiments, the solution is at a concentration of about 10^8 - 10^9 , 10^9 - 10^{10} , or 10^{10} - 10^{11} CFUs/ml. In some embodiments, the *N. eutropha* is administered as two 15 ml doses per day, where each dose is at a concentration of 10^9 CFU/ml.

[0478] In some embodiments, ammonia oxidizing bacteria, e.g., the *N. eutropha* is administered once, twice, three, or four times per day. In some embodiments, ammonia oxidizing bacteria, e.g., the *N. eutropha* is administered once, twice, three, four, five, or six times per week. In some embodiments, ammonia oxidizing bacteria, e.g., the *N. eutropha* is administered shortly after bathing. In some embodiments, ammonia oxidizing bacteria, e.g., the *N. eutropha* is administered shortly before sleep.

[0479] In certain aspects, the present disclosure provides combination therapies comprising ammonia oxidizing bacteria, e.g., a *N. eutropha* and a second therapeutic. For instance, the disclosure provides physical admixtures of the two (or more) therapies are physically admixed. In other embodiments, the two (or more) therapies are administered in combination as separate formulation. The second therapy may be, e.g., a pharmaceutical agent, surgery, or any other medical approach that treats the relevant disease or disorder. The following paragraphs describe combination therapies capable of treating diabetic ulcers, chronic wounds, acne, rosacea, eczema, and psoriasis.

[0480] For instance, in a combination therapy capable of treating diabetic ulcers, the second therapy may comprise, e.g., a wound dressing (e.g., absorptive fillers, hydrogel dressings, or hydrocolloids), angiotensin, angiotensin analogues, platelet-rich fibrin therapy, hyperbaric oxygen therapy, negative pressure wound therapy, debridement, drainage, arterial revascularization, hyperbaric oxygen therapy, low level laser therapy, and gastrocnemius recession. The combination therapy may comprise one or more of the above-mentioned treatments.

[0481] In a combination therapy capable of treating chronic wounds, the second therapy may comprise, e.g., an antibiotic (e.g., topical or systemic, and bacteriocidal or bacteriostatic) such as Penicillins, cephalosporins, polymyxins, rifamycins, lipiarmycins, quinolones, sulfonamides, macrolides, lincosamides, tetracyclines, cyclic lipopeptides, glycyliclones, oxazolidinones, and lipiarmycins; angiotensin, angiotensin analogues; debridement; drainage; wound irrigation; negative pressure wound therapy; application of heat; arterial revascularization; hyperbaric oxygen therapy; antioxidants such as ascorbic acid, glutathione, lipoic acid, carotenes, α -tocopherol, or ubiquinol; low level laser therapy; gastrocnemius recession; growth factors such as vascular endothelial growth factor, insulin-like growth factor 1-2, platelet derived growth factor, transforming growth factor- β , or epidermal growth factor; application of autologous platelets such as those that secrete one or more growth factors such as vascular endothelial growth factor, insulin-like growth factor 1-2, platelet derived growth factor, transforming growth factor- β , or epidermal growth

factor; implantation of cultured keratinocytes; allograft; collagen, for instance a dressing comprising collagen; or protease inhibitors such as SLPI. The combination therapy may comprise one or more of the above-mentioned treatments.

[0482] In a combination therapy capable of treating acne, the second therapy may comprise, e.g., a medication (e.g., systemic or topical) such as Benzoyl peroxide, antibiotics (such as erythromycin, clindamycin, or a tetracycline), Salicylic acid, hormones (e.g., comprising a progestin such as desogestrel, norgestimate or drospirenone), retinoids such as tretinoin, adapalene, tazarotene, or isotretinoin. The second therapy may also be a procedure such as comedo extraction, corticosteroid injection, or surgical lancing. The combination therapy may comprise one or more of the above-mentioned treatments.

[0483] In a combination therapy capable of treating rosacea, the second therapy may comprise, e.g., an antibiotic, e.g., an oral tetracycline antibiotic such as tetracycline, doxycycline, or minocycline, or a topical antibiotic such as metronidazole; azelaic acid; alpha-hydroxy acid; isotretinoin can be prescribed; sandalwood oil; clonidine; beta-blockers such as nadolol and propranolol; antihistamines (such as loratadine); mirtazapine; methylsulfonylmethane or silymarin, optionally in combination with each other; lasers such as dermatological vascular laser or CO₂ laser; or light therapies such as intense pulsed light, low-level light therapy or photorejuvenation. The combination therapy may comprise one or more of the above-mentioned treatments.

[0484] In a combination therapy capable of treating eczema, the second therapy may comprise, e.g., a corticosteroid such as hydrocortisone or clobetasol propionate, immunosuppressants (topical or systemic) such as pimecrolimus, tacrolimus, ciclosporin, azathioprine or methotrexate, or light therapy such as with ultraviolet light. The combination therapy may comprise one or more of the above-mentioned treatments.

[0485] In a combination therapy capable of treating psoriasis, the second therapy may comprise, e.g., a corticosteroid such as desoximetasone; a retinoid; coal tar; Vitamin D or an analogue thereof such as paricalcitol or calcipotriol; moisturizers and emollients such as mineral oil, vaseline, calcipotriol, decubal, or coconut oil; dithranol; or fluocinonide. The combination therapy may comprise one or more of the above-mentioned treatments.

[0486] While not wishing to be bound by theory, it is proposed that treatment of psoriasis with a therapeutically effective dose of ammonia oxidizing bacteria described herein may involve downregulation of inflammation due to NO generation and reduction in formation of human cathelicidin peptide LL-37.

[0487] While not wishing to be bound by theory, it is proposed that treatment of psoriasis with a therapeutically effective dose of ammonia oxidizing bacteria as described herein may involve downregulation of inflammation due to NO generation.

[0488] While not wishing to be bound by theory, it is proposed that treatment of impetigo or other skin and soft tissue infections with a therapeutically effective dose of ammonia oxidizing bacteria as described herein may involve limiting and/or inhibiting the spread and proliferation of *Staphylococcus aureus* (*S. aureus*), e.g., methicillin resistant *Staphylococcus aureus*, *Pseudomonas aeruginosa* (*P.*

aeruginosa), *Streptococcus pyogenes* (*S. pyogenes*), *Acinetobacter baumannii* (*A. baumannii*), Propionibacteria, and *Stenotrophomonas*.

[0489] The present disclosure also provides a method of promoting wound healing, comprising administering to a wound an effective dose of ammonia oxidizing bacteria as described herein. Similarly, the present disclosure provides ammonia oxidizing bacteria as described herein for use in treating a wound. Likewise, the present disclosure provides a use of ammonia oxidizing bacteria as described herein in the manufacture of a medicament or a composition for treating a wound.

[0490] Ammonia oxidizing bacteria as described herein may be used to promote wound healing in a patient that has an impaired healing ability, e.g., a diabetic patient.

[0491] In some embodiments, this disclosure provides methods of using ammonia oxidizing bacteria as described herein to prevent a disease or disorder, e.g., a skin disorder. Prevention, in certain embodiments, means reducing the risk of a subject developing a disease, compared to a similar untreated subject. The risk need not be reduced to zero.

[0492] Individuals having a reduced bathing frequency, such as astronauts, submarine crew members, military personnel during a campaign, civilian workers in remote locations, refugees, bedridden individuals and many others may maintain healthier skin by maintaining ammonia oxidizing bacteria on the skin. With regard to bedridden individuals, the ammonia oxidizing bacteria in some embodiments reduces the frequency or severity of bed sores by augmenting inadequate circulation.

[0493] It is appreciated that many modern degenerative diseases may be caused by a lack of NO species, and that ammonia oxidizing bacteria on the external skin can supply those species by diffusion, and that application of ammonia oxidizing bacteria to the skin resolves long standing medical conditions. In certain embodiments, ammonia oxidizing bacteria are applied to a subject to offset modern bathing practices, especially with anionic detergents remove ammonia oxidizing bacteria from the external skin.

[0494] One suitable method of topical application to apply sufficient ammonia oxidizing bacteria and then wear sufficient clothing so as to induce sweating. However, many people will want to derive the benefits of ammonia oxidizing bacteria while maintaining their current bathing habits, in which case, a culture of the bacteria can be applied along with sufficient substrate for them to produce NO. A nutrient solution approximating the inorganic composition of human sweat can be used for this purpose. Using bacteria adapted to media approximating human sweat minimizes the time for them to adapt when applied. Since sweat evaporates once excreted onto the skin surface, using a culture media that has a higher ionic strength is desirable. A concentration approximately twice that of human sweat is suitable, but other conditions are also contemplated. Ammonia oxidizing bacteria's nutritional needs are typically met with NH₃ or urea, O₂, CO₂, and minerals. In some embodiments, the substrate comprises trace minerals including iron, copper, zinc, cobalt, molybdenum, manganese, sodium, potassium, calcium, magnesium, chloride, phosphate, sulfate, or any combination thereof.

[0495] In some embodiments, the present disclosure provides a method of treating a wound by applying a bandage comprising ammonia oxidizing bacteria to the wound. Also provided are methods of producing such a bandage. The

bandage may comprise, for example, an adhesive portion to affix the bandage to undamaged skin near the wound and a soft, flexible portion to cover or overlay the wound. In some embodiments, the bandage contains no other organisms but ammonia oxidizing bacteria. The bandage may be made of a permeable material that allows gasses like oxygen and carbon dioxide to reach the ammonia oxidizing bacteria when the bandage is applied to the wound. In certain embodiments, the bandage comprises nutrients for ammonia oxidizing bacteria such as ammonium, ammonia, urea, or trace minerals. In certain embodiments, the bandage comprises an antibiotic to which the ammonia oxidizing bacteria is resistant. The antibiotic resistance may arise from one or more endogenous resistance gene or from one or more transgenes.

[0496] In some embodiments, the ammonia oxidizing bacteria, e.g., a preparation of ammonia oxidizing bacteria, is administered at a dose of about 10^8 - 10^9 CFU, 10^9 - 10^{10} CFU, 10^{10} - 10^{11} CFU, or 10^{11} - 10^{12} CFU per application or per day. In some embodiments, the ammonia oxidizing bacteria is administered topically at a dose of about 10^9 - 10^{10} CFU, e.g., about 1×10^9 - 5×10^9 , 1×10^9 - 3×10^9 , or 1×10^9 - 10×10^9 CFU.

[0497] In some embodiments, the ammonia oxidizing bacteria is administered in a volume of about 1-2, 2-5, 5-10, 10-15, 12-18, 15-20, 20-25, or 25-50 ml per dose. In some embodiments, the solution is at a concentration of about 10^8 - 10^9 , 10^9 - 10^{10} , or 10^{10} - 10^{11} CFU/ml. In some embodiments, the ammonia oxidizing bacteria is administered as two 15 ml doses per day, where each dose is at a concentration of 10^9 CFU/ml.

[0498] In some embodiments, the ammonia oxidizing bacteria is administered once, twice, three, or four times per day. In some embodiments, the ammonia oxidizing bacteria is administered once, twice, three, four, five, or six times per week. In some embodiments, the ammonia oxidizing bacteria is administered shortly after bathing. In some embodiments, the ammonia oxidizing bacteria is administered shortly before sleep.

[0499] In some embodiments, the ammonia oxidizing bacteria is administered for about 1-3, 3-5, 5-7, 7-9, 5-10, 10-14, 12-18, 12-21, 21-28, 28-35, 35-42, 42-49, 49-56, 46-63, 63-70, 70-77, 77-84, 84-91 days, e.g., for about 1 month, for about 2 months, for about 3 months. In some embodiments, the ammonia oxidizing bacteria is administered for an indefinite period of time, e.g., greater than one year, greater than 5 years, greater than 10 years, greater than 15 years, greater than 30 years, greater than 50 years, greater than 75 years.

6. EXPERIMENTAL MODELS FOR REFINING D23 TREATMENTS

[0500] Treatments comprising ammonia oxidizing bacteria as described herein (optionally in combination with another therapy) can be refined using a number of model systems. These model systems can be used to determine suitable doses and timing of administration.

[0501] For instance, with respect to chronic wounds and diabetic ulcers, one may use the mouse skin puncture model. Other models for these disorders include controlled cutaneous ischemia in a guinea pig model, rabbit ear ulcer model, application of calcium to a wound, or topical application of doxorubicin.

[0502] With respect to acne, one may use (for example) the Mexican hairless dog model, the Rhino mouse model, or

the rabbit ear assay. With respect to rosacea, one may use (for example) intradermal injection of LL-37 into mouse skin or the Syrian hamster model. With respect to eczema, one may use (for example) application of a crude extract of *Dermatophagoides farinae*, application of dinitrochlorobenzene to the ears of sensitized guinea pigs, or NC/Nga mice. With respect to psoriasis, one may use (for example) xenograft models in which involved and uninvolved psoriatic skin are transplanted onto immunodeficient mice, application of an antibody directed against interleukin 15 to the skin of SCID mice, and the Sharpin^{cpdm}/Sharpin^{cpdm} mouse model.

[0503] Treatments comprising ammonia oxidizing bacteria, e.g., *N. eutropha* as described herein (e.g., strain D23) (optionally in combination with another therapy) can be refined using a number of model systems. These model systems can be used to determine suitable doses and timing of administration.

[0504] For instance, with respect to chronic wounds and diabetic ulcers, one may use the mouse skin puncture model described herein in Example 6. Other models for these disorders include controlled cutaneous ischemia in a guinea pig model, rabbit ear ulcer model, application of calcium to a wound, or topical application of doxorubicin.

[0505] With respect to acne, one may use (for example) the Mexican Hairless Dog model, the Rhino mouse model, or the rabbit ear assay. With respect to rosacea, one may use (for example) intradermal injection of LL-37 into mouse skin or the Syrian hamster model. With respect to eczema, one may use (for example) application of a crude extract of *Dermatophagoides farinae*, application of dinitrochlorobenzene to the ears of sensitized guinea pigs, or NC/Nga mice. With respect to psoriasis, one may use (for example) xenograft models in which involved and uninvolved psoriatic skin are transplanted onto immunodeficient mice, application of an antibody directed against interleukin 15 to the skin of SCID mice, and the Sharpin^{cpdm}/Sharpin^{cpdm} mouse model.

7. MECHANISM OF THERAPEUTIC BENEFIT

[0506] While not wishing to be bound by theory, it is believed that one or more of the following mechanisms contributes to the beneficial effect of ammonia oxidizing bacteria, e.g., *N. eutropha* in treating the diseases and conditions discussed herein. Additional mechanistic details are found in International Application WO/2005/030147, which is herein incorporated by reference in its entirety.

[0507] In order to understand the beneficial aspects of these bacteria, it is helpful to understand angiogenesis. All body cells, except those within a few hundred microns of the external air, receive all metabolic oxygen from the blood supply. The oxygen is absorbed by the blood in the lung, is carried by red blood cells as oxygenated hemoglobin to the peripheral tissues, where it is exchanged for carbon dioxide, which is carried back and exhaled from the lung. Oxygen must diffuse from the erythrocyte, through the plasma, through the endothelium and through the various tissues until it reached the mitochondria in the cell which consumes it. The human body contains about 5 liters of blood, so the volume of the circulatory system is small compared to that of the body. Oxygen is not actively transported. It passively diffuses down a concentration gradient from the air to the erythrocyte, from the erythrocyte to the cell, and from the cell to cytochrome oxidase where it is consumed. The

concentration of oxygen at the site of consumption is the lowest in the body, and the O₂ flux is determined by the diffusion resistance and the concentration gradient. Achieving sufficient oxygen supply to all the peripheral tissues requires exquisite control of capillary size and location. If the spacing between capillaries were increased, achieving the same flux of oxygen would require a larger concentration difference and hence a lower O₂ concentration at cytochrome oxidase. With more cells between capillaries, the O₂ demand would be greater. If the spacing between capillaries were decreased, there would be less space available for the cells that perform the metabolic function of the organ.

[0508] In certain aspects, it is appreciated that NO from ammonia oxidizing bacteria is readily absorbed by the outer skin and converted into S-nitrosothiols since the outer skin is free from hemoglobin. M. Stucker et al. have shown that the external skin receives all of its oxygen from the external air in "The cutaneous uptake of atmospheric oxygen contributes significantly to the oxygen supply of human dermis and epidermis. (*Journal of Physiology* (2002), 538.3, pp. 985-994.) This is readily apparent, because the external skin can be seen to be essentially erythrocyte free. There is circulation of plasma through these layers because they are living and do require the other nutrients in blood, just not the oxygen. S-nitrosothiols formed are stable, can diffuse throughout the body, and constitute a volume source of authentic NO and a source of NO to transnitrosate protein thiols.

[0509] In some aspects, it is appreciated that capillary rarefaction may be one of the first indications of insufficient levels of NO. F. T. Tarek et al. have shown that sparse capillaries, or capillary rarefaction, is commonly seen in people with essential hypertension. (*Structural Skin Capillary Rarefaction in Essential Hypertension. Hypertension*. 1999; 33:998-1001

[0510] A great many conditions are associated with the capillary density becoming sparser. Hypertension is one, and researchers reported that sparse capillaries are also seen in the children of people with essential hypertension, and also in people with diabetes. Significant complications of diabetes are hypertension, diabetic nephropathy, diabetic retinopathy, and diabetic neuropathy. R. Candido et al. have found that the last two conditions are characterized by a reduction in blood flow to the affected areas prior to observed symptoms. (*Haemodynamics in microvascular complications in type 1 diabetes. Diabetes Metab Res Rev* 2002; 18: 286-304.) Reduced capillary density is associated with obesity, and simple weight loss increases capillary density as shown by A Philip et al. in "Effect of Weight Loss on Muscle Fiber Type, Fiber Size, Capilarity, and Succinate Dehydrogenase Activity in Humans. *The Journal of Clinical Endocrinology & Metabolism* Vol. 84, No. 11 4185-4190, 1999.

[0511] Researchers have shown that in primary Raynaud's phenomena (PRP), the nailfold capillaries are sparser (slightly) than in normal controls, and more abundant than in patients that have progressed to systemic sclerosis (SSc). M. Bukhari, Increased Nailfold Capillary Dimensions In Primary Raynaud's Phenomenon And Systemic Sclerosis. *British Journal of Rheumatology*, Vol. 24 No 35: 1127-1131, 1996. They found that the capillary density decreased from 35 loops/mm² (normal controls) to 33 (PRP), to 17 (SSc).

The average distance between capillary limbs was 18μ, 18μ, and 30μ for controls, PRP and SSc, respectively.

[0512] In certain aspects, it is appreciated that the mechanism that the body normally uses to sense "hypoxia" may affect the body's system that regulates capillary density. According to this aspect of the invention, a significant component of "hypoxia" is sensed, not by a decrease in O₂ levels, but rather by an increase in NO levels. Lowering of basal NO levels interferes with this "hypoxia" sensing, and so affects many bodily functions regulated through "hypoxia." For Example, anemia is commonly defined as "not enough hemoglobin," and one consequence of not enough hemoglobin is "hypoxia", which is defined as "not enough oxygen." According to some aspects, these common definitions do not account for the nitric oxide mediated aspects of both conditions.

[0513] At rest, acute isovolemic anemia is well tolerated. A 2/3 reduction in hematocrit has minimal effect on venous return PvO₂, indicating no reduction in either O₂ tension or delivery throughout the entire body. Weiskopf et al. Human cardiovascular and metabolic response to acute, severe isovolemic anemia. *JAMA* 1998, vol 279, No. 3, 217-221. At 50% reduction (from 140 to 70 g Hb/L), the average PvO₂ (over 32 subjects) declined from about 77% to about 74% (of saturation). The reduction in O₂ capacity of the blood is compensated for by vasodilatation and tachycardia with the heart rate increasing from 63 to 85 bpm. That the compensation is effective is readily apparent, however, the mechanism is not. A typical explanation is that "hypoxia" sensors detected "hypoxia" and compensated with vasodilatation and tachycardia. However, there was no "hypoxia" to detect. There was a slight decrease in blood lactate (a marker for anaerobic respiration) from 0.77 to 0.62 mM/L indicating less anaerobic respiration and less "hypoxia." The 3% reduction in venous return PvO₂ is the same level of "hypoxia" one would get by ascending 300 meters in altitude (which typically does not produce tachycardia). With the O₂ concentration in the venous return staying the same, and the O₂ consumption staying the same, there is no place in the body where there is a reduction in O₂ concentration. Compensation during isovolemic anemia may not occur because of O₂ sensing.

[0514] Thus the vasodilatation that is observed in acute isovolemic anemia may be due to the increased NO concentration at the vessel wall. NO mediates dilatation of vessels in response to shear stress and other factors. No change in levels of NO metabolites would be observed, because the production rate of NO is unchanged and continues to equal the destruction rate. The observation of no "hypoxic" compensation with metHb substitution can be understood because metHb binds NO just as Hb does, so there is no NO concentration increase with metHb substitution as there is with Hb withdrawal.

[0515] Nitric oxide plays a role in many metabolic pathways. It has been suggested that a basal level of NO exerts a tonal inhibitory response, and that reduction of this basal level leads to a dis-inhibition of those pathways. Zanzinger et al. have reported that NO has been shown to inhibit basal sympathetic tone and attenuate excitatory reflexes. (Inhibition of basal and reflex-mediated sympathetic activity in the RVLM by nitric oxide. *Am. J. Physiol.* 268 (Regulatory Integrative Comp. Physiol. 37): R958-R962, 1995.)

[0516] In some aspects, it is appreciated that one component of a volume source of NO is low molecular weight

S-nitrosothiols produced in the erythrocyte free skin from NO produced on the external skin by ammonia oxidizing bacteria. These low molecular weight S-nitrosothiols are stable for long periods, and can diffuse and circulate freely in the plasma. Various enzymes can cleave the NO from various S-nitrosothiols liberating NO at the enzyme site. It is the loss of this volume source of NO from AOB on the skin that leads to disruptions in normal physiology. The advantage to the body of using S-nitrosothiols to generate NO far from a capillary is that O₂ is not required for NO production from S-nitrosothiols. Production of NO from nitric oxide synthase (NOS) does require O₂. With a sufficient background of S-nitrosothiols, NO can be generated even in anoxic regions. Free NO is not needed either since NO only exerts effects when attached to another molecule, such as the thiol of a cysteine residue or the iron in a heme, so the effects of NO can be mediated by transnitrosation reactions even in the absence of free NO provided that S-nitrosothiols and transnitrosation enzymes are present.

[0517] Frank et al. have shown that the angiogenesis that accompanies normal wound healing is produced in part by elevated VEGF which is induced by increased nitric oxide. (Nitric oxide triggers enhanced induction of vascular endothelial growth factor expression in cultured keratinocytes (HaCaT) and during cutaneous wound repair. *FASEB J.* 13, 2002-2014 (1999).)

[0518] NO has a role in the development of cancer, indicating that the bacteria described herein may be used in methods of cancer treatment and prevention. According to certain aspects, it is appreciated that the presence of NO during hypoxia may prevent cells from dividing while under hypoxic stress, when cells are at greater risk for errors in copying DNA. One relevant cell function is the regulation of the cell cycle. This is the regulatory program which controls how and when the cell replicates DNA, assembles it into duplicate chromosomes, and divides. The regulation of the cell cycle is extremely complex, and is not fully understood. However, it is known that there are many points along the path of the cell cycle where the cycle can be arrested and division halted until conditions for doing so have improved. The p53 tumor suppressor protein is a key protein in the regulation of the cell cycle, and it serves to initiate both cell arrest and apoptosis from diverse cell stress signals including DNA damage and p53 is mutated in over half of human cancers as reported by Ashcroft et al. in "Stress Signals Utilize Multiple Pathways To Stabilize p53." (*Molecular And Cellular Biology*, May 2000, p. 3224-3233.) Hypoxia does initiate accumulation of p53, and while hypoxia is important in regulating the cell cycle, hypoxia alone fails to induce the downstream expression of p53 mRNA effector proteins and so fails to cause arrest of the cell cycle. Goda et al. have reported that hypoxic induction of cell arrest requires hypoxia-inducing factor-1 (HIF-1 α). (*Hypoxia-Inducible Factor 1 α Is Essential for Cell Cycle Arrest during Hypoxia*. *Molecular And Cellular Biology*, January 2003, p. 359-369.) Britta et al. have reported that NO is one of the main stimuli for HIF-1 α . (Accumulation of HIF-1 α under the influence of nitric oxide. *Blood*, 15 Feb. 2001, Volume 97, Number 4.) In contrast, NO does cause the accumulation of transcriptionally active p53 and does cause arrest of the cell cycle and does cause apoptosis. Wang et al., P53 Activation By Nitric Oxide Involves Down-Regulation Of Mdm2. *THE JOURNAL OF BIOLOGICAL CHEMISTRY* Vol. 277, No. 18, Issue Of May 3, Pp. 15697-15702, 2002.

[0519] In certain aspect of the invention, it is appreciated that preventing the necrotic death of cells by preventing the capillary rarefaction that leads to their hypoxic death may prevent autoimmune disorders. When cells are exposed to chronic hypoxia, the production of reactive oxygen species (ROS) is increased, and there is increased damage to the cells metabolic machinery and ultimately to the cells' DNA. Decreased metabolic capacity will decrease capacity for repair of damage due to ROS and due to exogenous carcinogen exposure. Over time, the damage accumulates and increases the chance of three events: the cell will undergo deletion of cancer-preventing genes and the cell will become cancerous, the cell will die through necrosis, or the cell will die through apoptosis. When cells die, either through necrosis or apoptosis, the cell debris must be cleared from the site. Dead cells are phagocytosed by immune cells, including dendritic cells and macrophages. When these cells phagocytose a body, it is digested by various proteolytic enzymes into antigenic fragments, and then these antigens are attached to the major histocompatibility complex (MHC1, MHC2) and the antigen-MHC complex is moved to the surface of the cell where it can interact with T cells and activate the T cells in various ways. Any cell injury releases adjuvants which stimulate the immune system in various ways. In general, cells that undergo necrosis stimulate a greater immune response than cells that undergo apoptosis. Chronic exposure of immune cells to dead and dying cells is therefore likely to lead to autoimmune disorders.

[0520] In certain aspects, it is appreciated that low basal NO leads to fibrotic hypertrophy. Once a dead cell has been cleared, a new cell cannot easily take its place, because there is insufficient O₂ to support it. Any such new cell would suffer the same fate. The space can remain empty, in which case the organ shrinks, the capillaries draw closer together, new cells are now deprived of the VEGF formerly produced by the now-missing cell, so capillaries ablate and the hypoxic zone reforms. This could result in a general shrinkage of the affected tissues. In tissues that support fibrosis, relatively inert collagen fibers can fill the space. Since the metabolic requirements of the body for the particular organ in question are not reduced, the organ may attempt to grow larger, but now with a significant fibrous content. This may result in fibrotic hypertrophy, such as of the heart and liver. Some organs, such as the brain, cannot grow larger or smaller because the three-dimensional connectivity of nerves and blood vessels are important, and cannot be continuously and simultaneously mapped onto an asymmetrically shrinking brain. The space must be filled with something, and β -amyloid might be the (not so inert) space filler. The kidney cannot grow larger because of the renal capsule, so the number of living cells becomes smaller and they are replaced with fibrotic tissue. If the dead cells are cleared, the tissue shrinks, and the ratio of NO/O₂ goes down again, and the capillaries again become sparser. This may set up the vicious circle of end stage renal disease, congestive heart failure/cardiac hypertrophy, primary biliary cirrhosis, Alzheimer's disease, atherosclerosis, inflammatory bowel disease, hypertrophic scar formation, and the multiple connective tissue diseases starting with Raynaud's phenomena and ending with Systemic Sclerosis and primary Sjogren's syndrome where capillary rarefaction is also observed. Ferrini et al, have shown that a reduction in basal NO levels through chronic inhibition of NOS with L-NAME leads to generalized fibrosis of the heart and kidneys. (Antifibrotic

Role of Inducible Nitric Oxide Synthase. Nitric Oxide: Biology and Chemistry Vol. 6, No. 3, pp. 283-294 (2002).) It may be that low basal NO leads to fibrotic hypertrophy.

[0521] In certain aspects, it is appreciated that capillary rarefaction affects a subject's ability to control their appetite. Capillary rarefaction is observed in the brains of aged humans and animals. Capillary rarefaction is associated with declines in circulating growth factors including insulin like growth factor-1. Neurogenesis in the adult brain is coordinated with angiogenesis. Since the brain regulates many homeostatic functions, increased diffusion lengths between capillaries to control elements of the brain might be "interpreted" as inadequate blood concentrations of those species. The flux of glucose in the brain is quite close to normal metabolic needs, where glucose flux is only 50 to 75% greater than glucose consumption and the glucose transporters across the blood brain barrier are saturable, stereospecific and independent of energy or ion gradients. A large part of the regulation of appetite is mediated through the brain, and capillary rarefaction may cause an adequate blood concentration of "nutrients" (or marker compounds proportional to "nutrients") to be interpreted as insufficient. This may be one cause of obesity.

[0522] According to certain aspects, it is appreciated that capillary rarefaction may be a cause of non-insulin dependent diabetes. Non-insulin dependent diabetes (NIDDM) is also known as the Metabolic Syndrome or Diabetes type 2, and is characterized by insulin resistance. The sensitivity of the body to insulin is reduced, and insulin levels increase. People with NIDDM have high blood glucose, high blood triglycerides, are typically obese, hypertensive, and typically have significant visceral fat.

[0523] Other symptoms accompany NIDDM, which may point to capillary rarefaction as the cause. In a study of 40 men, with and without NIDDM, obese (BMI 29) and lean (BMI 24) (10 of each), Konrad et al. report that blood lactate levels at rest were 1.78, 2.26, 2.42, and 2.76 (mM/L) for lean men without, obese men without, lean men with NIDDM, obese men with NIDDM respectively. (A-Lipoic acid treatment decreases serum lactate and pyruvate concentrations and improves glucose effectiveness in lean and obese patients with type 2 diabetes. Diabetes Care 22:280-287, 1999.) Lactate is a measure of anaerobic glycolysis. When O_2 is insufficient to generate ATP through oxidative phosphorylation, cells can produce ATP through anaerobic glycolysis. One of the products of anaerobic glycolysis is lactate, which must be exported from the cells, otherwise the pH drops and function is compromised. Blood lactate is commonly measured in exercise studies, where an increase indicates the work load at which maximum oxidative work can be done. Higher levels of lactate at rest would indicate increased anaerobic glycolysis at rest, which is consistent with capillary rarefaction.

[0524] Primary biliary cirrhosis is associated with Raynaud's phenomena, pruritus, sicca syndrome, osteoporosis, portal hypertension, neuropathy, and pancreatic insufficiency, and liver abnormalities are associated with rheumatic diseases. Elevated liver enzymes are a symptom of liver inflammation, and elevated liver enzymes are observed as an early symptom of "asymptomatic" primary biliary cirrhosis. Accordingly, the bacteria described herein may be used to treat liver inflammation.

[0525] Torre et al have reported that Alzheimer's disease (AD) is a microvascular disorder with neurological degen-

eration secondary to hypoperfusion, resulting in part from insufficient nitric oxide. (Review: Evidence that Alzheimer's disease is a microvascular disorder: the role of constitutive nitric oxide, Brain Research Reviews 34 (2000) 119-136.) Accordingly, the bacteria described herein may be used to treat AD.

[0526] Adverse health effects that are associated with hypertension may also be consequences of low basal NO. The decreased response to vasodilatation is also consistent with low basal NO. NO is a diffusible molecule that diffuses from a source to a sensor site where it has the signaling effect. With low NO levels, every NO source must produce more NO to generate an equivalent NO signal of a certain intensity a certain distance away. NO diffuses in three dimensions and the whole volume within that diffusion range must be raised to the level that will give the proper signal at the sensor location. This may result in higher NO levels at the source and between the source and the sensor. Adverse local effects of elevated NO near a source may then arise from too low a NO background. There is some evidence that this scenario actual occurs. In rat pancreatic islets, Henningson et al have reported that inhibition of NOS with L-NAME increases total NO production through the induction of iNOS. (Chronic blockade of NO synthase paradoxically increases islet NO production and modulates islet hormone release. Am J Physiol Endocrinol Metab 279: E95-E107, 2000.) Increasing NO by increasing NOS activity will only work up to some limit. When NOS is activated but is not supplied with sufficient tetrahydrobiopterin (BH4) or L-arginine, it becomes "uncoupled" and generates superoxide (O_2^-) instead of NO. This O_2^- may then destroy NO. Attempting to produce NO at a rate that exceeds the supply of BH4 or L-arginine may instead decrease NO levels. This may result in positive feedback where low NO levels are made worse by stimulation of NOS, and uncoupled NOS generates significant O_2^- which causes local reactive oxygen species (ROS) damage such as is observed in atherosclerosis, end stage renal disease, Alzheimer's, and diabetes.

[0527] The bacteria described herein may also be used to delay the signs of aging. Caloric restriction extends lifespan, and Holloszy reported that restricting food intake to 70% of ad lib controls, prolongs life in sedentary rats from 858 to 1,051 days, almost 25%. (Mortality rate and longevity of food restricted exercising male rats: a reevaluation. J. Appl. Physiol. 82(2): 399-403, 1997.) The link between calorie restriction and prolonged life is well established, however, the causal mechanism is not. Lopez-Torres et al. reported that the examination of liver mitochondrial enzymes in rats indicates a reduction in H_2O_2 production due to reduced complex I activity associated with calorie restriction. (Influence Of Aging And Long-Term Caloric Restriction On Oxygen Radical Generation And Oxidative DNA Damage In Rat Liver Mitochondria. Free Radical Biology & Medicine Vol. 32 No 9 pp 882-889, 2002.) H_2O_2 is produced by dismutation of O_2^- , which is a major ROS produced by the mitochondria during respiration. The main source of O_2^- has been suggested by Kushareva et al. and others to be complex I which catalyzes the NAD/NADH redox couple by reverse flow of electrons from complex III, the site of succinate reduction. The free radical theory, proposed by Beckman, of aging postulates, that free radical damage to cellular DNA, antioxidant systems and DNA repair systems accumulates with age and when critical systems are damaged beyond

repair, death ensues. (The Free Radical Theory of Aging Matures. *Physiol. Rev.* 78: 547-581, 1998.)

[0528] As an additional mechanism, NO has been demonstrated by Vasa et al. to activate telomerase and to delay senescence of endothelial cells. (Nitric Oxide Activates Telomerase and Delays Endothelial Cell Senescence. *Circ Res.* 2000; 87:540-542.) Low basal NO will increase basal metabolic rate by disinhibition of cytochrome oxidase. Increased basal metabolism will also increase cell turn-over and growth rate. Capillary rarefaction, by inducing chronic hypoxia may increase free radical damage and may also increase cell turn-over, and so accelerate aging by both mechanisms.

[0529] In some aspects, it is appreciated that autotrophic ammonia-oxidizing bacteria may produce protective aspects for allergies and autoimmune disorders. The best known autoimmune disease is perhaps Diabetes Type 1, which results from the destruction of the insulin producing cells in the pancreas by the immune system. Recurrent pregnancy loss is also associated with autoimmune disorders where the number of positive autoimmune antibodies correlated positively with numbers recurrent pregnancy losses. Systemic Sclerosis, Primary Biliary Cirrhosis, autoimmune hepatitis, and the various rheumatic disorders are other examples of autoimmune disorders. Application of AOB was observed to reduce an allergy, hay fever, as described in WO/2005/030147.

[0530] One mechanism by which AOB may exert their protective effect on allergies and autoimmune disorders is through the production of nitric oxide, primarily through the regulatory inhibition of NF- κ B and the prevention of activation of immune cells and the induction of inflammatory reactions. NF- κ B is a transcription factor that up-regulates gene expression and many of these genes are associated with inflammation and the immune response including genes which cause the release of cytokines, chemokines, and various adhesion factors. These various immune factors cause the migration of immune cells to the site of their release resulting in the inflammation response. Constitutive NO production has been shown to inhibit NF- κ B by stabilizing I κ B α (an inhibitor of NF- κ B) by preventing I κ B α degradation.

[0531] Administration of an NO donor has been shown by Xu et al. to prevent the development of experimental allergic encephalomyelitis in rats. (SIN-1, a Nitric Oxide Donor, Ameliorates Experimental Allergic Encephalomyelitis in Lewis Rats in the Incipient Phase: The Importance of the Time Window. *The Journal of Immunology*, 2001, 166: 5810-5816.) In this study, it was demonstrated that administering an NO donor, reduced the infiltration of macrophages into the central nervous system, reduced the proliferation of blood mononuclear cells, and increased apoptosis of blood mononuclear cells. All of these results are expected to reduce the extent and severity of the induced autoimmune response.

[0532] Low basal NO may lead to autism via the mechanism that new connections in the brain are insufficiently formed as a result of insufficient basal nitric oxide. While not wishing to be bound in theory, in some embodiments, formation of neural connections is modulated by NO. In these cases, any condition that lowers the range of NO diffusion may decrease the volume size of brain elements that can undergo connections. A brain which developed

under conditions of low basal NO levels may be arranged in smaller volume elements because the reduced effective range of NO.

[0533] Additional symptoms exhibited in autistic individuals may also point to low NO as a cause, including increased pitch discrimination, gut disturbances, immune system dysfunction, reduced cerebral blood flow, increased glucose consumption of the brain, increased plasma lactate, attachment disorders, and humming. Each of these symptoms may be attributed to a low basal NO level.

[0534] Takashi Ohnishi et al. have reported that autistic individuals show decreased blood flow. Takashi Ohnishi et al., Abnormal regional cerebral blood flow in childhood autism. *Brain* (2000), 123, 1838-1844. J. M. Rumsey et al. have reported that autistic individuals have increased glucose consumption. Rumsey J M, Duara R, Grady C, Rapoport J L, Margolin R A, Rapoport S I, Cutler N R. Brain metabolism in autism. Resting cerebral glucose utilization rates as measured with positron emission tomography. *Arch Gen Psychiatry*, 1985 May; 42(5):448-55 (abstract). D. C. Chugani has reported that autistic individuals have an increased plasma lactate levels. Chugani D C, et al., Evidence of altered energy metabolism in autistic children. *Prog Neuropsychopharmacol Biol Psychiatry*. 1999 May; 23(4): 635-41. The occurrence of these effects may be a result of capillary rarefaction in the brain, which may reduce blood flow and O₂ supply, such that some of the metabolic load of the brain may be produced through glycolysis instead of oxidative phosphorylation.

[0535] Nitric oxide has been demonstrated by B. A. Klyachko et al. to increase the excitability of neurons by increasing the after hyperpolarization through cGMP modification of ion channels. Vitaly A. Klyachko et al., cGMP-mediated facilitation in nerve terminals by enhancement of the spike after hyperpolarization. *Neuron*, Vol. 31, 1015-1025, Sep. 27, 2001. C. Sandie et al. have shown that inhibition of NOS reduces startle. Carmen Sandi et al., Decreased spontaneous motor activity and startle response in nitric oxide synthase inhibitor-treated rats. *European journal of pharmacology* 277 (1995) 89-97. Attention-Deficit Hyperactivity Disorder (ADHD) has been modeled using the spontaneously hypertensive rat (SHR) and the Naples high-excitability (NHE) rat. Both of these models have been shown by Raffaele Aspidi et al, to show increased attention deficits during periods of acute NOS inhibition. Raffaele Aspidi et al., Non-selective attention and nitric oxide in putative animal models of attention-deficit hyperactivity disorder. *Behavioral Brain Research* 95 (1998) 123-133. Accordingly, the bacteria herein may be used in the treatment of ADHD.

[0536] Inhibition of NOS has also been shown by M. R. Dzoljic to inhibit sleep. M. R. Dzoljic, R. de Vries, R. van Leeuwen. Sleep and nitric oxide: effects of 7-nitro indazole, inhibitor of brain nitric oxide synthase. *Brain Research* 718 (1996) 145-150. G. Zoccoli has reported that a number of the physiological effects seen during sleep are altered when NOS is inhibited, including rapid eye movement and sleep-wake differences in cerebral circulation. G. Zoccoli, et al., Nitric oxide inhibition abolishes sleep-wake differences in cerebral circulation. *Am. J. Physiol. Heart Circ Physiol* 280: H2598-2606, 2001. NO donors have been shown by L. Kapas et al. to promote non-REM sleep, however, these increases persisted much longer than the persistence of the NO donor, suggesting perhaps a rebound effect. Levente

Kapas et al. Nitric oxide donors SIN-1 and SNAP promote nonrapid-eye-movement sleep in rats. *Brain Research Bulletin*, vol 41, No 5, pp. 293-298, 1996. M. Rosaria et al., Central NO facilitates both penile erection and yawning. Maria Rosaria Melis and Antonio Argiolas. Role of central nitric oxide in the control of penile erection and yawning. *Prog Neuro-Psychopharmacol & Biol. Psychiatr.* 1997, vol 21, pp 899-922. P. Tani et al. have reported that insomnia is a frequent finding in adults with Asperger's. Pekka Tani et al., Insomnia is a frequent finding in adults with Asperger's syndrome. *BMC Psychiatry* 2003, 3:12. Y. Hoshino has also observed sleep disturbances in autistic children. Hoshino Y, Watanabe H, Yashima Y, Kaneko M, Kumashiro H. An investigation on sleep disturbance of autistic children. *Folia Psychiatr Neurol Jpn.* 1984; 38(1):45-51. (abstract) K. A. Schreck et al. has observed that the severity of sleep disturbances correlates with severity of autistic symptoms. Schreck K A, et al., Sleep problems as possible predictors of intensified symptoms of autism. *Res Dev Disabil.* 2004 Jan.-Feb.; 25(1):57-66. (abstract). Accordingly, the bacteria herein may be used in the treatment of insomnia.

[0537] W. D. Ratnasooriya et al reported that inhibition of NOS in male rats reduces pre-coital activity, reduces libido, and reduces fertility. W. D. Ratnasooriya et al., Reduction in libido and fertility of male rats by administration of the nitric oxide (NO) synthase inhibitor N-nitro-L-arginine methyl ester. *International journal of andrology*, 23: 187-191 (2000).

[0538] It may be that a number of seemingly disparate disorders, characterized by ATP depletion and eventual organ failure are actually "caused" by nitropenia, caused by a global deficiency in basal nitric oxide. When this occurs in the heart, the result is dilative cardiomyopathy. When this occurs in the brain, the result is white matter hyperintensity, Alzheimer's, vascular depression, vascular dementia, Parkinson's, and the Lewy body dementias. When this occurs in the kidney, the result is end stage renal disease, when this occurs in the liver, the result is primary biliary cirrhosis. When this occurs in muscle, the consequence is fibromyalgia, Gulf War Syndrome, or chronic fatigue syndrome. When this occurs in the bowel, the consequence is ischemic bowel disease. When this occurs in the pancreas, the consequence is first type 2 diabetes, followed by chronic inflammation of the pancreas, followed by autoimmune attack of the pancreas (or pancreatic cancer), followed by type 1 diabetes. When this occurs in the connective tissue, the consequence is systemic sclerosis.

[0539] In the remnant kidney model of end stage renal disease, part of the kidney is removed, (either surgically or with a toxin) which increases the metabolic load on the remainder. Superoxide is generated to decrease NO and increase O₂ diffusion to the kidney mitochondria. Chronic overload results in progressive kidney capillary rarefaction and progressive kidney failure. In acute kidney failure, putting people in dialysis can give the kidney a "rest", and allows it to recover. In acute renal failure induced by rhabdomyolysis (muscle damage which releases myoglobin into the blood stream) kidney damage is characterized by ischemic damage. Myoglobin scavenges NO, just as hemoglobin does, and would cause vasoconstriction in the kidney leading to ischemia. Myoglobin would also induce local nitropenia and the cascade of events leading to further ATP depletion.

[0540] In some aspects, low NO levels lead to reduced mitochondrial biogenesis. Producing the same ATP at a reduced mitochondria density will result in an increase in O₂ consumption, or an accelerated basal metabolic rate. An accelerated basal metabolic rate is observed in a number of conditions, including: Sickle cell anemia, Congestive heart failure, Diabetes, Liver Cirrhosis, Crohn's disease, Amyotrophic lateral sclerosis, Obesity, End stage renal disease, Alzheimer's, and chronic obstructive pulmonary disease.

[0541] While some increased O₂ consumption might be productively used, in many of these conditions uncoupling protein is also up-regulated, indicating that at least part of the increased metabolic rate is due to inefficiency. Conditions where uncoupling protein is known to be up-regulated include obesity and diabetes.

[0542] With fewer mitochondria consuming O₂ to a lower O₂ concentration, the O₂ gradient driving O₂ diffusion is greater, so the O₂ diffusion path length can increase resulting in capillary rarefaction, which is observed in dilative cardiomyopathy, hypertension, diabetes type 2, and renal hypertension.

[0543] Copper, either as Cu²⁺ or as ceruloplasmin (CP) (the main Cu containing serum protein which is present at 0.38 g/L in adult sera and which is 0.32% Cu and contains 94% of the serum copper) catalyzes the formation of S—NO-thiols from NO and thiol containing groups (RSH). The Cu content of plasma is variable and is increased under conditions of infection. Berger et al. reported that the Cu and Zn content of burn-wound exudates is considerable with patients with 1/3 of their skin burned, losing 20 to 40% of normal body Cu and 5 to 10% of Zn content in 7 days. (Cutaneous copper and zinc losses in burns. *Burns.* 1992 October; 18(5):373-80.) If the patients skin were colonized by AOB, wound exudates which contains urea and Fe, Cu, and Zn that AOB need, would be converted into NO and nitrite, greatly supplementing the local production of NO by iNOS, without consuming resources (such as O₂ and L-arginine) in the metabolically challenged wound. A high production of NO and nitrite by AOB on the surface of a wound would be expected to inhibit infection, especially by anaerobic bacteria such as the Clostridia which cause tetanus, gas gangrene, and botulism.

[0544] The practice of the present invention may employ, unless otherwise indicated, conventional methods of immunology, molecular biology, and recombinant DNA techniques within the skill of the art. Such techniques are explained fully in the literature. See, e.g., Sambrook, et al. *Molecular Cloning: A Laboratory Manual* (Current Edition); and *Current Protocols in Molecular Biology* (F. M. Ausubel, et al. eds., current edition).

8. NUCLEIC ACIDS AND PROTEINS FROM *N. EUTROPHA*

[0545] This disclosure provides, among other things, proteins and nucleic acids (optionally, isolated proteins and nucleic acids) that are identical to or similar to those found in strain D23. While not wishing to be bound by theory, it is believed that the sequenced strain of D23 has non-naturally occurring protein and nucleic acid sequences due to an extended period of culture and selection in the laboratory.

[0546] These nucleic acids and proteins have numerous uses. For instance, the proteins may be used to generate antibodies or other binding molecules that detect strain D23

or related strains. The proteins may also be used to carry out reactions under high-NH₄⁺ conditions, because D23 is adapted for growth and metabolism under these conditions. As another example, the nucleic acids may be used to produce proteins for generating antibodies or carrying out reactions as described above. The nucleic acids may also be used to detect strain D23 or related strains, e.g., using a microarray or another hybridization-based assay.

[0547] The genome of strain D23 is provided as SEQ ID NO: 1. The genome annotation (including the position and orientation of genes within SEQ ID NO: 1) is provided as Supplementary Table 1. Accordingly, this disclosure provides genes and proteins identical or similar to the genes listed in Supplementary Table 1.

[0548] Accordingly, this disclosure provides a nucleic acid (e.g., an isolated nucleic acid) comprising a sequence of a gene of Supplementary Table 1, as well as a protein encoded by said gene. In certain embodiments, the nucleic acid comprises a sequence that is similar (e.g., at least 70%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, or 99.5% identical) to a gene of Supplementary Table 1, or a protein encoded by said gene. The disclosure also provides a composition comprising a nucleic acid that is at least 1, 2, 3, 4, 5, 10, 15, 20, 50, 100, 200, 500, 1,000, 1,500, 2,000, 2,500, or all of the sequences of Supplementary Table 1, or a sequence that is similar thereto (e.g., at least 70%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, or 99.5% identical), or one or more proteins encoded by said nucleic acids. Also provided are fragments of said nucleic acids and proteins.

[0549] The present disclosure also provides, inter alia, one or more genes or proteins that are present in strain D23 and absent from strain C91, or a gene or protein similar to one of said genes or proteins. Examples of these genes are set out in FIGS. 6-8 and are described in more detail in Example 4 herein. Examples of these genes and proteins, as well as genes and proteins similar thereto, are described below.

[0550] Accordingly, with respect to FIG. 6, this application discloses nucleic acids that are identical or similar (e.g., at least 70%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, or 99.5% identical) to 1, 2, 3, 4, 5, 10, 15, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, or all of the sequences in FIG. 6. This application also discloses proteins that are identical or similar (e.g., at least 70%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, or 99.5% identical) to 1, 2, 3, 4, 5, 10, 15, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, or all of the proteins encoded by the genes listed in FIG. 6. Furthermore, the application discloses fragments of these genes and proteins, e.g., fragments of 0-20, 20-50, 50-100, 100-200, 200-500, 500-1000, or greater than 1000 nucleotides or amino acids. In some embodiments, a plurality of the above-mentioned genes or proteins are affixed to a solid support, e.g., to form a microarray.

[0551] With respect to FIG. 7, this application discloses nucleic acids that are identical or similar (e.g., at least 70%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, or 99.5% identical) to 1, 2, 3, 4, 5, 10, 15, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, or all of the sequences in FIG. 7. This application also discloses proteins that are identical or similar (e.g., at least 70%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, or 99.5% identical) to 1, 2, 3, 4, 5, 10, 15, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, or all of the proteins encoded by the genes listed in FIG. 7. Furthermore, the application discloses fragments of these genes and proteins, e.g., fragments of 0-20, 20-50, 50-100,

100-200, 200-500, 500-1000, or greater than 1000 nucleotides or amino acids. In some embodiments, a plurality of the above-mentioned genes or proteins are affixed to a solid support, e.g., to form a microarray.

[0552] With respect to FIG. 8, this application discloses nucleic acids that are identical or similar (e.g., at least 70%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, or 99.5% identical) to 1, 2, 3, 4, 5, 10, 15, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 200, or all of the sequences in FIG. 8. This application also discloses proteins that are identical or similar (e.g., at least 70%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, or 99.5% identical) to 1, 2, 3, 4, 5, 10, 15, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 200, or all of the proteins encoded by the genes listed in FIG. 8. Furthermore, the application discloses fragments of these genes and proteins, e.g., fragments of 0-20, 20-50, 50-100, 100-200, 200-500, 500-1000, or greater than 1000 nucleotides or amino acids. In some embodiments, a plurality of the above-mentioned genes or proteins are affixed to a solid support, e.g., to form a microarray.

[0553] With respect to FIGS. 6-8 collectively, this application discloses nucleic acids that are identical or similar (e.g., at least 70%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, or 99.5% identical) to 1, 2, 3, 4, 5, 10, 20, 40, 60, 80, 100, 150, 200, 250, 300, 350, 400, 450, 500, or all of the sequences in FIGS. 6-8. This application discloses proteins that are identical or similar (e.g., at least 70%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, or 99.5% identical) to 1, 2, 3, 4, 5, 10, 20, 40, 60, 80, 100, 150, 200, 250, 300, 350, 400, 450, 500, or all of the proteins encoded by genes listed in FIGS. 6-8. Furthermore, the application discloses fragments of these genes and proteins, e.g., fragments of 1-20, 20-50, 50-100, 100-200, 200-500, 500-1000, or greater than 1000 nucleotides or amino acids. In some embodiments, a plurality of the above-mentioned genes or proteins are affixed to a solid support, e.g., to form a microarray.

[0554] This disclosure also provides nucleic acid sequences that are fragments of SEQ ID NO: 1. The fragments may be, e.g., 1-20, 20-50, 50-100, 100-200, 200-500, 500-1000, 1,000-2,000, 2,000-5,000, or 10,000 or more nucleotides in length. The fragments may also be at least about 80%, 85%, 90%, 95%, 97%, 98%, 99%, or 99.5% identity to the corresponding portion of SEQ ID NO: 1 or its complement. The fragment may also be a fragment that hybridizes to SEQ ID NO: 1, or to the genome of the D23 strain deposited with the ATCC patent depository on Apr. 8, 2014, designated AOB D23-100 with the ATCC under accession number PTA-121157, or their complements, under low stringency, medium stringency, high stringency, or very high stringency, or other hybridization condition described herein.

[0555] The disclosure also provides nucleic acid sequences set out in Table 1 (which describes genes involved in ammonia metabolism). Accordingly, in some aspects, this application discloses genes that are identical or similar (e.g., at least 70%, 80%, 85%, 90%, 95%, 96%, 97%, 97.5%, 98%, 98.2%, 98.4%, 98.6%, 98.8%, 99%, 99.1%, 99.2%, 99.3%, 99.4%, 99.5%, 99.6%, 99.7%, 99.8%, 99.9%, or 100% identical) to 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, or all of the genes in Table 1. In embodiments, this application discloses proteins that are identical or similar (e.g., at least 70%, 80%, 85%, 90%, 95%, 96%, 97%, 97.5%, 98%, 98.2%, 98.4%, 98.6%, 98.8%, 99%, 99.1%, 99.2%, 99.3%, 99.4%, 99.5%, 99.6%, 99.7%, 99.8%, 99.9%, or

100% identical) to 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, or all of the proteins in Table 1.

[0556] Alignment of the nucleic acid sequences of Table 1 shows the percent identity between homologs in C91 and D23. The following paragraphs discuss this percent identity and describe various nucleic acids having homology to the D23 genes of Table 1.

[0557] More specifically, the amoA1 genes are about 98.8% identical (i.e., at 821/831 positions). Accordingly, in some embodiments, the amoA1 nucleic acid comprises D23 nucleotides at at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or all of the positions that differ in this gene between strains C91 and D23. In embodiments, the amoA1 nucleic acid comprise D23 nucleotides at at most 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or all of the positions that differ in this gene between strains C91 and D23. In embodiments, the amoA1 nucleic acid comprises a sequence at least about 98.8%, 98.9%, 99.0%, 99.2%, 99.4%, 99.6%, 99.8%, or 100% identical to the D23 amoA1 gene.

[0558] The amoA2 genes are about 98.8% identical (i.e., at 821/831 positions). Accordingly, in some embodiments, the amoA2 nucleic acid comprises D23 nucleotides at at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or all of the positions that differ in this gene between strains C91 and D23. In embodiments, the amoA2 nucleic acid comprises D23 nucleotides at at most 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or all of the positions that differ in this gene between strains C91 and D23. In embodiments, the amoA2 nucleic acid comprises a sequence at least about 98.8%, 98.9%, 99.0%, 99.2%, 99.4%, 99.6%, 99.8%, or 100% identical to the D23 amoA2 gene.

[0559] The amoB1 genes are about 99.1% identical (i.e., at 1255/1266 positions). Accordingly, in some embodiments, the amoB1 nucleic acid comprises D23 nucleotides at at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, or all of the positions that differ in this gene between strains C91 and D23. In embodiments, the amoB1 nucleic acid comprises D23 nucleotides at at most 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, or all of the positions that differ in this gene between strains C91 and D23. In embodiments, the amoB1 nucleic acid comprises a sequence at least about 99.1%, 99.2%, 99.4%, 99.6%, 99.8%, or 100% identical to the D23 amoB1 gene.

[0560] The amoB2 genes are about 99.1% identical (i.e., at 1254/1266 positions). Accordingly, in some embodiments, the amoB2 nucleic acid comprises D23 nucleotides at at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12 or all of the positions that differ in this gene between strains C91 and D23. In embodiments, the amoB2 nucleic acid comprises D23 nucleotides at at most 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12 or all of the positions that differ in this gene between strains C91 and D23. In embodiments, the amoB2 nucleic acid comprises a sequence at least about 99.1%, 99.2%, 99.4%, 99.6%, 99.8%, or 100% identical to the D23 amoB2 gene.

[0561] The amoC1 genes are about 99.8% identical (i.e., at 814/816 positions). Accordingly, in some embodiments, the amoC1 nucleic acid comprises D23 nucleotides at at least 1, 2, or all of the positions that differ in this gene between strains C91 and D23. In embodiments, the amoC1 nucleic acid comprises D23 nucleotides at at most 1, 2, or all of the positions that differ in this gene between strains C91 and D23. In embodiments, the amoC1 nucleic acid comprises a sequence at least about 99.8%, 99.9%, or 100% identical to the D23 amoC1 gene.

[0562] The amoC2 genes are about 99.8% identical (i.e., at 814/816 positions). Accordingly, in some embodiments, the

amoC2 nucleic acid comprises D23 nucleotides at at least 1, 2, or all of the positions that differ in this gene between strains C91 and D23. In embodiments, the amoC2 nucleic acid comprises D23 nucleotides at at most 1, 2, or all of the positions that differ in this gene between strains C91 and D23. In embodiments, the amoC2 nucleic acid comprises a sequence at least about 99.8%, 99.9%, or 100% identical to the D23 amoC2 gene.

[0563] The amoC3 genes are about 98.9% identical (i.e., at 816/825 positions). Accordingly, in some embodiments, the amoC3 nucleic acid comprises D23 nucleotides at at least 1, 2, 3, 4, 5, 6, 7, 8, 9, or all of the positions that differ in this gene between strains C91 and D23. In embodiments, the amoC3 nucleic acid comprises D23 nucleotides at at most 1, 2, 3, 4, 5, 6, 7, 8, 9, or all of the positions that differ in this gene between strains C91 and D23. In embodiments, the amoC3 nucleic acid comprises a sequence at least about 98.9%, 99.0%, 99.2%, 99.4%, 99.6%, 99.8%, or 100% identical to the D23 amoC3 gene.

[0564] The hao1 genes are about 99.0% identical (i.e., at 1696/1713 positions). Accordingly, in some embodiments, the hao1 nucleic acid comprises D23 nucleotides at at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, or all of the positions that differ in this gene between strains C91 and D23. In embodiments, the hao1 nucleic acid comprises D23 nucleotides at at most 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, or all of the positions that differ in this gene between strains C91 and D23. In embodiments, the hao1 nucleic acid comprises a sequence at least about 99.0%, 99.2%, 99.4%, 99.6%, 99.8%, or 100% identical to the D23 hao1 gene.

[0565] The hao2 genes are about 99.4% identical (i.e., at 1702/1713 positions). Accordingly, in some embodiments, the hao2 nucleic acid comprises D23 nucleotides at at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, or all of the positions that differ in this gene between strains C91 and D23. In embodiments, the hao2 nucleic acid comprises D23 nucleotides at at most 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, or all of the positions that differ in this gene between strains C91 and D23. In embodiments, the hao2 nucleic acid comprises a sequence at least about 99.4%, 99.6%, 99.8%, or 100% identical to the D23 hao2 gene.

[0566] The hao3 genes are about 99.2% identical (i.e., at 1700/1713 positions). Accordingly, in some embodiments, the hao3 nucleic acid comprises D23 nucleotides at at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, or all of the positions that differ in this gene between strains C91 and D23. In embodiments, the hao3 nucleic acid comprises D23 nucleotides at at most 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, or all of the positions that differ in this gene between strains C91 and D23. In embodiments, the hao3 nucleic acid comprises a sequence at least about 99.2%, 99.4%, 99.6%, 99.8%, or 100% identical to the D23 hao3 gene.

[0567] The cycA1 genes are about 98.0% identical (i.e., at 694/708 positions). Accordingly, in some embodiments, the cycA1 nucleic acid comprises D23 nucleotides at at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, or all of the positions that differ in this gene between strains C91 and D23. In embodiments, the cycA1 nucleic acid comprises D23 nucleotides at at most 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, or all of the positions that differ in this gene between strains C91 and D23. In embodiments, the cycA1 nucleic acid comprises a sequence at least about 98.0%, 98.2%,

98.4%, 98.6%, 98.8%, 99.0%, 99.2%, 99.4%, 99.6%, 99.8%, or 100% identical to the D23 cycA1 gene.

[0568] The cycA2 genes are about 98.7% identical (i.e., at 699/708 positions). Accordingly, in some embodiments, the cycA2 nucleic acid comprises D23 nucleotides at at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, or all of the positions that differ in this gene between strains C91 and D23. In embodiments, the cycA2 nucleic acid comprises D23 nucleotides at at most 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, or all of the positions that differ in this gene between strains C91 and D23. In embodiments, the cycA2 nucleic acid comprises a sequence at least about 98.7%, 98.8%, 99.0%, 99.2%, 99.4%, 99.6%, 99.8%, or 100% identical to the D23 cycA2 gene.

[0569] The cycA3 genes are about 99.3% identical (i.e., at 703/708 positions). Accordingly, in some embodiments, the cycA3 nucleic acid comprises D23 nucleotides at at least 1, 2, 3, 4, 5, or all of the positions that differ in this gene between strains C91 and D23. In embodiments, the cycA3 nucleic acid comprises D23 nucleotides at at most 1, 2, 3, 4, 5, or all of the positions that differ in this gene between strains C91 and D23. In embodiments, the cycA3 nucleic acid comprises a sequence at least about 99.3%, 99.4%, 99.6%, 99.8%, or 100% identical to the D23 cycA3 gene.

[0570] The cycB1 genes are about 96.7% identical (i.e., at 696/720 positions). Accordingly, in some embodiments, the cycB1 nucleic acid comprises D23 nucleotides at at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, or all of the positions that differ in this gene between strains C91 and D23. In embodiments, the cycB1 nucleic acid comprises D23 nucleotides at at most 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, or all of the positions that differ in this gene between strains C91 and D23. In embodiments, the cycB1 nucleic acid comprises a sequence at least about 96.7%, 96.8%, 97.0%, 97.2%, 97.4%, 97.6%, 97.8%, 98.0%, 98.2%, 98.4%, 98.4%, 98.6%, 98.8%, 99.0%, 99.2%, 99.4%, 99.6%, 99.8%, or 100% identical to the D23 cycB1 gene.

[0571] The cycB2 genes are about 97.1% identical (i.e., at 702/723 positions). Accordingly, in some embodiments, the cycB2 nucleic acid comprises D23 nucleotides at at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, or all of the positions that differ in this gene between strains C91 and D23. In embodiments, the cycB2 nucleic acid comprises D23 nucleotides at at most 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, or all of the positions that differ in this gene between strains C91 and D23. In embodiments, the cycB2 nucleic acid comprises a sequence at least about 97.1%, 97.2%, 97.4%, 97.6%, 97.8%, 98.0%, 98.2%, 98.4%, 98.4%, 98.6%, 98.8%, 99.0%, 99.2%, 99.4%, 99.6%, 99.8%, or 100% identical to the D23 cycB2 gene.

[0572] Further provided herein are vectors comprising nucleotide sequences described herein. In some embodiments, the vectors comprise nucleotides encoding a protein described herein. The vectors include, but are not limited to, a virus, plasmid, cosmid, lambda phage or a yeast artificial chromosome (YAC). Such vectors may include a promoter, an open reading frame with or without introns, and a termination signal.

[0573] The present disclosure also provides host cells comprising a nucleic acid as described herein, or a nucleic acid encoding a protein as described herein.

[0574] In certain embodiments, the host cells are genetically engineered by using an expression cassette. The phrase “expression cassette,” refers to nucleotide sequences, which are capable of affecting expression of a gene in hosts compatible with such sequences. Such cassettes may include a promoter, an open reading frame with or without introns, and a termination signal. Additional factors necessary or helpful in effecting expression may also be used, such as, for example, an inducible promoter.

[0575] The disclosure also provides host cells comprising the vectors described herein.

[0576] The cell can be, but is not limited to, a eukaryotic cell, a bacterial cell, an insect cell, or a human cell. If the cell is a bacterial cell, it may be, e.g., *E. coli* or an ammonia-oxidizing bacterium such as *Nitrosomonas* (e.g., *N. eutropha* or *N. europaea*), *Nitrosococcus*, *Nitrospira*, *Nitrosocystis*, *Nitrosolobus*, and *Nitrosovibrio*.

9. ADJUSTING THE SKIN MICROBIOME WITH AMMONIA OXIDIZING BACTERIA

[0577] The present disclosure provides for systems and methods for changing the skin microbiome, e.g., the human skin microbiome. The systems and methods may provide treatment of infections or conditions, e.g., related to the skin, e.g., skin infections and/or skin conditions.

[0578] Ammonia-oxidizing bacteria (AOB) of the genus *Nitrosomonas* are Gram-negative obligate autotrophic bacteria with a unique capacity to generate nitrite and nitric oxide exclusively from ammonia as an energy source. They are widely present both in soil and water environments and are essential components of environmental nitrification processes. Due to the roles of nitrite and nitric oxide on human skin as important components of several physiological functions, such as vasodilation, skin inflammation and wound healing, these bacteria may have beneficial properties for both healthy and immunopathological skin conditions. These bacteria may be safe for use in humans because they are slow-growing, cannot grow on organic carbon sources, may be sensitive to soaps and antibiotics, and have never been associated with any disease or infection in animals or humans.

[0579] Topical application of ammonia oxidizing bacteria to a subject, e.g., a human subject may lead to unexpected changes in the skin microbiome, and more specifically it may lead to increases in the proportion of normal commensal non-pathogenic species and reductions in the proportion of potentially pathogenic, pathogenic, or disease causing organisms.

EXAMPLES

Example 1

Initial Culturing of *N. eutropha*

[0580] A soil-derived culture enriched in various ammonia oxidizing bacteria was applied to the skin of an adult male subject as described in WO/2003/057380. The period of growth on the human body selected for a strain with the capacity to colonize human skin for an extended period of time. After several months, the strain was re-isolated from the skin of the individual and cultured in laboratory conditions for a sustained period as described in the subsequent examples. While not wishing to be bound by theory, it is believed that the sustained laboratory culture selected for

new mutations improving the strain properties, e.g., improved tolerance for high-ammonia conditions.

Example 2

Growing and Monitoring D23 or Mixtures of Strains that Comprise D23 Culture Conditions

[0581] D23 can be grown in batches or by continuous cultivation in a bioreactor. Batch preparation uses the medium of Table 3.

TABLE 3

Growth Medium for Batch culturing:		
	Weight/Volume (in~1.5 L)	Final Concentration (in~1.5 L)
(NH ₄) ₂ SO ₄ (MW 132.14)	4.95 g	50 mM NH ₄ ⁺
KH ₂ PO ₄ (MW 136.1)	0.616 g	3.0 mM
1M MgSO ₄	1.137 ml	0.76 mM
1M CaCl ₂	0.3 ml	0.2 mM
30 mM FeCl ₃ /50 mM EDTA	0.5 ml	10 μM/16.7 μM
50 mM CuSO ₄	30 μl	1.0 μM
Add 1400 ml ddH ₂ O to flask. Autoclave. Store at room temperature. After autoclaving add:		
Phosphate Buffer	100 ml	32 mM KH ₂ PO ₄ / 2.7 mM NaH ₂ PO ₄ •H ₂ O
5% Na ₂ CO ₃	12 ml	0.04%

[0582] The medium of Table 3 is inoculated with ~15 ml of a 3 day old culture of D23 (i.e. 1% volume). The cultures are incubated in the dark at 30° C. by shaking at 200 rpm.

[0583] Often, a *N. eutropha* D23 mixed culture is grown on complete *N. europaea* media. The culture medium is described below, and additional details on culturing ammonia-oxidizing bacteria are available on the World Wide Web at nitrificationnetwork.org/Nerecipe.php, Ensign et al., 1993, and Stein & Arp, 1998.

Step 1.

[0584] Add 900 ml of deionized water to a 2-liter Erlenmeyer flask.

[0585] Add in sequence:

[0586] 3.3 g (NH₄⁺)₂SO₄ (50 mM);

[0587] 0.41 g KH₂PO₄

[0588] 0.75 ml 1 M MgSO₄ stock solution

[0589] 0.2 ml 1 M CaCl₂ stock solution

[0590] 0.33 ml 30 mM FeSO₄/50 mM EDTA stock solution

[0591] 0.01 ml 50 mM CuSO₄ stock solution

[0592] Sterilize the solution by autoclaving.

Step 2.

[0593] Add 400 ml of deionized water to a beaker. Add:

[0594] 27.22 g KH₂PO₄

[0595] 2.4 g NaH₂PO₄

[0596] Adjust the pH to 8.0 with 10 N NaOH, and bring the final volume to 500 ml with deionized water.

[0597] Sterilize 100 ml fractions of the solution by autoclaving in 250-500 ml Erlenmeyer flasks.

Step 3

[0598] Prepare 500 ml of 5% (w/v) Na₂CO₃ (anhydrous)

[0599] Sterilize the solution by autoclaving.

Step 4

[0600] Add 1×100 ml aliquot of solution prepared in Step 2 to the flask prepared in Step 1.

Step 5

[0601] Add 8 ml of the solution prepared in Step 3 to the flask prepared in Step 1.

[0602] The D23 can also be cultured continuously in a bioreactor. Table 4 describes the appropriate media.

TABLE 4

Growth Medium for continuous culture:		
	Batch medium Weight/Volume (1 L) (Final concentration)	Feeding solution Weight/Volume (1 L) (Final concentration)
(NH ₄) ₂ SO ₄ (MW 132.14)	3.3 g (50 mM NH ₄ ⁺)	13.2 g (200 mM NH ₄ ⁺)
KH ₂ PO ₄ (MW 136.1)	1.23 g (9.0 mM)	0.41 g (3.0 mM)
1M MgSO ₄	0.758 ml (0.76 mM)	0.758 ml (0.76 mM)
1M CaCl ₂	0.2 ml (0.2 mM)	0.2 ml (0.2 mM)
30 mM FeCl ₃ /50 mM EDTA	0.333 ml (10 μM/16.7 μM)	0.333 ml (10 μM/16.7 μM)
50 mM CuSO ₄	20 μl (1.0 μM)	20 μl (1.0 μM)
ddH ₂ O	1000 ml	1000 ml

Autoclave each solution and store at room temperature.

[0603] The batch media, in a bioreactor vessel, is inoculated with ~10 ml of a 3 day old *N. eutropha* D23 culture (i.e. 1% volume). The pH is adjusted to 7.6 using 7.5% Na₂CO₃. The bioreactor is run in batch mode with below parameters: pH: 7.6 (lower limit: 7.45 & upper limit: 7.8), Temperature: 28° C. (lower limit: 25° C. & upper limit: 32° C.), DO (dissolved oxygen): 45% (lower limit: 10%, upper limit: 100%), Stirrer: 550 rpm.

[0604] The OD600 nm of the culture in the bioreactor reaches 0.15 to 0.18 in 3-4 days. At this point, the culture will consume most of the 50 mM NH₄⁺ present in the AOB growth media, and a user should start feeding the bioreactor with feeding solution at 0.59 ml/min (~10%). The outflow pump should also be turned on at 0.59 ml/min (~10%). The OD600 nm of the bioreactor reaches 0.5-0.6 in 1-2 days of continuous culture. The culture in the bioreactor is tested for heterotrophic contaminants by plating 1 ml of the bioreactor outflow on an LB plate.

Monitoring Growth of *N. eutropha* D23

[0605] Growth of *N. eutropha* D23 cells is monitored by measuring the OD600 nm of the culture. Typical growth in a batch culture as measured by OD600 nm is between 0.06 to 0.08.

[0606] The AOB growth medium contains NH₄⁺ that is stoichiometrically converted to NO₂⁻ by *N. eutropha* D23. Another way to monitor the growth of *N. eutropha* is to follow the release of nitrite (NO₂⁻) in the growth medium. NO₂⁻ concentration is determined with Griess reagents, sulfanilamide and N-naphthylethylenediamine (also called NNEQ). Briefly, sulfanilamide and NNEQ are added to a

sample and to known concentrations of sodium nitrite that make up a standard curve. Samples are incubated in the dark for 30 minutes. The absorbance is read at 540 nm.

[0607] Another way to follow nitrite production is by using a spectrophotometer by monitoring the optical density (OD) difference between 352 nm and 400 nm. The nitrite concentration is determined using a millimolar extinction coefficient of 0.0225 mM^{-1} . This assay can be performed directly by sampling the medium with the cells.

$$\text{NO}_2^- \text{ concentration (mM)} = (\text{OD}_{352} - \text{OD}_{400}) / 0.0225$$

[0608] The growth of a mixed culture comprising D23 was monitored by measuring optical density at 600 nm (OD600 nm) and by measuring Nitrite (NO_2^-), and the growth rate is shown in FIGS. 1 and 2. FIG. 1 shows that the optical density at a 600 nm wavelength plateaus slightly below 0.1, after 3 to 4 days. FIG. 2A shows that the amount of nitrite produced plateaus slightly below 25 mM after 3 to 4 days. NO_2^- concentrations in the cultures were determined colorimetrically by the Griess reagent (Hageman & Hucklesby, 1971), and is used as a second indicator for the growth rates and growth phases since the accumulation of NO_2^- is consistently proportional to the increase in cell mass during growth.

[0609] In FIG. 2B-I, increasing densities of D23 harvested from continuous culture were suspended in medium supplemented with 50 mM NH_4^+ and grown shaking at 30° C. for 48 hours. Nitrite production was measured in supernatant samples using the Griess assay at the time points indicated. Results shown are mean values $\pm$ SD from three independent experiments.

[0610] In FIG. 2B-II, Nitrite production by *N. eutropha* D23 in vitro is shown. Increasing densities of D23 were suspended in mineral salt medium supplemented with 50 mM NH_4^+ and grown shaking at 30° C. for 24 hr. Nitrite production was measured in supernatant samples using the Griess assay at the time points indicated.

[0611] Storage Conditions

[0612] *N. eutropha* suspensions obtained from the continuous culture system showed remarkable stability upon storage at 4° C. for several months, as indicated by the highly consistent nitrite concentrations generated upon sub-culture under batch growth conditions. Protocols for storing and recovering *N. eutropha* are set out below.

[0613] Obtain 500 ml of a *N. eutropha* D23 culture grown to late-exponential phase ($\text{OD}_{600} = 0.5\text{--}0.6$ in continuous culture). Centrifuge at $10,000\times g$ for 15 min at 20° C. Remove supernatant and resuspend the pellet in 50 ml of AOB storage buffer. Spin as above. Remove supernatant and resuspend thoroughly in a total of 50 ml storage buffer. This would be the $10\times$ AOB stock. Store upright at 4° C. in 50 ml polypropylene tubes.

[0614] AOB Storage Buffer (for AOB storage at 4° C.): 50 mM Na_2HPO_4 - 2 mM MgCl_2 (pH 7.6) can be made as follows.

[0615] In 1 Liter ddH₂O: Na_2HPO_4 - 7.098 g

[0616] MgCl_2 - 0.1904 g

[0617] Adjust pH to 7.6. Filter-sterilize.

[0618] *N. eutropha* may be cryopreserved as follows. Transfer 1.25 ml of *N. eutropha* D23 mid-log culture to a 2 ml cryotube and 0.75 ml of sterile 80% glycerol. Shake tubes gently, incubate at room temperature for 15 min to enable uptake of the cryoprotective agents by the cells. Then, put tubes directly in a -80°C freezer for freezing and storage.

For resuscitation of cultures, thaw frozen stocks on ice for 10-20 minutes. Centrifuge, at $8,000\times g$ for 3 minutes at 4° C. Discard supernatant and wash the pellet by suspending it in 2 ml AOB medium followed by another centrifugation at $8,000\times g$ for 3 minutes at 4° C. to reduce potential toxicity of the cryoprotective agents in subsequent growth experiments. Discard the supernatant and resuspend the pellet in 2 ml of AOB medium, inoculate into 50 ml of AOB medium containing 50 mM NH_4^+ , and incubate in dark at 30° C. by shaking at 200 rpm.

[0619] In FIG. 2C, stability upon storage at 4° C. was studied. *N. eutropha* D23 previously harvested from continuous culture and stored at 4° C. was inoculated at 10^9 CFU/ml in mineral salt medium supplemented with 50 mM NH_4^+ and grown shaking at 30° C. Nitrite production was determined at 24 and 48 hours post-incubation (left and right panel, respectively). Data shown are representative of a D23 suspension sampled repeatedly over a 6-month period.

Example 3

Creation of an Axenic D23 Culture

[0620] To isolate *N. eutropha* D23 in pure culture, four types of media (described below) were made, autoclaved and poured in plates. Sterile nylon membranes were placed on the plates.

N. europaea media+1.2% R₂A agar

N. europaea media+1.2% agar

N. europaea media+1.2% agarose

N. europaea media+1.2% agarose+0.3 g/L pyruvate

[0621] 3 day old *N. eutropha* D23 culture was streaked onto the nylon membranes and the plates were incubated at 30° C. The plates were monitored daily for growth of red colored *N. eutropha* cells. Nylon membranes were transferred to fresh plates once a week.

[0622] Reddish colored colonies appeared on plates with R₂A agar or agar by end of 1 week. Single colonies were picked from plates with R₂A agar and grown in *N. europaea* media. The cultures grew well in 6 days to 0.08 OD600 nm. Heterotrophic colonies appeared when the culture was plated on LB-Agar plates.

[0623] Reddish colored colonies on plates with R₂A agar, agar, agarose, or agarose+pyruvate appeared by end of 2 weeks. Single colonies were picked from plates with agar or agarose and grown in *N. europaea* media. The cultures grew well in 6-8 days to 0.08 OD600 nm. Heterotrophic colonies appeared when the culture was plated on LB-Agar plates.

[0624] Bright reddish colonies on plates with R₂A agar, agar, agarose, or agarose+pyruvate appeared by end of 4 weeks. Single colonies were picked from plates with agarose and grown in *N. europaea* media. The cultures grew well in 6-8 days to 0.08 OD600 nm. White colonies appeared when the culture was plated on LB-Agar plates.

[0625] Contaminating bacteria (e.g., non-*N. eutropha* bacteria present in the mixed culture) were identified by culturing, amplifying 16S rRNA by PCR, and sequencing of the PCR products. Contaminants were identified as *Microbacterium* sp. and *Alcaligenaceae* bacterium.

[0626] To create an axenic culture of D23 (i.e., free of contaminating bacteria) serial dilution was used. Eight single colonies (designated A-H) were picked, and each was placed into a 10 ml culture of *N. europaea* medium. For each

culture, five sequential 1:10 dilutions were created. For each culture A-H, growth was observed in the two or three most concentrated of the dilutions.

[0627] A second serial dilution was carried out. 50 ml of media was inoculated with approximately 2×10^8 *N. eutropha* cells, and sequential dilutions of 1:50 were made, such that after the fifth dilution, a flask was expected to have approximately one cell. Flasks that exhibited bacterial growth were plated on LB-agar to assay for contaminating bacteria, and no contaminating bacteria were observed. In addition, no contaminating gram positive cells were observed under the microscope.

[0628] Accordingly, the serial dilution process yielded an axenic or substantially axenic culture of *N. eutropha*.

Example 4

Sequencing of the D23 Genome

[0629] Strain D23 was obtained as described in Example 1, and was made axenic as described in Example 3.

[0630] A 10 ml aliquot the bacterial sample was inoculated into approximately 1 L of *N. europaea* growth medium described in Example 2. The culture grew well to optical density of 0.08 at 600 nm in a batch culture in 3 days.

[0631] Total DNA of the culture was prepared and sequenced using Illumina® technology and/or SMRT® DNA Sequencing System technology, Pacific Biosciences. The strain was identified as *Nitrosomonas eutropha* and was designated D23.

[0632] The genome sequence of D23 was compared to that of *N. eutropha* C91, which is believed to be the only other sequenced strain of *N. eutropha*.

[0633] The length of the D23 chromosome is 2,537,572 base pairs, which is shorter than the 2,661,057 base pair chromosome of *N. eutropha* strain C91 chromosome. Based on the 16S-23S operon, strain D23 has 99.46% identity to C91 and 95.38% identity to *N. europaea*. DNA sequencing of *N. eutropha* D23 indicated that this strain lacks plasmids. This contrasts with the sequence of strain C91, which has two plasmids.

[0634] Protein-encoding regions and RNA-encoding sequences were identified by sequence analysis. Supplementary Table 1 is a table of annotations that lists the positions of 2,777 genes in the D23 genome (SEQ ID NO: 1).

[0635] On the level of individual genes, several genes are present in D23 that are absent in C91. These genes are summarized in FIGS. 6-8. FIG. 6 is a table displaying unique D23 genes with an assigned ORF number and a function based on sequence analysis, or a hypothetical gene above 200 base pairs in length. There are 162 genes in this category. FIG. 7 is a table displaying unique D23 genes below 200 base pairs that have an assigned ORF number. There are 164 of these genes. FIG. 8 is a table displaying unique D23 genes with no assigned ORF number. There are 219 of these genes (of which 180 are below 200 bp in length).

[0636] Strain D23 also lacks a number of genes that are present (or lack close homologs) in strain C91. These genes are sometimes referred to as unique C91 genes. These genes include the about 300 genes listed in FIG. 9.

[0637] D23 contains several ammonia metabolism genes that differ from their homologs in C91. Certain of these genes are enumerated in Table 1 of the Detailed Description. Sequence alignments were performed between the D23

proteins and their homologs in strain C91. The sequence alignments are shown in FIGS. 10-16 and sequence differences between the two strains are shown in Table 2 of the Detailed Description.

[0638] The sequence comparisons revealed the percent sequence identities between the C91 and D23 homologs of each protein. More specifically, FIG. 10 is an alignment between AmoA1 and AmoA2 of strains C91 and D23. Each protein is identical at 273/276 residues, and so each is about 98.9% identical between strains. FIG. 11 is an alignment between AmoB1 and AmoB2 of strains C91 and D23. Both proteins are identical at 419/421 positions, and so are about 99.5% identical between strains. FIG. 12 is an alignment between AmoC1 and AmoC2 of strains C91 and D23. Both proteins are identical throughout. FIG. 13 is an alignment between AmoC3 of strains C91 and D23. This protein is identical at 272/274 positions, and so are about 99.3% identical between strains.

[0639] As to the Hao proteins, FIG. 14 (A and B) is an alignment between Hao1, Hao2, and Hao3 of strains C91 and D23. Hao1 is identical at 567/570 positions, and so each is about 99.5% identical between strains. Hao2 and Hao3 are each identical at 568/570 positions, and so are about 99.6% identical between strains.

[0640] Turning now to cytochrome c554 proteins, FIG. 15 is an alignment between CycA1, CycA2, and CycA3 of strains C91 and D23. CycA1 is identical at 233/235 positions, and so is about 99.1% identical between strains. CycA2 and CycA3 are each identical at 234/235 positions, and so each is about 99.6% identical between strains.

[0641] As to the cytochrome c_M552 proteins, FIG. 16 is an alignment between CycB1 and CycB2 of strains C91 and D23. CycB1 is identical at 232/239 positions, and so is about 97.1% identical between strains. CycB2 is identical at 236/239 positions, and so is about 98.7% identical between strains. Here, the length of the protein is considered 239 amino acids because that is its length in strain D23.

[0642] Alignment of the nucleic acid sequences of Table 1 shows the percent identity between homologs in C91 and D23. The amoA1 genes are about 98.8% identical (i.e., at 821/831 positions), the amoA2 genes are about 98.8% identical (i.e., at 821/831 positions), the amoB1 genes are about 99.1% identical (i.e., at 1255/1266 positions), the amoB2 genes are about 99.1% identical (i.e., at 1254/1266 positions), the amoC1 genes are about 99.8% identical (i.e., at 814/816 positions), the amoC2 genes are about 99.8% identical (i.e., at 814/816 positions), and the amoC3 genes are about 98.9% identical (i.e., at 816/825 positions). The hao1 genes are about 99.0% identical (i.e., at 1696/1713 positions), the hao2 genes are about 99.4% identical (i.e., at 1702/1713 positions), and the hao3 genes are about 99.2% identical (i.e., at 1700/1713 positions). Of the cytochrome c554 genes, the cycA1 genes are about 98.0% identical (i.e., at 694/708 positions), the cycA2 genes are about 98.7% identical (i.e., at 699/708 positions), and the cycA3 genes are about 99.3% identical (i.e., at 703/708 positions). Of the cytochrome c_M552 genes, the cycB1 genes are about 96.7% identical (i.e., at 696/720 positions) and the cycB2 genes are about 97.1% identical (i.e., at 702/723 positions).

Example 5

Competitive Growth Studies

[0643] A study was designed to determine whether *N. eutropha* strain D23 could inhibit the growth of undesirable

bacteria such as *Pseudomonas aeruginosa* (*P. aeruginosa* or PA), *Staphylococcus aureus* (*S. aureus* or SA), *Streptococcus pyogenes* (*S. pyogenes* or SP), *Acinetobacter baumannii* (*A. baumannii* or AB), and *Propionibacterium acnes*, all of which are important pathogenic agents frequently isolated from either one or both of infected skin and wound sites. This protocol may also be used to test other *N. eutropha* strains for the ability to inhibit the growth of undesirable bacteria.

[0644] Briefly, a suitable protocol can comprise the following steps. At $t=0$, a culture is inoculated with *N. eutropha*, and then the *N. eutropha* is incubated for 24 hours. Culture characteristics (e.g., pH and nitrite levels) are assayed. At $t=24$ hours, the undesirable bacterium is added to the culture. Immediately upon addition, samples are obtained for determining CFU/ml of the undesirable bacteria and optionally CFU/ml of *N. eutropha*, pH, and nitrite levels. Incubation is allowed to proceed for an additional 24 hours. At subsequent timepoints, e.g., $t=30$ and $t=48$, one can take the same measurements as at $t=24$. To determine CFU/ml, one can plate neat/-1/-2/-3/-4/-5 (or higher) to obtain accurate counts.

[0645] A more detailed protocol is set out below.

Day 1

[0646] 1. Mix the 10×AOB stock suspension stored at 4° C. by inverting several times until a homogenous suspension is obtained.

[0647] 2. Aliquot 10 ml of the suspension in 8×1.5 ml polypropylene tubes.

[0648] 3. Centrifuge at 17,000×g for 3 min at room temperature.

suspension in 16×1.5 ml polypropylene tubes. Obtain pellets as above and resuspend in a total of 26 ml complete AOB medium/0.5× Phosphate Buffer in a 50 ml polypropylene tube.

[0652] 7. Pipet 5 ml of the 10×AOB suspension in each of five 50 ml polypropylene tubes (Tube 4-8).

[0653] 8. Also, prepare two tubes with 10× Heat-killed AOB suspensions in either complete AOB medium (Tube 3) or complete AOB medium/0.5× Phosphate Buffer (Tube 9). Aliquot 10 ml of the Heat-killed suspension stored at 4° C. in 8×1.5 ml polypropylene tubes. Centrifuge at 17,000×g for 3 min at room temperature and remove supernatant, as described above for live AOB. Resuspend four pellets in a total of 5 ml complete AOB medium in one 50 ml polypropylene tube (Tube 3) and the remaining four pellets in a total of 5 ml complete AOB medium/0.5× Phosphate Buffer in a second 50 ml polypropylene tube (Tube 9).

[0654] 9. Add 141 μ l of 1 M ammonium sulfate to obtain 25 mM final concentration (Tube 1, 3, 4, 5, 9). Add an equal volume of dH₂O to corresponding control tubes (Tube 2, 6, 7).

[0655] 10. To Tube 8, add 141 μ l of fresh 1 M NaNO₂.

[0656] 11. Swirl all tubes gently, but thoroughly, to mix.

[0657] 12. Immediately after mixing each suspension, remove 0.5 ml from each tube and centrifuge all samples at 17,000×g, 3 min, RT. Transfer supernatants into fresh tubes after completing step 13, and measure both pH and nitrite levels using Griess Reagent to obtain TO values.

[0658] 13. Incubate all 50 ml tubes at 30° C. with mixing on an orbital shaker at 150 rpm (upright position) for 24 hr.

TABLE 5

T0							
SAMPLE	Tube	10x AOB (ml)	10x Killed AOB (ml)	1M (NH ₄) ₂ SO ₄ (μl)	H ₂ O (μl)	1M NaNO ₂ (μl)	T24 SA/PA in saline (ml)
Complete AOB medium							
10x AOB + NH ₄ ⁺	1	5	—	141	—	—	0.5
10x AOB	2	5	—	—	141	—	0.5
10x Killed AOB + NH ₄ ⁺	3	—	5	141	—	—	0.5
Complete AOB medium/0.5x Phosphate Buffer							
10x AOB + NH ₄ ⁺	4	5	—	141	—	—	0.5
10x AOB + NH ₄ ⁺	5	5	—	141	—	—	0.5
10x AOB	6	5	—	—	141	—	0.5
10x AOB	7	5	—	—	141	—	0.5
10x AOB + NaNO ₂	8	5	—	—	—	141	0.5
10x Killed AOB + NH ₄ ⁺	9	—	5	141	—	—	0.5

[0649] 4. Remove supernatant and any residual buffer from each pellet and resuspend all pellets thoroughly into a total of 10 ml complete AOB medium in a 50 ml polypropylene tube.

[0650] 5. Pipet 5 ml of 10×AOB suspension in each of two 50 ml polypropylene tubes (Tube 1-2).

[0651] 6. Prepare five additional tubes (Tube 4-8) containing 10×AOB suspensions in complete AOB medium/0.5× Phosphate Buffer. Aliquot 26 ml of the 10×AOB stock

Day 2

[0659] 14. At 24 hr, prepare SA, PA, SP or AB inocula to add to the suspensions.

[0660] 15. From an overnight (20-24 hr) SA or PA culture grown on Tryptic Soy Agar (TSA), or a SP or AB culture prepared on Brain Heart Infusion (BHI) Agar, prepare bacterial suspension in Tryptic Soy Broth (TSB) or BHI broth (BHIB) at $\sim 2 \times 10^8$ CFU/ml.

- [0661] 16. Pipet 50 μ l of the SA/PA/SP/AB suspension in 9.95 ml saline to obtain $\sim 10^6$ CFU/ml. Keep on ice, as needed.
- [0662] 17. Vortex SA/PA/SP/AB suspension and add 0.5 ml to Tube 1-9.
- [0663] 18. Swirl all tubes gently, but thoroughly, to mix.
- [0664] 19. Immediately after mixing each suspension, transfer 100 μ l from each tube into 0.9 ml TSB or BHIB (10^{-1} dilution) to neutralize samples for CFU determination. In addition, remove 0.5 ml from each tube and centrifuge at $17,000\times g$, 3 min, RT. Recover supernatants in fresh tubes after completing Step 20 and measure both pH and nitrite levels using Griess Reagent after Step 21 to obtain T24 values.
- [0665] 20. Incubate all 50 ml tubes at 30°C . with mixing on an orbital shaker (150 rpm) for an additional 24 hr.
- [0666] 21. Dilute T24 samples further in TSB or BHIB and plate $-2/-3/-4$ dilutions on TSA or BHI agar. Incubate plates at 37°C . for 24 hr to obtain SA, PA, SP, or AB viable counts.
- [0667] 22. At 6 and 24 hr post-mixing of SA/PA/SP/AB with AOB, vortex tubes and pipet 100 μ l samples into 0.9 ml TSB. Dilute further in TSB or BHIB and plate neat through -5 dilutions on TSA or BHI agar. At each time point, also remove 0.5 ml from each tube and measure both pH and nitrite levels in each supernatant sample, as described above.
- [0668] 23. Incubate TSA or BHI agar plates at 37°C . for 24 hr to obtain T30 (6 hr) and T48 (24 hr) viable counts.
- [0669] 24. Count CFU to determine % killing rates for each time point

Griess Reagent Assay for Nitrite Quantification

- [0670] 1. Use the 0.5 ml supernatant samples obtained for pH determination at 0, 2, 6, and 24 hr.
2. Serially dilute 56 μ l of the supernatant in 0.5 ml dH₂O to obtain 10-100- and 1000-fold dilutions, as needed. For TO samples, use $1/10$ for Tube 1-6, 8, 9, and $1/1000$ for Tube 7. For T24/T30/T48 samples, use $1/10$, $1/100$, $1/1000$ for all tubes.
3. To prepare sodium nitrite standards, dilute 10 μ l of a fresh 1 M stock in 990 μ l complete AOB medium-10% saline to obtain a 10 mM solution.
4. Dilute 10 μ l of the 10 mM stock in 990 μ l dH₂O to obtain a 100 μ M working solution.
5. Prepare standards in dH₂O as shown below. Run standards only with TO samples.

TABLE 6

100 μ M sodium nitrite (μ l)	dH ₂ O (μ l)	Nitrite conc (μ M)	A _{540nm} (indicative values)
0 (blank)	500	0	0
62.5	437.5	12.5	0.307
125	375	25	0.607
250	250	50	1.164
500	0	100	2.35

6. To each 0.5 ml sample (or sodium nitrite standard), add 0.25 ml each of Reagent A (58 mM sulfanilamide in 1.5 N HCl) and Reagent B (0.77 mM n-(1-naphthyl) ethylene diamine-2HCl in H₂O (light-sensitive; store in dark)).

7. Mix and let stand at room temperature for 30 min in the dark (or cover with foil). The color should change to a vivid pink/violet.

8. Read absorbance at 540 nm and determine nitrite concentrations from standard curve.

[0671] This protocol was used to test *N. eutropha* D23's ability to inhibit the growth of *P. aeruginosa* (PA), *S. aureus* (SA), *S. pyogenes* (SP), *A. baumannii* (AB), or *P. acnes*. The results of this experiment are shown in FIGS. 3A, 3B, and 3C.

[0672] The left panel of FIG. 3A plots CFU/ml of PA versus time, when PA is co-cultured with live *N. eutropha* and ammonium (squares), live *N. eutropha* without ammonium (circles), killed *N. eutropha* and ammonium (triangles), or live *N. eutropha* with NaNO₂ (inverted triangles). The right panel of FIG. 3A plots CFU/ml of SA versus time, under the same conditions. The left panel of FIG. 3B plots CFU/ml of SP versus time, under the same conditions. The right panel of FIG. 3B plots CFU/ml AB versus time, under the same conditions. FIG. 3C plots CFU/ml of *P. acnes* versus time, when *P. acnes* is co-cultured with live *N. eutropha* and ammonium (squares), live *N. eutropha* without ammonium (circles), killed *N. eutropha* and ammonium (triangles), or live *N. eutropha* with NaNO₂ (inverted triangles). In all cases, live *N. eutropha* with ammonium results in declining numbers of PA, SA, SP, AB, or *P. acnes* whereas the other culture conditions allow the undesirable bacteria to grow. Without being bound by theory, these experiments suggest that nitrite generation from ammonia concurrently with medium acidification by D23 led to strong antibacterial effects, e.g., an approximately 100-fold reduction in viable counts of methicillin-resistant *Staphylococcus aureus*, *Pseudomonas aeruginosa*, *Streptococcus pyogenes*, *Acinetobacter baumannii*, or *P. acnes*. By contrast, control co-cultures of pathogenic bacteria either with heat-killed D23 supplemented with ammonia, or with live D23 without ammonia, did not produce comparable antibacterial effects. The control comprising live *N. eutropha* culture without ammonium is consistent with the model that *N. eutropha*'s ammonia oxidation activity contributes to its antibacterial effects. The control comprising killed *N. eutropha* and ammonium indicates that some biological activity of the *N. eutropha* (e.g., its ammonia oxidation activity) contributes to antibacterial activity. The control comprising live *N. eutropha* with NaNO₂ indicates that comparable nitrite levels at neutral pH (versus low pH when the bacteria use ammonia) do not have a strong antimicrobial effect, and is consistent with the model that *N. eutropha*'s oxidation of ammonia, rather than nitrite alone, contributes to the antibacterial activity.

[0673] The top panel of FIG. 4A plots the NO₂⁻ concentration over time in the co-cultures described in the paragraph above. NO₂⁻ concentration is an indication of the rate of NH₃ metabolism in the cultures. As above, PA is co-cultured with *N. eutropha* and ammonium (squares), *N. eutropha* without ammonium (circles), or killed *N. eutropha* and ammonium (triangles). Live *N. eutropha* with ammonium produces dramatically higher NO₂⁻ levels than the two control cultures, indicating that the live *N. eutropha* converts ammonium into NO₂⁻ under the culture conditions.

[0674] The bottom panel of FIG. 4A plots pH over time in the same co-culturing conditions. pH indicates the metabolic activity of the *N. eutropha* because the conversion of ammonia to nitrite produces hydrogen ions. PA is co-cultured with

N. eutropha and ammonium (squares), *N. eutropha* without ammonium (circles), killed *N. eutropha* and ammonium (triangles), or live *N. eutropha* with NaNO₂ (inverted triangles). Live *N. eutropha* with ammonium acidifies the medium, in contrast to the three control cultures, indicating that the live *N. eutropha* metabolizes ammonium under the culture conditions.

[0675] The top panels of FIG. 4B plot the NO₂⁻ concentration over time in the co-cultures described above. NO₂⁻ concentration is an indication of the rate of NH₃ metabolism in the cultures. As above, *S. pyogenes* (SP) and *A. baumannii* (AB) are co-cultured with *N. eutropha* and ammonium (squares), *N. eutropha* without ammonium (circles), or killed *N. eutropha* and ammonium (triangles). Live *N. eutropha* with ammonium produces dramatically higher NO₂ levels than the two control cultures, indicating that the live *N. eutropha* converts ammonium into NO₂⁻ under the culture conditions.

[0676] The bottom panels of FIG. 4B plot pH over time in the same co-culturing conditions. pH indicates the metabolic activity of the *N. eutropha* because the conversion of ammonia to nitrite produces hydrogen ions. SP and AB are co-cultured with *N. eutropha* and ammonium (squares), *N. eutropha* without ammonium (circles), killed *N. eutropha* and ammonium (triangles), or live *N. eutropha* with NaNO₂ (inverted triangles). Live *N. eutropha* with ammonium acidifies the medium, in contrast to the three control cultures, indicating that the live *N. eutropha* metabolizes ammonium under the culture conditions.

[0677] FIG. 4E shows an alternative visualization the data of FIGS. 4A and 4B.

[0678] The capacity of *Nitrosomonas eutropha* D23 to inhibit proliferation of pathogenic bacteria due to nitrite production concurrent with acidification (acidified nitrite) was assessed by testing the survival of 5 strains of pathogenic bacteria in co-culture studies with D23 in vitro. The five strains of pathogenic bacteria included *Propionibacterium acnes*, *Streptococcus pyogenes*, methicillin-resistant *Staphylococcus aureus* (MRSA), *Pseudomonas aeruginosa*, and multidrug-resistant *Acinetobacter baumannii*. Incubation of *N. eutropha* D23 (10¹⁰ cells/ml) in the presence of ammonium led to nitrite concentrations of 10 mM or higher and acidification to pH 6 or lower (FIG. 4B). The combination of increased nitrite concentration and lowering of pH led to bactericidal or bacteriostatic effects and a marked reduction (up to 965-fold) in viable counts of the pathogenic bacterial species tested. The results of these studies are summarized in FIG. 4D and Table 7, below. In contrast to the D23 co-cultures incubated in the presence of ammonium, control co-cultures of the five pathogenic agents with D23 without ammonium, or with heat-killed D23 (B244) supplemented with ammonium, did not lead to any inhibitory or antimicrobial effects.

TABLE 7

Effect of <i>N. eutropha</i> D23 (D23) on relative survival of pathogenic bacteria in vitro			
Pathogen Tested	Relative Survival (Fold Change)		
	AOB + NH ₃	AOB - NH ₃	Heat-Killed AOB + NH ₃
<i>Propionibacterium acnes</i> ATCC 6919	-114	-19,067	-1.05

TABLE 7-continued

Effect of <i>N. eutropha</i> D23 (D23) on relative survival of pathogenic bacteria in vitro			
Pathogen Tested	Relative Survival (Fold Change)		
	AOB + NH ₃	AOB - NH ₃	Heat-Killed AOB + NH ₃
<i>Staphylococcus aureus</i> (MRSA) ATCC BAA-1717	-117.6	8.2	2.03
<i>Pseudomonas aeruginosa</i> ATCC 15442	-84.3	2.65	379
<i>Streptococcus pyogenes</i> ATCC 19615	-965	-2.88	-3.81
<i>Acinetobacter baumannii</i> (MDR) ATCC BAA-1605	-5.43	92.4	89.8

Example 6

Wound Healing

[0679] The effect of *Nitrosomonas eutropha* D23 (sometimes also called B244) on wound closure in diabetic mice was evaluated in two separate studies. In Study 1, db/db mice (8 mice/group) were pre-treated by body immersion daily for one week with 3 concentrations of D23 (10⁷, 10⁸ or 10⁹ cells/ml) supplemented with ammonium chloride, or with vehicle control suspension only. Subsequently, full-thickness wounds generated on the back of each animal were treated topically once daily for 14 days with vehicle alone or equal volumes of 3 concentrations of D23 (10⁷, 10⁸ or 10⁹ cells/ml) in PBS supplemented ammonium chloride. Of the three D23-treated groups, the group receiving the highest dose showed significant improvement in wound closure from day 5 to day 15, with the most pronounced improvement of 83% observed on day 9 post-wounding. The median time to 50% wound closure was significantly reduced (P<0.05) for the animals treated with 10⁹ cells/ml of D23, as compared to the animal group receiving vehicle treatment alone.

[0680] Initial histopathology analyses of wound tissue samples collected on Day 15 upon study completion did not reveal any gross differences between vehicle- and D23-treated animals. Subsequently, a more in-depth examination of the tissue sections was performed according to the scoring system and parameters adapted and modified from Altavilla, et al (2001). This analysis suggested a trend of increased levels of angiogenesis and maturity of granulation tissue with decreased levels of dermal inflammation in animals treated with 10⁹ cells/ml of D23 versus the vehicle control group, which was consistent with the observed improvement in wound healing rates of the D23-treated animals.

[0681] *N. eutropha* strain D23 was tested for its ability to accelerate wound healing in a diabetic mouse model, using C57BLKS/J Iar-+Lepr^{db}/+Lepr^{db} male mice (non-GLP). A detailed protocol is set out below.

Day -6 to Day 1: Whole-Body Immersion Pre-Treatment of Mice with Test Organism

[0682] 1. Mix the 10×D23 stock suspension stored at 4° C. by inverting several times until a homogenous suspension is obtained.

[0683] 2. Pipet 2×29 ml of the 10× stock suspension into two 50 ml polypropylene centrifuge tubes.

[0684] 3. Centrifuge at 8,000×g for 15 min at 20° C.

- [0685] 4. Remove supernatant and any residual buffer from the pellets and resuspend the two pellets gently but thoroughly into a total of 58 ml room-temperature Phosphate Buffered Saline, pH 7.4 (PBS). This is the 10×D23 (Test Organism) suspension to use for the following steps.
- [0686] 5. Prepare 500 ml baths containing the Test Organism at 1×, 0.1× and 0.01× strength in pre-warmed PBS at 30° C. supplemented with 2 mM NH₄Cl, or a Vehicle control bath, as shown below. Prepare and use one bath at a time from the 10×D23 suspension kept at room temperature before continuing with the next bath. This will prevent keeping the Test Substance at 30° C. for long time periods without ammonium. To prevent contamination of the Vehicle control group with the Test Substance, begin with the Vehicle control group before proceeding with the D23 baths.
- [0687] 6. Immerse each group of mice in corresponding baths for 60 sec daily for seven days.
- [0688] 7. Use a fresh 500 ml baths for each daily immersion into the Test Organism or Vehicle control.

TABLE 8

BATH/ GROUP	10× D23 (room temp.) (ml)	PBS pH 7.4 (ml)	1M NH ₄ /Cl (ml)	CFU/ ml
Vehicle (control)	—	500	1.0	0
1× D23	50	450	1.0	109
0.1× D23	5	495	1.0	108
0.01× D23	0.5	499	1.0	107

Day 1: Wounding of Mice by Skin Puncture

- [0689] 1. Generate skin wounds on the back of each mouse by skin puncture after shaving of the back and shoulders.
- [0690] 2. House each mouse separately for the remainder of the study.

Day 1 to Day 15: Topical Treatment of Skin Wounds with Test Organism

- [0691] 1. Mix the 10×D23 stock suspension stored at 4° C. by inverting several times until a homogenous suspension is obtained.
- [0692] 2. Pipet 1 ml of the 10× stock suspension into a 1.5 ml polypropylene tube.
- [0693] 3. Centrifuge at 17,000×g for 3 min at room temperature.
- [0694] 4. Remove supernatant and any residual buffer from the pellet and resuspend pellet gently but thoroughly into a total of 10 ml pre-warmed Phosphate Buffered Saline, pH 7.4 (PBS) at 30° C. This is the 1×D23 (Test Organism) suspension to use for the following steps.
- [0695] 5. Prepare 1×, 0.1× and 0.01× suspensions of the Test Organism in pre-warmed PBS supplemented with 2 mM NH₄Cl, or a Vehicle control solution, in 50 ml polypropylene tubes as shown below.
- [0696] 6. Draw 2.0 ml of each suspension using a repetitive pipet.
- [0697] 7. Drip slowly 0.2 ml of the Test Organism (1×, 0.1×, 0.01× groups), or an equal volume of Vehicle control, onto each wound and surrounding shaved skin

area. Gently spread applied suspension onto the wound and the entire shaved skin area using a pipet tip.

- [0698] 8. Repeat application of Test Organism or Vehicle control daily for a total of 14 days.
- [0699] 9. Measure wound size by wound planimetry and obtain photo images of each wound on Day 1, 3, 5, 7, 9, 11, 13 and 15 using Image Analyzer (Image-pro plus version 4.5, Media Cybernetics Inc).
- [0700] 10. Calculate % wound closure and wound half-closure time (CT₅₀) for each group.

TABLE 9

GROUP	1× D23 (ml)	PBS pH 7.4 (ml)	1M NH ₄ Cl (μl)	CFU/ ml	CFU/ wound
Vehicle (control)		5.0	10	0	0
1× D23	5.0	0	10	10 ⁹	2 × 10 ⁸
0.1× D23	0.5	4.5	10	10 ⁸	2 × 10 ⁷
0.01× D23	0.05	4.95	10	10 ⁷	2 × 10 ⁶

Day 15 (Upon Study Completion): Collection of Wound Tissues Samples and Histopathology Analyses

- [0701] 1. Obtain half-wound tissue samples from four mice per group using aseptic technique to avoid cross-contamination of tissues.
- [0702] 2. Proceed with histopathology analyses.
- [0703] 3. Store temporarily at -70° C. the remainder half-wound samples and the additional four full-size wound tissues from each group for further evaluation.
- [0704] As shown in FIG. 5A, topical application of 10⁹ CFU/ml of strain D23 significantly (*p<0.05) accelerated wound healing. The sample size was N=8 animals/gp. The group receiving the highest doses showed significant improvement in wound closure from day 5 to day 15, with the most pronounced improvement of 83% observed on day 9, post-wounding. This study demonstrates the potential therapeutic benefit of ammonia oxidizing bacteria, e.g., D23, to diabetic foot ulcers, chronic wounds, and other related indications.
- [0705] FIG. 5B is a plot showing CT₅₀ versus control (vehicle) and 10⁹ CFU/ml D23. CT₅₀ is the time required to achieve a 50% wound closure. As shown in the plot, those wounds having application of D23 provided for lower CT₅₀ values.
- [0706] FIG. 5C is a plot of another experiment in which the protocol above was carried out to obtain wound closure measurements versus time. Control (vehicle) wounds were tested and compared to D23 at 10⁹ CFU/ml wounds. This plot shows the effects of D23 when immersion pre-treatment and topical application was carried out.
- [0707] FIG. 5D is a plot of another experiment in which the protocol above was carried out, without immersion pre-treatment, to obtain wound closure measurements versus time. Control (vehicle) wounds were tested and compared to applications of D23 at 10⁹ CFU/ml and 10¹⁰ CFU/ml to wounds. This plot shows the effects of D23 when topical application was carried out.
- [0708] FIG. 5E is a plot showing CT₅₀ versus control (vehicle) and 10⁹ CFU/ml D23, with and without immersion pre-treatment, and 10⁹ CFU/ml D23 without pre-treatment.

As shown in the plot, those wounds having application of D23 provided for lower CT50 values.

[0709] FIG. 5F are images of the wound healing experiments, at Day 1, Day 11, and Day 15. AOB represents D23.

[0710] Possible modulation of inflammatory responses coupled with anti-infective action of D23 could prove an effective topical treatment against diabetic and other chronic wounds.

[0711] FIG. 5G are plots of blood glucose levels in the mice tested for the control (vehicle) and various concentrations of D23. "IM" shown in the x-axis of the right-hand panel plot represents those tests down with an immersion pre-treatment of D23. FIG. 5H is a plot of body weight of the animals used in testing for the study including immersion pre-treatment, over the time of the study. FIG. 5I are plots of body weight of the animals used in testing for the study, including the immersion pre-treatment study, and the study done without immersion pre-treatment, over the time of the studies.

[0712] In Study 2, the effect of pretreatment of db/db mice with 10^9 cells/ml of B244 on wound closure was examined. Groups of seven mice were treated topically with 10^9 cells/ml of B244 with and without prior body immersion. One additional group of seven mice was treated topically with 10^{10} cells/ml of D23 (B244). Corresponding vehicle groups (seven mice) were run in parallel with and without body immersion as negative controls. Wound surface area and photo images of each wound were obtained as before. These studies reproduced the findings of Study 1 suggesting improvement of wound closure with a B244 dose of approximately 10^9 cells/ml. Moreover, topical treatment alone with 10^9 cells/ml improved wound closure rates simi-

[0713] Cytokine and growth factor expression in D23-treated diabetic animals was investigated using Luminex technology. Specifically, expression of growth-regulated oncogene/keratinocyte chemoattractant (Gro/KC), interleukin-1 (IL-1), interleukin-6 (IL-6), macrophage inflammatory protein-2 (MIP-2), tumor necrosis factor (TNF), and vascular endothelial growth factor (VEGF) was compared between D23-treated and control diabetic animals in serum samples obtained on Day 5 and Day 15 from four mice per group treated with or without prior body immersion. In similar Luminex analyses, lysates of tissues from D23-treated or Vehicle control animals obtained upon completion of the study (Day 15) were also analyzed. Abnormally high and sustained expression of inflammation markers, including MIP-2, TNF α and IL-1 β , has been previously associated with a dysregulated inflammatory response and impaired wound healing processes in db/db mice (Wetzler, 2000). Analyses of Day 5 and Day 15 serum samples yielded very low signal for all six cytokines in both D23-treated and vehicle control animals, a result indicating the lack of systemic effects following wound treatment with high D23 doses. In wound tissue lysates obtained on Day 15, MIP-2 levels (1155-1516 pg/100 g total protein) were significantly higher than the remaining five cytokines, with IL-6 and Gro/KC measured at much lower levels (44-48 pg/100 g total protein) and both IL-1 and VEGF being close to undetectable (≤ 3.8 pg/100 g total protein). Overall, no difference was observed between D23-treated animals and vehicle control animals with or without full-body immersion in D23 suspensions. The levels of all six cytokines or growth factors measured in tissue lysates of all four groups of mice examined are summarized in Table 10 below.

TABLE 10

Cytokine levels measured in wound tissue lysates of D23-treated and vehicle control-treated db/db mice							
Treatment	Animal	Gro/KC (pg/100 g protein)	IL-1 β (pg/100 g protein)	IL-6 (pg/100 g protein)	MIP-2 (pg/ 100 g protein)	TNF α (pg/100 g protein)	VEGF (pg/100 g protein)
Vehicle (with prior immersion)	1-1	49	2.4	78	1089	28.7	5.8
	1-3	66	2.4	134	1335	31.2	4.5
	1-5	59	2.7	128	1112	25.7	4.2
	1-7	76	1.2	148	1013	9.4	4.1
	MEAN	62	2.2	122	1137	23.7	4.7
D23 10^9 cells/ml (with prior immersion)	3-1	49	2.1	66	1830	24.7	4.1
	3-3	75	1.8	162	1615	32.3	3.6
	3-5	50	2.4	132	1896	23.9	4.3
	3-7	17	1.5	28	720	9.0	3.4
	MEAN	48	1.9	97	1516	22.5	3.8
Vehicle (topical only)	5-1	43	1.5	90	833	13.2	3.6
	5-3	55	2.2	104	1312	18.6	3.6
	5-5	44	1.4	59	644	17.6	3.2
	5-7	100	3.8	168	1308	48.6	4.0
	MEAN	60	2.2	105	1024	24.5	3.6
D23 10^9 cells/ml (topical only)	6-1	82	2.2	105	1573	28.5	2.9
	6-3	18	0.8	36	943	8.0	2.5
	6-5	25	1.2	45	1027	9.5	2.2
	6-7	49	1.5	92	1077	18.5	2.9
	MEAN	44	1.4	69	1155	16.1	2.6

lar to the animals receiving topical treatments with immersion. Additional histopathology analyses by of H & E—stained wound tissue sections recovered on Day 5 did not reveal any differences between vehicle and D23 (B244)-treated wounds.

[0714] Pharmacokinetic evaluation of D23 (B244) in rodents was conducted during a 28-day repeat dose toxicology study as described in the section below. No separate single dose pharmacokinetic studies were run for D23 (B244).

Example 7

Toxicology

[0715] 28-Day Safety Study of *Nitrosomonas eutropha* D23 (B244) Application on Full-Thickness Wounds of Streptozotocin-Induced Diabetic Sprague-Dawley Rats

[0716] The objectives of this study were to determine the potential toxicity of *Nitrosomonas eutropha* D23 (B244) in rats when given dermally on wounded skin for a minimum of 28 days, and to evaluate the potential reversibility of any findings. In addition, the toxicokinetic characteristics of D23 (B244) were determined.

Study Design and Methods

[0717] The design was based on the study objectives, the overall product development strategy for the test article, and the following study design guidelines: OECD Guidelines 407 and 417, Committee for Human Medicinal Products (CHMP), and ICH Harmonised Tripartite Guidelines M3 (R2), S3a, and S6 (R1). The study design is outlined herein and results are shown in Table 11.

TABLE 11

28-Day Safety Study design								
Group	Test	Dose Level	Dose Volume	Dose	No. of Animals			
			(mL/kg)	Conc.	Main Study		Recovery	
No.	Material	(CFU/kg/day)	Split	(CFU/mL)	M	F	M	F
1	Control Article	0	0.8	0	10	10	5	5
2	AOB-D23-100	6×10^7	0.8	8×10^7	10	10	0	0
3	AOB-D23-100	6×10^8	0.8	8×10^8	10	10	0	0
4	AOB-D23-100	6×10^9	0.8	8×10^9	10	10	5	5

M = Male,

F = Female,

Conc. = Concentration,

CFU = Colony Forming Unit.

Control Article = 99.998% Phosphate Buffered Saline, pH 7.4 (PBS), 0.002% 1M NH₄Cl

[0718] For induction of diabetes, Streptozotocin was administered to Sprague Dawley rats via intraperitoneal injection on Day -4. Animals with blood glucose levels of >200 mg/dL were considered as responders to the Streptozotocin treatment and were used for the dosing phase of the study. Two full-thickness skin wounds were created per animal (1 on each side of the back of each anesthetized animal) using an 8-mm skin biopsy punch. The wounds were left uncovered during administration of the control and test article and also for the duration of the study. The test and control articles were administered to the appropriate animals dermally once daily (for 24 hours±1 hour) from Days 1 to

28. The end points evaluated in this study were the following: clinical signs, dermal findings, body weights, body weight changes, food consumption, ophthalmology, glucose analysis, clinical pathology parameters (hematology, coagulation, clinical chemistry, urinalysis, hemoglobin A1c, and methemoglobin), C-reactive protein and serum ferritin analysis, toxicokinetic parameters, gross necropsy findings, organ weights, and wound histopathology.

Results

[0719] The results for the endpoints evaluated in the 28-day GLP toxicology study are outlined below in Table 12.

TABLE 12

28-Day Safety Study-Results		
End points	Observations	Comments
Mortality	No unscheduled deaths during the course of the study were attributed to D23 (B244).	

TABLE 12-continued

28-Day Safety Study-Results		
End points	Observations	Comments
Clinical Observations	One control male was found dead on Day 41; the cause of death due to necrosis in the kidney, liver, pancreas, and spleen	Similar clinical signs have been previously associated with an uncontrolled diabetic state in rats and other animal models
	No test article D23 (B244)-related clinical signs were observed during the study. Clinical signs including abdominal distension, prominent backbone, fur staining, soft stools and ungroomed appearance were related to the diabetic state of the animals	
Dermal Scores	Skin discoloration (red/black) was present in both control and treated animals	
	No dermal irritation occurred during the study	
Body Weights and Body Weight Changes	No erythema or edema was observed following dermal administration of the test article	
	No D23 (B244)-related effects on body weight or body weight change were noted during the study.	
Food Consumption	Mean weight gain was observed throughout the study interval, with isolated instances of slight loss in individual animals across the dose groups which did not follow specific dose-related trends	
	There were no test article-related effects on food consumption.	
Ophthalmic Examinations	There were no D23 (B244)-related ophthalmologic changes during the study. The majority of the animals on study developed cataracts and there were no differences among dose groups.	The appearance of cataracts is a known complication of diabetes
Hematology, Coagulation, Hemoglobin A1c, and Methemoglobin	No test article-related changes were noted in hematology, coagulation, hemoglobin A1c, and methemoglobin parameters on Day 29 or 43.	
Clinical Chemistry	Isolated statistically significant differences were noted during the study; however, the values were within the historical control ranges and were not considered meaningful	
	No test article-related changes were noted on Days 29 or 43.	
Urinalysis	Isolated statistically significant differences were noted during the study; however, the values were within the historical control ranges and not considered meaningful	
	No test article-related effects	
C-reactive Protein and Serum Ferritin Analysis	No test article-related effects	
Gross Pathology	No test article-related gross findings were noted on Day 29 or Day 43	Any gross findings observed were considered to be related to the diabetic condition of the rats and incidental in nature
Organ Weights	There was an increase in adrenal weight in females at $\geq 6 \times 10^8$ CFU/kg/day on Day 29, whereas adrenal weight was decreased in males and there were no associated gross pathology findings making the association of this finding to D23 (B244) administration equivocal	
	Potential D23 (B244)-related organ weight changes noted at the terminal euthanasia (Day 29) were not observed at the end of the recovery period (Day 43)	
Histopathology Terminal	No D23 (B244)-related microscopic findings on Day 29.	
Euthanasia (Day 29)	Changes observed in the kidneys, large and small intestine, and urinary bladder were related to the diabetic state of the animals. The incidence and severity of these	

TABLE 12-continued

28-Day Safety Study-Results		
End points	Observations	Comments
	findings were similar in all study groups including controls.	
	Changes at the administration/wound sites included epidermal regeneration, fibrosis, and granulomatous inflammation. The incidence and severity of these findings were similar in all groups including controls	
Histopathology	Changes observed on Day 43 were similar to those reported on Day 29	
Recovery		
Euthanasia (Day 43)		

Conclusions

- [0720] Once daily application of D23 (B244) on rat wounds was well tolerated at levels of 6×10^7 , 6×10^8 , and 6×10^9 CFU/kg/day.
- [0721] No D23 (B244)-related mortality observed during the study
- [0722] Healing of full tissue thickness excisions was similar in all groups
- [0723] No D23 (B244)-related clinical signs or dermal irritation were observed
- [0724] No effects observed during the study on body weight, food consumption, clinical pathology parameters, c-reactive protein, or serum ferritin
- [0725] No test article-related gross necropsy findings or histopathologic findings
- [0726] The no-observed-adverse-effect level (NOAEL) was determined to be 6×10^9 CFU/kg/day (8×10^9 cells/ml)
- [0727] No specific target organs were identified
- [0728] No D23 related mortality occurred during the study. There were no D23-related clinical signs or dermal irritation, and there were no effects on body weight, body weight changes, food consumption, clinical pathology parameters, C-reactive protein, or serum ferritin during the study. There were no test article-related gross necropsy findings or histopathologic findings. Increases in adrenal weights were noted in the $>6 \times 10^8$ CFU/kg/day females on Day 29; however, association with D23 was considered equivocal based on the lack of a similar effect in the males, the lack of corresponding gross findings, and the lack of microscopic evaluation of this tissue.
- [0729] All wound sites were completely covered by epidermis and appeared to be in the remodeling/resolution phase, which was characterized by stratification of the epidermis with keratinization and refinement of the dermal collagen (synthesis, bundling, and degradation) and capillaries to restore the normal architecture of the epidermis and dermis. The incidence and severity were similar in all groups, including controls.

Example 8

Antibiotic Susceptibility

[0730] The activities of five antibiotics, each representing a different antibiotic class, were tested against *Nitrosomonas eutropha* D23. The antibiotics tested included clindamycin,

erythromycin, gentamicin, piperacillin with or without the β -lactamase inhibitor Tazobactam, and tetracycline. These were chosen based on the Clinical and Laboratory Standards Institute (CLSI) recommendations for routine testing and reporting of phylogenetically-related proteobacteria (*Pseudomonas aeruginosa*) listed under Non-fastidious organisms and Non-Enterobacteriaceae in the CLSI 24th Informational Supplement (M100-524), and also included topical or systemic antimicrobial agents commonly used against acne, such as clindamycin or tetracycline. Studies with clindamycin were included even though this antibiotic was not expected to be very effective at inhibiting *Nitrosomonas*, as is the case for other aerobic Gram-negative bacteria.

[0731] Minimal Inhibitory Concentrations (MICs) were determined by culturing *N. eutropha* D23 in decreasing concentrations of each of the five antibiotics. Bacterial growth at 30° C. was monitored for 48-72 hr by determining optical density (OD₆₀₀) values in samples collected at 24 hr intervals. MIC values were identified as the lowest antibiotic concentration from a two-fold dilution series leading to no increase in OD₆₀₀ measurements for the 2 or 3-day incubation period. The *N. eutropha* D23 phenotype in each antibiotic test was determined as Susceptible, Intermediate, or Resistant according to the MIC Interpretive Criteria provided by the CLSI. As summarized in Table 13, these studies demonstrated susceptibility of *N. eutropha* D23 to erythromycin and gentamicin and intermediate resistance to tetracycline and piperacillin suggesting the lack of strong antibiotic-resistance potential by the Drug Substance. Clindamycin resistance observed for *N. eutropha* D23 is in agreement with previous reports for natural resistance of aerobic Gram-negative bacteria to this antibiotic. In addition to testing the β -lactam antibiotic piperacillin alone, the broad range β -lactamase inhibitor Tazobactam was also tested in combination with piperacillin to assess the possible expression of β -lactamase(s) by *N. eutropha* D23. The results from this comparison showed no increase in *N. eutropha* D23 susceptibility, indicating the absence of β -lactamase expression by *N. eutropha* D23, at least under the conditions tested.

TABLE 13

MIC values for five antibiotics tested against <i>N. eutropha</i> D23 cultures in vitro			
Antibiotic	Antibiotic Class	MIC (μ g/ml)	MIC Interpretive Criteria*
Clindamycin	Lincosamide	>16	Resistant (≥ 4 μ g/ml)
Erythromycin	Macrolide	0.16	Susceptible (≤ 0.5 μ g/ml)
Gentamicin	Aminoglycoside	0.25	Susceptible (≤ 4 μ g/ml)
Piperacillin	β -lactam	64	Intermediate (32-64 μ g/ml)
Piperacillin/ Tazobactam	β -lactam/ β -lactamase inhibitor	64/4	Intermediate (32/4-64/4 μ g/ml)
Tetracycline	Tetracycline	8	Intermediate (8 μ g/ml)

*as recommended by the Clinical and Laboratory Standards Institute (values in parentheses represent MIC levels for corresponding Susceptible, Intermediate or Resistant outcomes)

Conclusions

[0732] These studies demonstrate susceptibility of D23 (B244) to macrolide and aminoglycoside antibiotics and resistance to lincosamides, results that indicate the lack of strong antibiotic-resistance potential by the Drug Substance.

Example 9

Elucidation of Structure of *N. eutropha*

[0733] *N. eutropha* was defined at the species and the strain level using PCR and gene sequencing methodologies. The species level was defined as *N. eutropha* by sequencing of the V1-V5 variable regions of the 16S rRNA gene. *N. eutropha* was defined as a novel *N. eutropha* strain D23 by identification of a unique gene from whole genome sequence analysis. *N. eutropha* was defined at the species level as *N. eutropha* by 16S rRNA gene sequencing using the MicroSeq 500 rDNA Bacterial Identification PCL and sequencing kit.

[0734] Strain identity may be determined using custom primers, which correspond to the underlined portions of the following sequence and the D23 1c1355 sequence & primers Table 14 below. While not wishing to be bound by theory, it is believed that gene D23 1c1355 is unique to *N. eutropha* D23, and thus performing a PCR amplification reaction within gene D23 1c1355 will indicate whether *N. eutropha* D23 is present in a given sample.

TABLE 14

D23_1c1355 sequence & primers				
Primer	Sequence (5'-3')	T _m (° C.)	Position	Product size (bp)
D23_1c1355-F (SEQ ID NO: 64)	AATCTGTCTCCACAGGCAGC	54	287-305	595
D23_1c1355-R (SEQ ID NO: 65)	TATACCCACACCCACGCTA	54	881-862	

D23_1c1355 outer membrane autotransporter
barrel domain-containing protein (SEQ ID NO: 66)

TTGGTTGGTTTGAACAGGTAAGGAGAGGAGGAAAATCGCCAGAATAT
10 20 30 40 50

CGTCGCCAAAGGTTATCGGATCACCATAGCTTATCCACTCAAAGGGGAGA
60 70 80 90 100

TTATCATGAGCAAGGTTCTGTCGATTAAAAAGAGTTTATATACGGTTACT
110 120 130 140 150

GCACTAACTCTCGGTTTCGGACCATTTGTGACAGCGAGTGGACAATCATT
160 170 180 190 200

CGAAGAAACACCCGTACAAACACCCGGACGAGCTTTTGCAGTGGACAATT
210 220 230 240 250

TAAAGGGTATCTGTGTACAAACACAAGTGAAGACCCCTCATTAGCAATA
260 270 280 290 300

GCTTGCACCTTCGCACTGGGCGGGATAAATGATATTACCGCGCAGAACTCT
310 320 330 340 350

GTCTCCACAGGCAGCGATTCTAGGCCGAGTCGCGATTACTTCTCCCT
360 370 380 390 400

ATCAGTTTATTCGCGACGAATGAAAGCATACAGCGGCTAACAGGTCGC
410 420 430 440 450

TCTGCTGAGAAACGTCAGCAACAATCCTCTTTTACTACAAAGCTCAGC
460 470 480 490 500

GTGGTAGCAGGCACGCCATCATTTGGCACTTCTGGTTTTATAGGCGCTG
510 520 530 540 550

TAGGGGTTTCGCTGAGCGGTGGCGGAGCTTTGGTGAACGCAATACCGCT
560 570 580 590 600

GAAGGGCAGACCGGTTTTCAATTGAATACCCGGCAAACCGCTGATGAT
610 620 630 640 650

CGATTATTCATTTAATCAAAAATTGATTGGCGGCTTTTCTTTAATTATC
660 670 680 690 700

TGGGGACAGATCGTAATTTGAGATTGGCGAGTGGGACTTGAATCCGAT
710 720 730 740 750

AGCTATCGGTTTGCACCCTTTGTGCTTTTTCAGACCACTACCAATAGCTA
760 770 780 790 800

CTTAACCTCTGATGGGAGGGTATGCTTTGGTTAATTATCGTTCCACGCGCA
810 820 830 840 850

GCGTTTCGAGTCAAAATGACATCACGTTTGATAACGCCACAGCCAATAT
860 870 880 890 900

GATGCTAATCAGTTTTTTGCTAGCGTGGGTGGGTATACCTTTACTTT
910 920 930 940 950

AATGGATGGATGGAATCTGCGAGGATATGGTCGCGGGGACTTTAGTGATA
960 970 980 990 1000

TTAGTATCCAGAGCTTTTCAGGAAAAGGTGGCGTTGCTCATAGTGGGAAC
1010 1020 1030 1040 1050

GATAGTTTATCTCTTGCTATGCTGTGAATAAACAACCATACGCTCGGT
1060 1070 1080 1090 1100

TACCAGTACATTAGGCGTTGAACTTAGTCATGCAATTAGCACCAGAACTT
1110 1120 1130 1140 1150

TTATTCCCGTCATTATCCCGAGACTGCGTGCAAGATGGGTGCATGAATTT
1160 1170 1180 1190 1200

GAAACAATGCCAGAACTATCACGGCCGGTTTCTAGGCCAGAACTATAG
1210 1220 1230 1240 1250

-continued
TCCCACCTTCTGCATCAATGGCAGTTGCAAGCTCAGTGCGTAATTGGGCAA
1260 1270 1280 1290 1300
ACCTGGGGGTTGGAGTGCAAAATGCTGTTTGCCCGCTCGATTATCGGGTAC
1310 1320 1330 1340 1350
ATTAATTACGACAGATTAATTATCAAGCACGCGGAGAACAATATCATTTTC
1360 1370 1380
TGGTGGGATTTCGTATGAATTCTAA

Example 10

Administering Ammonia Oxidizing Bacteria to the Back of the Head to Change the Skin Microbiome

[0735] Ammonia oxidizing bacteria (*N. eutropha* D23) was applied topically to the back of the head of a subject for over 2 weeks. The dose was 3×10¹⁰ CFU applied per day. The product concentration was 1×10⁹ CFU/ml (15 ml, two times a day) in a phosphate buffer with magnesium chloride. On each day a skin swab was taken to isolate and sequence all the bacterial DNA that was present, using isolation and sequencing protocols known in the art.

[0736] Ammonia oxidizing bacteria of the genus *Nitrosomonas* was not present in the Day 0 sample, and was detected and present in the Day 7, 14, and 16 skin swabs.

[0737] As shown in FIGS. 17 and 18, which plots the proportion versus bacterial genus for Day 0, 1, 8, 14, and 16, the application of ammonia oxidizing bacteria led to proportional increases in commensal non-pathogenic *Staphylococcus* (which was at least 98% *Staphylococcus epidermidis*) from close to 0% on day 0 to approximately 50% on day 16. Additionally, application of ammonia oxidizing bacteria led to a proportional reduction in potentially pathogenic or disease associated *Propionibacteria* over the time period tested (from over 75% on day 0 to less than 50% on day 16). Application of ammonia oxidizing bacteria also led to reductions in potentially pathogenic or disease associated *Stenotrophomonas* over the time period tested (from 0.1% on day 0 to less than 0.01% on day 16.)

[0738] Some of the data shown in FIGS. 1 and 2 is also presented below in Table 15.

TABLE 15

Genera by Day			
Day	Proportion by genus: <i>Propionibacteria</i>	Proportion by genus: <i>Staphylococci</i>	Proportion by genus: <i>Stenotrophomonas</i>
0	0.78	0.01	0.13
1	0.79	0.1	0
8	0.8	0.15	0
14	0.55	0.45	0.001
16	0.48	0.49	0

[0739] As shown in Table 15, the proportion of *Propionibacteria* was reduced after about 14 days (compare data for Day 0, 1, and 8 with Day 14 and 16 in Table 15). The proportion of *Staphylococci* increased after about two weeks (compare data for Day 0, 1, and 8 with Day 14 and 16 in Table 15). The proportion of *Stenotrophomonas* decreased after about 1 day (compare data for Day 0 with Day 1, 8, 14, and 16 in Table 15).

[0740] These changes in the skin microbiome composition to a less pathogenic state indicate that application of ammonia oxidizing bacteria would be useful in treatment of dermatologic diseases including but not limited to acne, eczema, skin infections, and rosacea.

Example 11

Studies with Ammonia-Oxidizing Bacteria for the Human Skin: Cosmetic Effects, Safety, Detection and Skin Metagenomics

[0741] A blinded, placebo-controlled 24 human volunteer study randomized 4:1 AOB to placebo control was performed. Subjects applied a *Nitrosomonas* suspension (10⁹ CFU/ml, 2 times per day, for a total of 3×10¹⁰ CFU per day) to their face and scalp twice daily for one week and were followed for two additional weeks post-application. Volunteers were instructed to refrain from using hair products during the one-week AOB application as well as the week following application, then returned to regular shampoo use for the third week. Scalp swabs were obtained on Day 0 as baseline controls and on Day 1, 3, 8, 14 and 21 to assess presence/absence of AOB by PCR and 16S rRNA sequencing analyses.

[0742] No serious adverse events were associated with AOB application for one week and the product was deemed safe. AOB users reported a clear improvement in skin condition and quality, as indicated by self-assessment reports completed after the seven-day application period. Using AOB-specific PCR analyses of the skin samples, we could demonstrate presence of the bacteria in 83-100% of AOB users during the application period, whereas no AOB were detected in the placebo control samples. All subjects lacked AOB from baseline swabs obtained prior to study initiation, consistent with the predicted sensitivity of these bacteria to soaps and other commercial products. Amplification of the 16S rRNA gene and sequencing of a subset of samples confirmed presence of AOB in corresponding samples and suggested potential trends in modulating the skin microbiome by topical AOB application. In summary, live AOB-based products are safe and could hold great promise as novel self-regulating topical delivery agents of nitrite and nitric oxide to the human skin.

[0743] As shown in Table 16, below, the proportion of *Nitrosomonas* (AOB) went up when comparing Day 0 versus Day 8. The proportion of other bacteria, *Propionibacterium*, *Enterobacter*, and *Citrobacter* went down, when comparing Day 0 versus Day 8. The p-values indicated in Table 16 demonstrate that the most significant change between Day 0 and Day 8 was observed with *Nitrosomonas* (AOB) followed by *Propionibacterium*. *Enterobacter* and *Citrobacter* also showed changes between Day 0 and Day 8 to a lesser degree.

TABLE 16

Trends in microbiome composition following AOB application (Day 0 versus Day 8)		
Genus	P-value (unadjusted)	Trend
<i>Nitrosomonas</i> (AOB)	0.0039	Up
<i>Propionibacterium</i>	0.0078	Down

TABLE 16-continued

Trends in microbiome composition following AOB application (Day 0 versus Day 8)		
Genus	P-value (unadjusted)	Trend
<i>Enterobacter</i>	0.0346	Down
<i>Citrobacter</i>	0.036	Down

[0744] Because nitrite and nitric oxide have been implicated in critical physiological functions, such as vasodilation, skin inflammation and wound healing, we have hypothesized that AOB may have beneficial effects on both healthy and immunopathological skin conditions by metabolizing ammonia from sweat while concurrently driving skin acidification. We reasoned that *Nitrosomonas* would be safe for human use because they are slow-growing and incapable of utilizing organic carbon sources, they are sensitive to antibiotics, and they have never been linked to animal or human disease. Here we describe a blinded, placebo-controlled 24 human volunteer study where subjects applied a live *Nitrosomonas* suspension to their face and scalp twice daily for one week and were subsequently followed for two additional weeks. Volunteers did not use hair products during the first and second week, then they returned to their regular routine for the third week. Scalp swabs were obtained on Day 0 as baseline controls and on Day 1, 3, 8, 14 and 21 to assess presence/absence of *Nitrosomonas* and to examine microbial diversity. Importantly, no adverse events were associated with topical application. PCR analyses demonstrated presence of the bacteria in 83%-100% of skin swabs obtained from AOB users during or immediately after completion of the one-week application period (Day 1, 3 or 8) and in 60% of the users on Day 14, but not in any of the placebo control samples. All subjects lacked AOB from baseline swabs obtained prior to study initiation. Increased levels of AOB during the one-week application period correlated with a qualitative improvement in skin condition, in contrast to no improvement reported by placebo control subjects. Sequencing of the 16S rRNA gene amplification product obtained from a subset of subjects verified the presence of AOB in corresponding samples and suggested potential modulation of the skin microbiome composition. In summary, live *Nitrosomonas* are well tolerated and may hold promise as novel self-regulating topical delivery agents of nitrite and nitric oxide to the human skin.

[0745] Here, we present the results from preliminary studies in humans where we have begun evaluating topical application of a *Nitrosomonas* suspension to the human skin and the potential of using AOB as natural delivery systems of NO/NO₂⁻ in vivo. We have explored methodologies for AOB detection in skin specimens and the possible effects of AOB in skin microbial communities, as well as collected important user feedback from the early adopters of our topical cosmetic.

[0746] Methods

[0747] Culture Conditions.

[0748] *N. eutropha* D23 was propagated in batch culture at 28-30° C. in mineral salt medium supplemented with 20-50 mM NH₄⁺ and sodium carbonate as the carbon source [Ensign et al, 1993]. For continuous culture, D23 was grown at ~10⁹ cells/ml in a 1 liter mini-Bioreactor (Applikon Biotechnology) at 28° C. using sodium carbonate for both pH neutralization and the carbon source.

[0749] Nitrite Quantification.

[0750] Nitrite concentrations in culture supernatants were determined using the Griess colorimetric assay [Hageman and Kucklesby, 1971] and sodium nitrite as standards.

[0751] DNA Extraction from Skin Swabs.

[0752] Samples were maintained in 1 ml of 10% AssayAsure Bioservative (Thermo Scientific) diluted in PBS. Biomass was centrifuged and cells were lysed using a method developed for skin specimens [Grice, 2009] with modifications to the buffer designed to maintain long DNA integrity. DNA was then purified using the PowerLyzer UltraClean microbial DNA isolation kit (Mo Bio Laboratories). *N. eutropha* D23 was identified using a 3-gene PCR signature amplifying the ammonia monooxygenase encoding locus amoCAB.

[0753] PCR and Library Preparation.

[0754] Full-length 16S rRNA genes were amplified in duplicate reactions using a cocktail of primers and AccuPrime DNA polymerase SuperMix kit (Life Technologies). All PCR products were directly treated with the SMRTbell Template Prep Kit followed by the DNA/Polymerase Binding Kit P4 (Pacific Biosciences).

[0755] 16S rDNA Sequencing and Analysis.

[0756] PCR products were sequenced using the Pacific Biosciences RS instrument [Eid, 2009]. Raw base calls were transformed to consensus DNA sequences using the Pacific Biosciences Consensus Tools package and then processed with the Whole Biome Microbiome Profiling Platform to obtain phylum-genus and strain-level frequency measures for each sample.

[0757] Human Volunteer Study.

[0758] A total of 24 male volunteers were included in a blinded, placebo-controlled, study each for a total of three weeks according to a protocol for topical AOB-001 use approved by the Allendale Institutional Review Board (Old Lyme, Conn.). Written informed consent was obtained from each study participant. Subjects applied 15 ml of an aqueous suspension of *N. eutropha* (AOB-001), or placebo (vehicle), twice daily containing ~10⁹ cells/ml.

[0759] The human volunteer study design for the preliminary evaluation of a *Nitrosomonas*-containing topical suspension (AOB-001) is shown in FIG. 5K. Detection of AOB was performed by PCR in scalp swab samples. FIG. 5L shows PCR analyses of scalp swabs collected during the study. The left panel indicates the percent-positive samples for AOB-specific three-gene signature (amoA, amoB, amoC). The right panel indicates the Composite PCR scores for a total of six samples collected from each of 23 volunteers. The scoring scheme used for the positive samples collected at each of six sampling points is indicated.

[0760] Skin microbiome composition prior and during AOB-001 application were obtained by 16S rDNA sequencing. FIG. 5M indicates that genus-level bacterial diversity as determined by 16S rDNA sequencing in skin swab samples collected before and after topical application of AOB-001.

[0761] The percentage of the total sequence reads representing each of twelve bacterial genera in samples collected at baseline prior to application (Day 0) and immediately after the one week application (Day 8), or one week after stopping topical application (Day 14), are shown.

[0762] FIG. 5N indicates changes in abundance of *Nitrosomonas* and other species in skin samples collected before and after AOB-001 application. Panel A shows percentages of the total 16S rDNA sequence reads representing

Nitrosomonas prior to application (Day 0), immediately after the one-week application (Day 8), or one week after terminating application (Day 14) are shown. Panel B shows a change in patterns in abundance of species detected by 16S rDNA sequencing in Day 0 versus Day 8 samples collected from AOB users.

[0763] AOB-001 users report an improvement in skin condition. FIG. 5O shows a user evaluation of AOB-001. Assessment of AOB-001 cosmetic effects was provided by 23 volunteers upon completion of the one week application to their scalp and face. Subjects were plotted in order of increasing composite PCR scores. (The responses were categorized as 2=agree strongly; 0=no change; -2=disagree strongly). In summary, AOB-001 is well-tolerated. The user responses in a blind study indicate improved skin/scalp condition. AOB (*Nitrosomonas*) are readily detectable in skin microbiome samples by PCR and 16S rRNA gene sequencing. Preliminary microbiome analyses indicate modulation of skin microbiota by AOB.

SUPPLEMENTARY TABLE 1

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld	C91 Alias
fig/66666666.60966.peg.1	CDS	35	1414	2	+	1380	Chromosomal replication initiator protein DnaA	Cell Division Subsystem including YidCD; DNA replication cluster 1	D23_1c0001	Neut_0001
fig/66666666.60966.peg.2	CDS	1619	2740	2	+	1122	DNA polymerase III beta subunit (EC 2.7.7.7)	Cell Division Subsystem including YidCD; DNA replication cluster 1	D23_1c0002	Neut_0002
fig/66666666.60966.peg.3	CDS	2798	5227	2	+	2430	DNA gyrase subunit B (EC 5.99.1.3)	Cell Division Subsystem including YidCD; DNA gyrase subunits; DNA replication cluster 1; DNA topoisomerases, Type II, ATP-dependent; Resistance to fluoroquinolones	D23_1c0003	Neut_0003
fig/66666666.60966.peg.4	CDS	5248	5691	1	+	444	FIG039061: hypothetical protein related to heme utilization	-none-	D23_1c0004	Neut_0004
fig/66666666.60966.peg.5	CDS	5748	6479	3	+	732	tRNA pseudouridine synthase A (EC 4.2.1.70)	Colicin V and Bacteriocin Production Cluster; RNA pseudouridine synthases; RNA modification Bacteria; RNA processing	D23_1c0005	Neut_0005
fig/66666666.60966.peg.6	CDS	7261	7518	1	+	258	4Fe—4S ferredoxin, iron-sulfur binding	Inorganic Sulfur Assimilation	D23_1c0009	Neut_0127
fig/66666666.60966.peg.7	CDS	7584	7946	3	+	363	FIG00858425; hypothetical protein	-none-	D23_1c0010	Neut_0128
fig/66666666.60966.peg.8	CDS	11430	7966	-3	-	3465	Transcription-repair coupling factor	Transcription factors bacterial; Transcription repair cluster	D23_1c0011	Neut_0129
fig/66666666.60966.peg.9	CDS	12737	11457	-2	-	1281	InterPro IPR003416 COGs COG3174	-none-	D23_1c0012	Neut_0130
fig/66666666.60966.peg.10	CDS	14499	12730	-3	-	1770	Single-stranded-DNA-specific exonuclease RecJ (EC 3.1.—.—)	DNA Repair Base Excision	D23_1c0013	Neut_0131
fig/66666666.60966.peg.11	CDS	15277	14681	-1	-	597	InterPro IPR000345	-none-	D23_1c0014	Neut_0132
fig/66666666.60966.peg.12	CDS	16285	15365	-1	-	921	Indole-3-glycerol phosphate synthase (EC 4.1.1.48)	Chorismate: Intermediate for synthesis of Tryptophan, PABA antibiotics, PABA,	D23_1c0015	Neut_0133

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.									
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld C91 Alias
fig/66666666.60966.peg.13	CDS	17321	16296	-2	-	1026	Anthraniolate phosphoribosyltransferase (EC 2.4.2.18)	3-hydroxyanthranilate and more;; Tryptophan synthesis Auxin biosynthesis; Chorismate: Intermediate for synthesis of Tryptophan, PAPA antibiotics, PABA, 3-hydroxyanthranilate and more;; Tryptophan synthesis Chorismate: Intermediate for synthesis of Tryptophan, PAPA antibiotics, PABA, 3-hydroxyanthranilate and more;; Tryptophan synthesis -none-	D23_1c0016 Neut_0134
fig/66666666.60966.peg.14	CDS	17920	17318	-1	-	603	Anthraniolate synthase, amidotransferase component (EC 4.1.3.27)	Chorismate: Intermediate for synthesis of Tryptophan, PAPA antibiotics, PABA, 3-hydroxyanthranilate and more;; Tryptophan synthesis -none-	D23_1c0017 Neut_0135
fig/66666666.60966.peg.15	CDS	18046	19545	1	+	1500	Putative sensor-like histidine kinase YfhK	De Novo Purine Biosynthesis; De Novo Purine Biosynthesis	D23_1c0018 Neut_0136
fig/66666666.60966.peg.16	CDS	19644	20081	3	+	438	FIG00858754; hypothetical protein	-none-	D23_1c0019 Neut_0137
fig/66666666.60966.peg.17	CDS	20101	21465	1	+	1365	Putative sensory histidine kinase YfhA	-none-	D23_1c0020 Neut_0138
fig/66666666.60966.peg.18	CDS	22742	21474	-2	-	1269	PDZ/DHR/GLGF domain protein	-none-	D23_1c0021 Neut_0139
fig/66666666.60966.peg.19	CDS	26700	22798	-3	-	3903	Phosphoribosylformylglycinamide synthase, synthetase subunit (EC 6.3.5.3) / Phosphoribosylformylglycinamide synthase, glutamine amidotransferase subunit (EC 6.3.5.3)	De Novo Purine Biosynthesis; De Novo Purine Biosynthesis	D23_1c0022 Neut_0140
fig/66666666.60966.peg.20	CDS	26942	28510	2	+	1569	hypothetical protein	-none-	D23_1c0023 Neut_0141
fig/66666666.60966.peg.22	CDS	28682	28867	2	+	186	hypothetical protein	-none-	D23_1c0024 NA
fig/66666666.60966.peg.23	CDS	29060	28851	-2	-	210	Death on curing protein, Doc toxin	Phd-Doc, YdeE-YdeD toxin-antitoxin (programmed cell death) systems	D23_1c0025 NA
fig/66666666.60966.peg.24	CDS	29367	29227	-3	-	141	Prevent host death protein, Phd antitoxin	Phd-Doc, YdeE-YdeD toxin-antitoxin (programmed cell death) systems	D23_1c0026 Neut_0143

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.									
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	C91 Alias
fig/66666666.60966.psg.25	CDS	29726	30082	2	+	357	birl1219; hypothetical protein	-none-	Neut_0144
fig/66666666.60966.psg.26	CDS	30113	31672	2	+	1560	NAD(P)HX epimerase/ NAD(P)HX dehydratase	YjeE; <h>YjeE	Neut_0145
fig/66666666.60966.psg.29	CDS	31959	32078	3	+	120	hypothetical protein	-none-	NA
fig/66666666.60966.psg.30	CDS	32096	32914	2	+	819	O-antigen export system permease	-none-	Neut_0146
fig/66666666.60966.psg.31	CDS	33063	33266	3	+	204	hypothetical protein	-none-	Neut_0147
fig/66666666.60966.psg.32	CDS	33441	33995	3	+	555	hypothetical protein	-none-	Neut_0148
fig/66666666.60966.psg.33	CDS	34044	34424	3	+	381	hypothetical protein	-none-	NA
fig/66666666.60966.psg.34	CDS	34530	35588	3	+	1059	putative transposase	-none-	Neut_0149
fig/66666666.60966.psg.36	CDS	36348	36064	-3	-	285	HigA protein (antitoxin to HigB)	Toxin-antitoxin replicon stabilization systems	Neut_0150
fig/66666666.60966.psg.37	CDS	36621	36379	-3	-	243	HigB toxin protein	Toxin-antitoxin replicon stabilization systems	Neut_0151
fig/66666666.60966.psg.38	CDS	36580	36750	1	+	171	hypothetical protein	-none-	NA
fig/66666666.60966.psg.39	CDS	36747	38108	3	+	1362	Tetchoic acid export ATP-binding protein	Rhamnose containing glycans	Neut_0152
fig/66666666.60966.psg.40	CDS	38105	42433	2	+	4329	Glycosyl transferase, group 2 family protein	-none-	Neut_0153
fig/66666666.60966.psg.41	CDS	42537	43733	3	+	1197	glycosyl transferase, group 1/2 family protein	-none-	NA
fig/66666666.60966.psg.42	CDS	43945	44838	1	+	894	Alpha-L-Rha alpha-1,3-L-rhamnosyltransferase (EC 2.4.1.-)	Rhamnose containing glycans	Neut_0166
fig/66666666.60966.psg.43	CDS	45457	45140	-1	-	318	HigA protein (antitoxin to HigB)	Toxin-antitoxin replicon stabilization systems	Neut_0167
fig/66666666.60966.psg.44	CDS	45610	45470	-1	-	141	HigB toxin protein	Toxin-antitoxin replicon stabilization systems	Neut_0168
fig/66666666.60966.psg.45	CDS	45950	46279	2	+	330	Glycosyl transferase, group 2 family protein	-none-	Neut_0169
fig/66666666.60966.psg.47	CDS	47082	46804	-3	-	279	hypothetical protein	-none-	NA
fig/66666666.60966.psg.49	CDS	48719	47757	-2	-	963	Mobile element protein	-none-	Neut_0978
fig/66666666.60966.psg.50	CDS	48899	48777	-2	-	123	Mobile element protein	-none-	Neut_0357
fig/66666666.60966.psg.51	CDS	49218	48970	-3	-	249	Mobile element protein	-none-	Neut_2405
fig/66666666.60966.psg.52	CDS	49615	49502	-1	-	114	hypothetical protein	-none-	NA
fig/66666666.60966.psg.53	CDS	49842	50255	3	+	414	Nucleotidyltransferase (EC 2.7.7.-)	-none-	Neut_0172
fig/66666666.60966.psg.54	CDS	50257	50622	1	+	366	Nucleotidyltransferase (EC 2.7.7.-)	-none-	Neut_0173
fig/66666666.60966.psg.55	CDS	51293	50880	-2	-	414	Mobile element protein	-none-	NA
fig/66666666.60966.psg.56	CDS	51432	51253	-3	-	180	hypothetical protein	-none-	Neut_0176
fig/66666666.60966.psg.57	CDS	51530	52492	2	+	963	Mobile element protein	-none-	Neut_1746
fig/66666666.60966.psg.58	CDS	52657	52908	1	+	252	Mobile element protein	-none-	Neut_0884
fig/66666666.60966.psg.59	CDS	52964	53326	2	+	363	Mobile element protein	-none-	Neut_2499

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 GbkId	C91 Alias
fig/66666666.60966.peg.60	CDS	54452	53361	-2	-	1092	putative transposase	-none-	D23_1c0061	Neut_0177
fig/66666666.60966.peg.61	CDS	54765	54430	-3	-	336	FIG00859125;	-none-	D23_1c0062	Neut_0178
fig/66666666.60966.peg.62	CDS	55016	55774	2	+	759	hypothetical protein dTDP-Rha:A-D-GlcNAc- diphosphoryl polyprenol, A-3-L- rhamnosyl transferase WbbL	dTDP-rhamnose synthesis	D23_1c0063	Neut_0179
fig/66666666.60966.peg.63	CDS	56735	55788	-2	-	948	UDP-glucose 4- epimerase (EC 5.1.3.2)	CBSS- 296591.1.peg.2330; N-linked Glycosylation in Bacteria; Rhamnose containing glycans	D23_1c0064	Neut_0180
fig/66666666.60966.peg.64	CDS	56874	56746	-3	-	129	hypothetical protein	-none-	D23_1c0065	NA
fig/66666666.60966.peg.65	CDS	60470	57075	-2	-	3396	Adenylate cyclase (EC 4.6.1.1)Guanylate cyclase (EC 4.6.1.2)	cAMP signaling in bacteria	D23_1c0066	Neut_0181
fig/66666666.60966.peg.66	CDS	60633	60755	3	+	123	hypothetical protein	-none-	D23_1c0067	NA
fig/66666666.60966.peg.67	CDS	62853	60769	-3	-	2085	Ubiquinone biosynthesis monooxygenase UbiB	Ubiquinone Biosynthesis; Ubiquinone Biosynthesis-gio	D23_1c0068	Neut_0182
fig/66666666.60966.peg.68	CDS	63084	63821	3	+	738	hypothetical protein	-none-	D23_1c0069	Neut_0183
fig/66666666.60966.peg.69	CDS	64515	66023	3	+	1509	CBSS- 498211.3.peg.1514; hypothetical protein FIG039767;	-none-	D23_1c0070	NA
fig/66666666.60966.peg.70	CDS	66074	66751	2	+	678	hypothetical protein	-none-	D23_1c0071	NA
fig/66666666.60966.peg.71	CDS	66741	70157	3	+	3417	hypothetical protein FIG007317;	-none-	D23_1c0072	NA
fig/66666666.60966.peg.72	CDS	70190	71326	2	+	1137	hypothetical protein FIG005429;	-none-	D23_1c0073	Neut_0184
fig/66666666.60966.peg.73	CDS	71379	71939	3	+	561	hypothetical protein Lipid carrier: UDP-N- acetylglactosaminyltransferase (EC 2.4.1.--)	CBSS- 296591.1.peg.2330; N-linked Glycosylation in Bacteria	D23_1c0074	Neut_0185
fig/66666666.60966.peg.74	CDS	71949	73931	3	+	1983	Nucleoside-diphosphate sugar epimerase/dehydratase	CBSS-296591.1.peg.2330	D23_1c0075	Neut_0186
fig/66666666.60966.peg.75	CDS	74467	73949	-1	-	519	cytidine and deoxycytidylate deaminase family protein	-none-	D23_1c0076	Neut_0187
fig/66666666.60966.peg.76	CDS	74956	74594	-1	-	363	Mobile element protein	-none-	D23_1c0077	Neut_2499
fig/66666666.60966.peg.77	CDS	75263	75012	-2	-	252	Mobile element protein	-none-	D23_1c0078	Neut_0884
fig/66666666.60966.peg.78	CDS	75586	76446	1	+	861	Flagellar motor rotation protein MotA	Flagellar motility; Flagellum	D23_1c0079	Neut_0188

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 GbkId	C91 Alias
fig/66666666.60966.peg.79	CDS	76489	77433	1	+	945	Flagellar motor rotation protein MotB	Flagellar motility; Flagellum	D23_1c0080	Neut_0189
fig/66666666.60966.peg.80	CDS	77408	78205	2	+	798	FIG00858624: hypothetical protein	-none-	D23_1c0081	Neut_0190
fig/66666666.60966.peg.81	CDS	79621	78218	-1	-	1404	CysteinyI-RNA synthetase (EC 6.1.1.16)	Zinc regulated enzymes; RNA	D23_1c0082	Neut_0191
fig/66666666.60966.peg.83	CDS	79830	80384	3	+	555	Peptidyl-prolyl cis-trans isomerase PpiB (EC 5.2.1.8)	aminoacylation, Cys Peptidyl-prolyl cis-trans isomerase; Queuosine-Archaeosine	D23_1c0083	Neut_0192
fig/66666666.60966.peg.84	CDS	80403	80894	3	+	492	Peptidyl-prolyl cis-trans isomerase PpiB (EC 5.2.1.8)	Biosynthesis Peptidyl-prolyl cis-trans isomerase; Queuosine-Archaeosine	D23_1c0084	Neut_0193
fig/66666666.60966.peg.85	CDS	80972	81424	2	+	453	Rhodanese-related sulfurtransferase	Biosynthesis	D23_1c0085	Neut_0194
fig/66666666.60966.peg.86	CDS	82260	81439	-3	-	822	Undecaprenyl-diphosphatase (EC 3.6.1.27)	-none-	D23_1c0086	Neut_0195
fig/66666666.60966.peg.87	CDS	84206	82308	-2	-	1899	Thiamin biosynthesis protein ThiC	Thiamin biosynthesis	D23_1c0087	Neut_0196
fig/66666666.60966.peg.88	CDS	84412	85068	1	+	657	Protein-L-isopartate O-methyltransferase (EC 2.1.1.77)	Protein-L-isopartate O-methyltransferase; Stationary phase repair cluster; Ton and Tol transport systems	D23_1c0088	Neut_0197
fig/66666666.60966.peg.90	CDS	85216	86493	1	+	1278	Type I secretion outer membrane protein, TolC precursor	Multidrug Resistance Efflux Pumps; Ton and Tol transport systems	D23_1c0089	Neut_0198
fig/66666666.60966.peg.91	CDS	89009	86556	-2	-	2454	ATP-dependent protease La (EC 3.4.21.53) Type II	Proteasome bacterial; Proteolysis in bacteria, ATP-dependent	D23_1c0090	Neut_0199
fig/66666666.60966.peg.92	CDS	89253	89375	3	+	123	hypothetical protein	-none-	D23_1c0091	NA
fig/66666666.60966.peg.93	CDS	89433	89579	3	+	147	hypothetical protein	-none-	D23_1c0092	NA
fig/66666666.60966.peg.94	CDS	90769	89555	-1	-	1215	Serine-pyruvate aminotransferase (EC 2.6.1.51) L-alanine:glyoxylate aminotransferase (EC 2.6.1.44)	Photorespiration (oxidative C2 cycle); Pyruvate Alanine Serine Interconversions	D23_1c0093	Neut_0200
fig/66666666.60966.peg.95	CDS	93514	91088	-1	-	2427	ATP-dependent protease La (EC 3.4.21.53) Type I	Proteasome bacterial; Proteolysis in bacteria, ATP-dependent	D23_1c0095	Neut_0201

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld	C91 Alias
fig/66666666.60966.peg.96	CDS	94903	93620	-1	-	1284	ATP-dependent Clp protease ATP-binding subunit ClpX	Proteasome bacterial; Proteolysis in bacteria, ATP-dependent	D23_1c0096	Neut_0202
fig/66666666.60966.peg.97	CDS	95607	94963	-3	-	645	ATP-dependent Clp protease proteolytic subunit (EC 3.4.21.92)	Proteasome bacterial; Proteolysis in bacteria, ATP-dependent; cAMP signaling in bacteria	D23_1c0097	Neut_0203
fig/66666666.60966.peg.98	CDS	96907	95591	-1	-	1317	Cell division trigger factor (EC 5.2.1.8)	Bacterial Cell Division	D23_1c0098	Neut_0204
fig/66666666.60966.peg.99	CDS	97996	97241	-1	-	756	Short-chain dehydrogenase/reductase SDR	Transcription repair cluster	D23_1c0100	Neut_0205
fig/66666666.60966.peg.100	CDS	99750	98107	-3	-	1644	Heat shock protein 60 family chaperone GroEL	GroEL GroES	D23_1c0101	Neut_0206
fig/66666666.60966.peg.101	CDS	100080	99790	-3	-	291	Heat shock protein 60 family co-chaperone GroES	GroEL GroES	D23_1c0102	Neut_0207
fig/66666666.60966.peg.102	CDS	100244	101554	2	+	1311	Adenosylmethionine-8-amino-7-oxononanoate aminotransferase (EC 2.6.1.62)	Biotin biosynthesis; Biotin biosynthesis Experimental; Biotin synthesis cluster	D23_1c0103	Neut_0208
fig/66666666.60966.peg.103	CDS	101561	102967	2	+	1407	Metallo-beta-lactamase family protein, RNA-specific	Bacterial RNA-metabolizing Zn-dependent hydrolases; Ribonucleases in Bacillus	D23_1c0104	Neut_0209
fig/66666666.60966.peg.104	CDS	103374	103066	-3	-	309	Cytochrome c, class I	-none-	D23_1c0105	Neut_0210
fig/66666666.60966.peg.105	CDS	103536	104300	3	+	765	Exodeoxyribonuclease III (EC 3.1.11.2)	DNA repair, bacterial	D23_1c0106	Neut_0211
fig/66666666.60966.peg.106	CDS	104347	105459	1	+	1113	Alanine dehydrogenase (EC 1.4.1.1)	Pyruvate Alanine Serine Interconversions	D23_1c0107	Neut_0212
fig/66666666.60966.peg.107	CDS	106118	105597	-2	-	522	Conserved uncharacterized protein	Tolerance to colicin E2	D23_1c0108	Neut_0213
fig/66666666.60966.peg.109	CDS	107425	106253	-1	-	1173	Permeases of the major facilitator superfamily	-none-	D23_1c0109	Neut_0214
fig/66666666.60966.peg.110	CDS	108032	107454	-2	-	579	Uncharacterized protein family UPF0016	-none-	D23_1c0110	Neut_0215
fig/66666666.60966.peg.112	CDS	108821	109603	2	+	783	Ribulose-5-phosphate 4-epimerase and related epimerases and aldolases	-none-	D23_1c0113	Neut_0218
fig/66666666.60966.peg.113	CDS	109609	113274	1	+	3666	InterPro IPR000014:IPR001789:IPR002106: IPR002570:IPR003594: IPR003660: IPR003661:IPR004358:IPR005467	-none-	D23_1c0114	Neut_0219

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.									
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkid C91 Alias
fig/66666666.60966.psg.114	CDS	113292	114485	3	+	1194	COGs COG0642 Succinyl-CoA ligase [ADP-forming] beta chain (EC 6.2.1.5)	TCA Cycle	D23_1c0115 Neut_0220
fig/66666666.60966.psg.115	CDS	114489	115364	3	+	876	Succinyl-CoA ligase [ADP-forming] alpha chain (EC 6.2.1.5)	TCA Cycle	D23_1c0116 Neut_0221
fig/66666666.60966.psg.116	CDS	115402	115722	1	+	321	FIG00858523: hypothetical protein	-none-	D23_1c0117 Neut_0222
fig/66666666.60966.psg.117	CDS	115750	117177	1	+	1428	D-alanyl-D-alanine carboxypeptidase (EC 3.4.16.4)	CBSS-84588.1.psg.1247; Metalloprotease; (EC 3.4.17.---); Murein Hydrolases; Peptidoglycan Biosynthesis	D23_1c0118 Neut_0223
fig/66666666.60966.psg.118	CDS	117265	118227	1	+	963	Mobile element protein	-none-	D23_1c0119 Neut_1278
fig/66666666.60966.psg.119	CDS	120193	120056	-1	-	138	hypothetical protein	-none-	D23_1c0120 NA
fig/66666666.60966.ma.5	RNA	118725	120255	3	+	1531	Small Subunit Ribosomal RNA; ssuRNA; SSU rRNA	-none-	D23_1c0120 NA
fig/66666666.60966.psg.121	CDS	122376	122495	3	+	120	hypothetical protein	-none-	D23_1c0124 NA
fig/66666666.60966.psg.120	CDS	121863	121994	3	+	132	hypothetical protein	-none-	D23_1c0124 NA
fig/66666666.60966.ma.8	RNA	120652	123535	1	+	2884	Large Subunit Ribosomal RNA; lsuRNA; LSU rRNA	-none-	D23_1c0124 NA
fig/66666666.60966.ma.9	RNA	123600	123716	3	+	117	5S RNA	-none-	D23_1c0126 NA
fig/66666666.60966.psg.123	CDS	124878	124708	-3	-	171	hypothetical protein	-none-	D23_1c0127 NA
fig/66666666.60966.psg.124	CDS	125317	125496	1	+	180	hypothetical protein	-none-	D23_1c0129 Neut_0547
fig/66666666.60966.psg.125	CDS	125792	126799	2	+	1008	NAD-dependent epimerase/dehydratase	CBSS-296591.1.psg.2330	D23_1c0130 Neut_0225
fig/66666666.60966.psg.126	CDS	126808	128082	1	+	1275	UDP-glucose dehydrogenase (EC 1.1.1.22)	-none-	D23_1c0131 Neut_0226
fig/66666666.60966.psg.127	CDS	128985	128089	-3	-	897	Permeases of the drug/metabolite transporter (DMT) superfamily	-none-	D23_1c0132 Neut_0227
fig/66666666.60966.psg.128	CDS	129078	130283	3	+	1206	N-succinyl-L,L-diaminopimelate aminotransferase alternative (EC 2.6.1.17)	Lysine Biosynthesis DAP Pathway, GJO scratch	D23_1c0133 Neut_0228
fig/66666666.60966.psg.129	CDS	130311	131132	3	+	822	2,3,4,5-tetrahydropyridine-2,6-dicarboxylate N-succinyltransferase (EC 2.3.1.117)	Lysine Biosynthesis DAP Pathway, GJO scratch	D23_1c0134 Neut_0229

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.									
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld C91 Alias
fig/66666666.60966.peg.130	CDS	131322	131693	3	+	372	FIG00858507: hypothetical protein	-none-	D23_1c0135 Neut_0230
fig/66666666.60966.peg.131	CDS	131801	132127	2	+	327	FIG00858507: hypothetical protein	-none-	D23_1c0136 Neut_0231
fig/66666666.60966.peg.132	CDS	132190	132312	1	+	123	hypothetical protein	-none-	D23_1c0137 NA
fig/66666666.60966.peg.133	CDS	132314	133303	2	+	990	Biotin operon repressor/ Biotin-protein ligase (EC 6.3.4.15)	Biotin biosynthesis; Biotin biosynthesis; Biotin synthesis cluster; Biotin synthesis cluster	D23_1c0138 Neut_0232
fig/66666666.60966.peg.134	CDS	133331	134104	2	+	774	Pantothenate kinase type III, CoaX-like (EC 2.7.1.33)	Coenzyme A Biosynthesis; Coenzyme A Biosynthesis cluster	D23_1c0139 Neut_0233
fig/66666666.60966.peg.135	CDS	134123	134794	2	+	672	GTP-binding protein EngB	Universal GTPases	D23_1c0140 Neut_0234
fig/66666666.60966.peg.136	CDS	134938	135945	1	+	1008	Porphobilinogen synthase (EC 4.2.1.24)	Heme and Siroheme Biosynthesis; Zinc regulated enzymes	D23_1c0141 Neut_0235
fig/66666666.60966.peg.137	CDS	136861	136064	-1	-	798	Phosphate transport ATP-binding protein PstB (TC 3.A.1.7.1)	High affinity phosphate transporter and control of PHO regulon; Phosphate metabolism	D23_1c0142 Neut_0236
fig/66666666.60966.peg.138	CDS	137797	136871	-1	-	927	Phosphate transport system permease protein PstA (TC 3.A.1.7.1)	High affinity phosphate transporter and control of PHO regulon; Phosphate metabolism	D23_1c0143 Neut_0237
fig/66666666.60966.peg.139	CDS	138817	137876	-1	-	942	Phosphate transport system permease protein PstC (TC 3.A.1.7.1)	High affinity phosphate transporter and control of PHO regulon; Phosphate metabolism	D23_1c0144 Neut_0238
fig/66666666.60966.peg.140	CDS	139048	139299	1	+	252	FIG00858998: hypothetical protein	-none-	D23_1c0146 Neut_0239
fig/66666666.60966.peg.141	CDS	140580	139432	-3	-	1149	Cell division protein FtsZ (EC 3.4.24.-)	Bacterial Cell Division; Bacterial Cytoskeleton; cell division cluster containing FisQ; cell division core of larger cluster	D23_1c0147 Neut_0240
fig/66666666.60966.peg.142	CDS	141909	140650	-3	-	1260	Cell division protein FtsA	Bacterial Cell Division; Bacterial Cytoskeleton; cell division cluster containing FisQ; cell	D23_1c0149 Neut_0241

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld	C91 Alias
fig 66666666.60966.peg.143	CDS	142678	141950	-1	-	729	Cell division protein FisQ	division core of larger cluster Bacterial Cell Division; Bacterial Cytoskeleton; cell division cluster containing FisQ; cell division core of larger cluster	D23_1c0150	Neut_0242
fig 66666666.60966.peg.144	CDS	143654	142734	-2	-	921	D-alanine--D-alanine ligase (EC 6.3.2.4)	Peptidoglycan Biosynthesis; Peptidoglycan biosynthesis--glo; cell division cluster containing FisQ	D23_1c0152	Neut_0243
fig 66666666.60966.peg.145	CDS	144649	143651	-1	-	999	UDP-N-acetylglucosamine-6-phosphate uridylyltransferase (EC 2.4.1.158)	Peptidoglycan Biosynthesis; UDP-N-acetylglucosamine-6-phosphate	D23_1c0153	Neut_0244
fig 66666666.60966.peg.146	CDS	146080	144659	-1	-	1422	UDP-N-acetylglucosamine-6-phosphate uridylyltransferase (EC 2.4.1.227)	Peptidoglycan Biosynthesis; Peptidoglycan biosynthesis--glo; cell division cluster containing FisQ	D23_1c0154	Neut_0245
fig 66666666.60966.peg.147	CDS	147159	146077	-3	-	1083	UDP-N-acetylglucosamine-6-phosphate uridylyltransferase (EC 2.4.1.227)	Peptidoglycan Biosynthesis; cell division core of larger cluster	D23_1c0155	Neut_0246
fig 66666666.60966.peg.148	CDS	148372	147212	-1	-	1161	Cell division protein FisW	Bacterial Cell Division; Bacterial Cytoskeleton; cell division cluster containing FisQ	D23_1c0156	Neut_0247
fig 66666666.60966.peg.149	CDS	149789	148377	-2	-	1413	UDP-N-acetylglucosamine-6-phosphate uridylyltransferase (EC 2.4.1.227)	Peptidoglycan Biosynthesis; Peptidoglycan biosynthesis--glo	D23_1c0157	Neut_0248
fig 66666666.60966.peg.150	CDS	150871	149786	-1	-	1086	Phospho-N-acetylglucosamine-6-phosphate uridylyltransferase (EC 2.4.1.227)	Peptidoglycan Biosynthesis	D23_1c0158	Neut_0249

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.									
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld C91 Alias
fig/66666666.60966.peg.151	CDS	152316	150943	-3	-	1374	transferase (EC 2.7.8.13) UDP-N-acetylglucosaminyl- D-glucosyl-2,6-diaminopimelate-D-alanyl-D-alanine ligase (EC 6.3.2.10)	Peptidoglycan Biosynthesis; Peptidoglycan biosynthesis--glo	D23_1c0159 Neut_0250
fig/66666666.60966.peg.152	CDS	153875	152313	-2	-	1563	UDP-N-acetylglucosaminyl- D-glucosyl-2,6-diaminopimelate ligase (EC 6.3.2.13)	Peptidoglycan Biosynthesis; Peptidoglycan biosynthesis--glo	D23_1c0160 Neut_0251
fig/66666666.60966.peg.153	CDS	155545	153872	-1	-	1674	Cell division protein FtsI [Peptidoglycan synthetase] (EC 2.4.1.129)	16S rRNA modification within P site of ribosome; Bacterial Cell Division; Bacterial Cytoskeleton; CBSS-83331.1.peg.3039; Flagellum in <i>Campylobacter</i> ; Peptidoglycan Biosynthesis	D23_1c0161 Neut_0252
fig/66666666.60966.peg.155	CDS	155895	155608	-3	-	288	Cell division protein FtsL	16S rRNA modification within P site of ribosome; Bacterial Cell Division; Bacterial Cytoskeleton; Stationary phase repair cluster	D23_1c0162 Neut_0253
fig/66666666.60966.peg.156	CDS	156845	155892	-2	-	954	rRNA small subunit methyltransferase H	16S rRNA modification within P site of ribosome; Bacterial Cell Division	D23_1c0163 Neut_0254
fig/66666666.60966.peg.157	CDS	157094	156861	-2	-	234	Cell division protein MraZ	16S rRNA modification within P site of ribosome; Bacterial Cell Division	D23_1c0164 Neut_0255
fig/66666666.60966.peg.158	CDS	157584	157859	3	+	276	DNA-3-methyladenine glycosylase II (EC 3.2.2.21)	ribosome; Bacterial Cell Division; Bacterial Cytoskeleton	D23_1c0165 NA
fig/66666666.60966.peg.159	CDS	158202	158420	3	+	219	Rho-specific inhibitor of transcription termination (YacO)	DNA Repair Base Excision Transcription factors bacterial	D23_1c0166 Neut_0257

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.									
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkid C91 Alias
fig/66666666.60966.peg.160	CDS	159328	158561	-1	-	768	InterPro IPR001173 COGs COG0463	-none-	D23_1c0167 Neut_0258
fig/66666666.60966.peg.161	CDS	159475	159924	1	+	450	InterPro IPR000086 COGs COG0494	-none-	D23_1c0168 Neut_0259
fig/66666666.60966.peg.162	CDS	160257	160814	3	+	558	possible (U92432) ORF4 [<i>Mitrosopira</i> sp. NpAV]	-none-	D23_1c0169 Neut_0260
fig/66666666.60966.peg.164	CDS	160969	161451	1	+	483	FIG00859298: hypothetical protein	-none-	D23_1c0170 Neut_0261
fig/66666666.60966.peg.165	CDS	161593	162063	1	+	471	Adenine phosphoribosyltransferase (EC 2.4.2.7)	Purine conversions; cAMP signaling in bacteria	D23_1c0171 Neut_0262
fig/66666666.60966.peg.166	CDS	162260	163573	2	+	1314	Seryl-tRNA synthetase (EC 6.1.1.11)	CBS- 326442.4.peg.1852; Glycine and Serine Utilization, tRNA aminoacylation, Ser	D23_1c0172 Neut_0263
fig/66666666.60966.peg.167	CDS	163620	164306	3	+	687	FIG00858527: hypothetical protein	-none-	D23_1c0173 Neut_0264
fig/66666666.60966.peg.168	CDS	165061	164351	-1	-	711	Phosphoglycerate mutase (EC 5.4.2.1)	Glycolysis and Gluconeogenesis; Phosphoglycerate mutase protein family	D23_1c0174 Neut_0265
fig/66666666.60966.peg.169	CDS	166178	165111	-2	-	1068	InterPro IPR001225	-none-	D23_1c0175 Neut_0266
fig/66666666.60966.peg.170	CDS	166643	166200	-2	-	444	FIG00858776: hypothetical protein	-none-	D23_1c0176 Neut_0267
fig/66666666.60966.peg.171	CDS	167465	166659	-2	-	807	CTP:inositol-1- phosphate cytidyltransferase (2.7.7.→)	-none-	D23_1c0177 Neut_0268
fig/66666666.60966.peg.172	CDS	168669	167509	-3	-	1161	Cysteine desulfurase (EC 2.8.1.7)	Alanine biosynthesis; CBS- 84588.1.peg.1247; mm5U34 biosynthesis bacteria; tRNA modification	D23_1c0178 Neut_0269
fig/66666666.60966.peg.174	CDS	169251	169631	3	+	381	FIG048548: ATP synthase protein I2	Bacteria -none-	D23_1c0180 Neut_0270
fig/66666666.60966.peg.175	CDS	169720	170472	1	+	753	ATP synthase A chain (EC 3.6.3.14)	-none-	D23_1c0181 Neut_0271
fig/66666666.60966.peg.176	CDS	170516	170788	2	+	273	ATP synthase C chain (EC 3.6.3.14)	-none-	D23_1c0182 Neut_0272
fig/66666666.60966.peg.177	CDS	170900	171301	2	+	402	ATP synthase B chain (EC 3.6.3.14)	-none-	D23_1c0183 Neut_0273
fig/66666666.60966.peg.178	CDS	171302	171838	2	+	537	ATP synthase delta chain (EC 3.6.3.14)	-none-	D23_1c0184 Neut_0274
fig/66666666.60966.peg.179	CDS	171851	173392	2	+	1542	ATP synthase alpha chain (EC 3.6.3.14)	-none-	D23_1c0185 Neut_0275

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld	C91 Alias
fig/66666666.60966.peg.180	CDS	173396	174280	2	+	885	ATP synthase gamma chain (EC 3.6.3.14)	-none-	D23_1c0186	Neut_0276
fig/66666666.60966.peg.181	CDS	174311	175693	2	+	1383	ATP synthase beta chain (EC 3.6.3.14)	-none-	D23_1c0187	Neut_0277
fig/66666666.60966.peg.182	CDS	175842	176141	3	+	300	ATP synthase epsilon chain (EC 3.6.3.14)	-none-	D23_1c0188	Neut_0278
fig/66666666.60966.peg.183	CDS	176389	177765	1	+	1377	N-acetylglucosamine-1-phosphate uridylyltransferase (EC 2.7.7.23) Glucosamine-1-phosphate N-acetyltransferase (EC 2.3.1.157)	Peptidoglycan Biosynthesis; Peptidoglycan Biosynthesis; Sialic Acid Metabolism; Sialic Acid Metabolism; Transcription repair cluster; Transcription repair cluster; UDP-N-acetylmuramate from Fructose-6-phosphate Biosynthesis; UDP-N-acetylmuramate from Fructose-6-phosphate Biosynthesis	D23_1c0189	Neut_0279
fig/66666666.60966.peg.184	CDS	177805	179652	1	+	1848	Glucosamine-fructose-6-phosphate aminotransferase [isomerizing] (EC 2.6.1.16)	Sialic Acid Metabolism; UDP-N-acetylmuramate from Fructose-6-phosphate Biosynthesis	D23_1c0190	Neut_0280
fig/66666666.60966.peg.185	CDS	179795	180520	2	+	726	FIG000859; hypothetical protein YebC	Fructose-6-phosphate Biosynthesis Riboflavin, FMN and FAD metabolism in plants; RuvABC plus a hypothetical	D23_1c0191	Neut_0281
fig/66666666.60966.peg.186	CDS	180523	181059	1	+	537	Crossover junction endonuclease RuvC (EC 3.1.22.4)	RuvABC plus a hypothetical	D23_1c0192	Neut_0282
fig/66666666.60966.peg.187	CDS	181056	181640	3	+	585	Holliday junction DNA helicase RuvA	RuvABC plus a hypothetical	D23_1c0193	Neut_0283
fig/66666666.60966.peg.188	CDS	181659	182699	3	+	1041	Holliday junction DNA helicase RuvB	RuvABC plus a hypothetical	D23_1c0194	Neut_0284
fig/66666666.60966.peg.189	CDS	182760	183173	3	+	414	4-hydroxybenzoyl-CoA thioesterase family active site	Ton and Tol transport systems	D23_1c0195	Neut_0285
fig/66666666.60966.peg.190	CDS	183166	183870	1	+	705	MotA/TolQ/ExbB proton channel family protein	Ton and Tol transport systems	D23_1c0196	Neut_0286
fig/66666666.60966.peg.191	CDS	183867	184283	3	+	417	Tol biopolymer transport system, TolR protein	Ton and Tol transport systems	D23_1c0197	Neut_0287

SUPPLEMENTARY TABLE 1-continued

Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld	C91 Alias
fig/66666666.60966.peg.192	CDS	184304	185200	2	+	897	TolA protein	Ton and Tol transport systems	D23_1c0198	Neut_0288
fig/66666666.60966.peg.193	CDS	185239	186510	1	+	1272	tolB protein precursor; periplasmic protein involved in the tonB-independent uptake of group A colicins	Ton and Tol transport systems	D23_1c0199	Neut_0289
fig/66666666.60966.peg.194	CDS	186565	187086	1	+	522	18K peptidoglycan-associated outer membrane lipoprotein; Peptidoglycan-associated lipoprotein precursor; Outer membrane protein P6; OmpA/MotB precursor TPR repeat containing exported protein; Putative periplasmic protein contains a protein	Ton and Tol transport systems	D23_1c0200	Neut_0290
fig/66666666.60966.peg.195	CDS	187086	187907	3	+	822	protein contains a prenyltransferase domain	Ton and Tol transport systems	D23_1c0201	Neut_0291
fig/66666666.60966.peg.196	CDS	188060	188644	2	+	585	Queuosine Biosynthesis QueE Radical SAM	Queuosine-Archaosine Biosynthesis; RNA modification Bacteria	D23_1c0202	Neut_0292
fig/66666666.60966.peg.197	CDS	188666	189346	2	+	681	Queuosine Biosynthesis QueC ATPase	Queuosine-Archaosine Biosynthesis; RNA modification Bacteria	D23_1c0203	Neut_0293
fig/66666666.60966.peg.198	CDS	189700	189347	-1	-	354	Dihydrocyclopterin aldolase (EC 4.1.2.25)	Folate Biosynthesis	D23_1c0204	Neut_0294
fig/66666666.60966.peg.199	CDS	189786	190388	3	+	603	Acyl-phosphate:glycerol-3-phosphate O-acyltransferase PlsY	Glycerolipid and Glycerophospholipid Metabolism in Bacteria	D23_1c0205	Neut_0295
fig/66666666.60966.peg.200	CDS	191422	190406	-1	-	1017	Tsai/Kae1/Orf7 protein, required for threonylcarbamoyl/adenosine t(6)A37 formation in tRNA	Bacterial RNA-metabolizing Zn-dependent hydrolases; Macromolecular synthesis operon; YgiD and YeaZ	D23_1c0206	Neut_0296
fig/66666666.60966.peg.201	CDS	191698	191910	1	+	213	SSU ribosomal protein S21p	Macromolecular synthesis operon	D23_1c0207	Neut_0297
fig/66666666.60966.peg.202	CDS	191984	192391	2	+	408	Transamidase GatB domain protein	Macromolecular synthesis operon	D23_1c0208	Neut_0298
fig/66666666.60966.peg.203	CDS	192486	194279	3	+	1794	DNA primase (EC 2.7.7.—)	CBS-349161.4.peg.2417; Macromolecular synthesis operon	D23_1c0209	Neut_0299

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 GbkId	C91 Alias
fig/66666666.60966.psg.204	CDS	194461	196710	1	+	2250	RNA polymerase sigma factor RpoD	CBSS-349161.4.psg.2417; Flagellum; Macromolecular synthesis operon; Transcription initiation, bacterial sigma factors	D23_1c0210	Neut_0300
fig/66666666.60966.psg.206	CDS	197605	197180	-1	-	426	Mobile element protein	-none-	D23_1c0212	Neut_0357
fig/66666666.60966.psg.207	CDS	198088	198819	1	+	732	Mobile element protein	-none-	D23_1c0213	NA
fig/66666666.60966.psg.209	CDS	200235	199564	-3	-	672	transposase and inactivated derivatives	-none-	D23_1c0214	Neut_2192
fig/66666666.60966.psg.210	CDS	200398	200210	-1	-	189	hypothetical protein	-none-	D23_1c0215	NA
fig/66666666.60966.psg.211	CDS	200852	200995	2	+	144	Mobile element protein	-none-	D23_1c0216	Neut_0978
fig/66666666.60966.psg.212	CDS	201848	200970	-2	-	879	Mobile element protein	-none-	D23_1c0217	Neut_1720
fig/66666666.60966.psg.213	CDS	202240	201947	-1	-	294	Mobile element protein	-none-	D23_1c0218	Neut_1719
fig/66666666.60966.psg.214	CDS	202367	203209	2	+	843	Mobile element protein	-none-	D23_1c0219	Neut_1524
fig/66666666.60966.psg.215	CDS	203592	203461	-3	-	132	Phage Rha protein	-none-	D23_1c0220	NA
fig/66666666.60966.psg.216	CDS	203906	203571	-2	-	336	Mobile element protein	-none-	D23_1c0221	Neut_2450
fig/66666666.60966.psg.218	CDS	204442	204113	-1	-	330	hypothetical protein	Soluble cytochromes and functionally related electron carriers	D23_1c0222	Neut_2449
fig/66666666.60966.psg.219	CDS	205381	204746	-1	-	636	Cytochrome c4	-none-	D23_1c0223	Neut_0305
fig/66666666.60966.psg.220	CDS	205494	206096	3	+	603	FIG00859469: hypothetical protein	-none-	D23_1c0224	Neut_0306
fig/66666666.60966.psg.221	CDS	206204	207016	2	+	813	Methionine aminopeptidase (EC 3.4.11.18)	CBSS-312309.3.psg.1965; Translation termination factors	D23_1c0225	Neut_0307
fig/66666666.60966.psg.222	CDS	207076	207840	1	+	765	Ribonuclease PH (EC 2.7.7.56)	bacterial Heat shock dnaK gene cluster extended;	D23_1c0226	Neut_0308
fig/66666666.60966.psg.223	CDS	207825	208439	3	+	615	Xanthosine/inosine triphosphate pyrophosphatase; HAM1-like protein	 tRNA processing CBSS-630.2.psg.3360; Heat shock dnaK gene cluster extended	D23_1c0227	Neut_0309
fig/66666666.60966.psg.224	CDS	208474	209709	1	+	1236	Radical SAM family enzyme, similar to coproporphyrinogen III oxidase, oxygen-independent, clustered with nucleoside-triphosphatase RdgB	CBSS-630.2.psg.3360; Heat shock dnaK gene cluster extended; Heme and Siroheme Biosynthesis; Queuosine-Archaeosine	D23_1c0228	Neut_0310
fig/66666666.60966.psg.225	CDS	209741	211540	2	+	1800	Multicopper oxidase	Biosynthesis	D23_1c0229	Neut_0311
fig/66666666.60966.psg.226	CDS	211537	212352	1	+	816	Copper resistance protein B	Copper homeostasis Copper homeostasis	D23_1c0230	Neut_0312

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 GbkId	C91 Alias
fig/66666666.60966.peg.227	CDS	213327	212398	-3	-	930	hypothetical protein	-none-	D23_1c0231	Neut_0313
fig/66666666.60966.peg.228	CDS	213918	213340	-3	-	579	LemA PROTEIN	-none-	D23_1c0232	Neut_1392
fig/66666666.60966.peg.229	CDS	214368	214553	3	+	186	Mobile element protein	-none-	D23_1c0233	Neut_2500
fig/66666666.60966.peg.230	CDS	214610	215206	2	+	597	Mobile element protein	-none-	D23_1c0234	Neut_1375
fig/66666666.60966.peg.231	CDS	215510	215623	2	+	114	hypothetical protein	-none-	D23_1c0235	NA
fig/66666666.60966.peg.232	CDS	215668	215847	1	+	180	hypothetical protein	-none-	D23_1c0236	Neut_0314
fig/66666666.60966.peg.233	CDS	217943	216069	-2	-	1875	Glutathione-regulated potassium-efflux system ATP-binding protein	Glutathione-regulated potassium-efflux system and associated functions; Potassium homeostasis	D23_1c0237	Neut_0315
fig/66666666.60966.peg.234	CDS	218233	219195	1	+	963	Mobile element protein	-none-	D23_1c0238	Neut_1862
fig/66666666.60966.peg.235	CDS	219960	219271	-3	-	690	InterPro IPR001687	-none-	D23_1c0239	NA
fig/66666666.60966.peg.236	CDS	220560	222266	3	+	1707	Glutathione-regulated potassium-efflux system protein KefB	Glutathione-regulated potassium-efflux system and associated functions	D23_1c0241	Neut_0318
fig/66666666.60966.peg.237	CDS	222848	223903	2	+	1056	SAM-dependent methyltransferase	-none-	D23_1c0242	Neut_0320
fig/66666666.60966.peg.238	CDS	223971	224243	3	+	273	Phosphate transport system permease protein PstA (TC 3.A.1.7.1)	High affinity phosphate transporter and control of PHO regulon; Phosphate metabolism	D23_1c0244	Neut_0321
fig/66666666.60966.peg.239	CDS	225095	224421	-2	-	675	tRNA (guanine46-N7-)-methyltransferase (EC 2.1.1.33)	RNA methylation; RNA modification	D23_1c0245	Neut_0322
fig/66666666.60966.peg.240	CDS	225934	225128	-1	-	807	Thiazole biosynthesis protein ThiG	Bacteria	D23_1c0246	Neut_0323
fig/66666666.60966.peg.241	CDS	226194	225994	-3	-	201	Sulfur carrier protein ThiS	Thiamin biosynthesis	D23_1c0247	Neut_0324
fig/66666666.60966.peg.242	CDS	226421	227008	2	+	588	FIG008443: hypothetical protein	Thiamin biosynthesis	D23_1c0249	Neut_0325
fig/66666666.60966.peg.243	CDS	227005	228537	1	+	1533	FIG139976: hypothetical protein	CBSS-208964.1.peg.1768	D23_1c0250	Neut_0326
fig/66666666.60966.peg.244	CDS	228587	229492	2	+	906	hypothetical protein FIG002781: Alpha-L-glutamate ligase family protein	CBSS-208964.1.peg.1768	D23_1c0251	Neut_0327
fig/66666666.60966.peg.245	CDS	231155	229677	-2	-	1479	Cardiolipin synthetase (EC 2.7.8.—)	Cardiolipin synthesis; Glycerolipid and Glycerophospholipid Metabolism in Bacteria	D23_1c0252	Neut_0328
fig/66666666.60966.peg.246	CDS	231229	231411	1	+	183	hypothetical protein	-none-	D23_1c0253	NA
fig/66666666.60966.peg.248	CDS	232352	231813	-2	-	540	Urea channel Urel	Urea decomposition	D23_1c0255	Neut_0329
fig/66666666.60966.peg.249	CDS	232767	232585	-3	-	183	hypothetical protein	-none-	D23_1c0256	NA
fig/66666666.60966.peg.251	CDS	233588	234748	2	+	1161	FIG00855934: hypothetical protein	-none-	D23_1c0259	Neut_0331

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 GbkId	C91 Alias
fig 66666666.60966.peg.252	CDS	235271	234831	-2	-	441	Mobile element protein	-none-	D23_1c0260	Neut_0332
fig 66666666.60966.peg.253	CDS	235792	235397	-1	-	396	NAD-dependent glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12)	Calvin-Benson cycle; Glycolysis and Gluconeogenesis; Pyridoxin (Vitamin B6) Biosynthesis	D23_1c0261	Neut_0333
fig 66666666.60966.peg.254	CDS	235832	236260	2	+	429	hypothetical protein	-none-	D23_1c0261	Neut_0333
fig 66666666.60966.peg.255	CDS	237167	236592	-2	-	576	Flagellar basal-body P-ring formation protein	Flagellum	D23_1c0262	Neut_0334
fig 66666666.60966.peg.256	CDS	237299	237415	2	+	117	hypothetical protein	-none-	D23_1c0264	NA
fig 66666666.60966.peg.257	CDS	237436	237924	1	+	489	Flagellar basal-body rod protein FigB	Flagellum; Flagellum in <i>Campylobacter</i>	D23_1c0265	Neut_0335
fig 66666666.60966.peg.258	CDS	237930	238334	3	+	405	Flagellar basal-body rod protein FigC	Flagellum; Flagellum in <i>Campylobacter</i>	D23_1c0266	Neut_0336
fig 66666666.60966.peg.259	CDS	238347	239021	3	+	675	Flagellar basal-body rod modification protein	Flagellar motility; Flagellum	D23_1c0267	Neut_0337
fig 66666666.60966.peg.260	CDS	239037	240296	3	+	1260	Flagellar hook protein	Flagellum	D23_1c0268	Neut_0338
fig 66666666.60966.peg.261	CDS	240337	241080	1	+	744	Flagellar basal-body rod protein FigF	Flagellum	D23_1c0269	Neut_0339
fig 66666666.60966.peg.262	CDS	241119	241901	3	+	783	Flagellar basal-body rod protein FigG	Flagellum	D23_1c0270	Neut_0340
fig 66666666.60966.peg.263	CDS	242034	242828	3	+	795	Flagellar L-ring protein	Flagellar motility; Flagellum	D23_1c0271	Neut_0341
fig 66666666.60966.peg.264	CDS	242850	243977	3	+	1128	Flagellar P-ring protein	Flagellum	D23_1c0272	Neut_0342
fig 66666666.60966.peg.265	CDS	243991	244998	1	+	1008	Flagellar protein FigJ	Flagellum	D23_1c0273	Neut_0343
fig 66666666.60966.peg.266	CDS	245257	246660	1	+	1404	[peptidoglycan hydrolase] (EC 3.2.1.—)	Flagellum	D23_1c0274	Neut_0344
fig 66666666.60966.peg.267	CDS	246638	247588	2	+	951	Flagellar hook-associated protein FigK	Flagellum	D23_1c0275	Neut_0345
fig 66666666.60966.peg.268	CDS	247665	248210	3	+	546	Flagellar hook-associated protein FigL	-none-	D23_1c0276	Neut_0346
fig 66666666.60966.peg.269	CDS	249330	248200	-3	-	1131	hypothetical protein	-none-	D23_1c0277	Neut_0347
fig 66666666.60966.peg.270	CDS	249439	249960	1	+	522	hypothetical protein	-none-	D23_1c0278	Neut_0348
fig 66666666.60966.peg.271	CDS	249932	250513	2	+	582	GCN5-related N-acetyltransferase	-none-	D23_1c0279	Neut_0349
fig 66666666.60966.peg.272	CDS	250589	250861	2	+	273	FIG001341; Probable Fe(2+)-trafficking protein YggX	Heat shock dnaK gene cluster extended	D23_1c0280	Neut_0350
fig 66666666.60966.peg.273	CDS	250912	253038	1	+	2127	Polyposphate kinase (EC 2.7.4.1)	High affinity phosphate transporter and control	D23_1c0281	Neut_0351

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.											
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld	C91 Alias	
fig/66666666.60966.peg.274	CDS	254786	253059	-2	-	1728	Sulfate permease	of PHO regulon; Phosphate metabolism; Polyphosphate; Purine conversions Cysteine Biosynthesis -none- -none- -none- -none- cAMP signaling in bacteria	D23_1c0282	Neut_0352	
fig/66666666.60966.peg.275	CDS	255133	254783	-1	-	351	Transcriptional regulator, ArsR family		D23_1c0283	Neut_0353	
fig/66666666.60966.peg.277	CDS	256153	255827	-1	-	327	hypothetical protein		D23_1c0285	Neut_0355	
fig/66666666.60966.peg.278	CDS	256608	257603	3	+	996	hypothetical protein		D23_1c0286	Neut_0356	
fig/66666666.60966.peg.279	CDS	258986	257739	-2	-	1248	Mobile element protein		D23_1c0287	Neut_0357	
fig/66666666.60966.peg.280	CDS	259004	259126	2	+	123	patatin family protein		D23_1c0288	Neut_1317	
fig/66666666.60966.peg.281	CDS	259254	259123	-3	-	132	cAMP-binding proteins-catabolite gene		D23_1c0289	NA	
							activator and regulatory subunit of cAMP-dependent protein kinases				
fig/66666666.60966.peg.282	CDS	259543	260031	1	+	489	Cytochrome c'		-none-	D23_1c0291	NA
fig/66666666.60966.peg.283	CDS	260060	260947	2	+	888	Putative dihenne cytochrome c-553		Soluble cytochromes and functionally related electron carriers	D23_1c0292	Neut_1381
fig/66666666.60966.peg.285	CDS	261917	261708	-2	-	210	hypothetical protein	-none-	D23_1c0294	Neut_0363	
fig/66666666.60966.peg.288	CDS	262640	262440	-2	-	201	Mobile element protein	-none-	D23_1c0296	Neut_1696	
fig/66666666.60966.peg.289	CDS	263106	264041	3	+	936	hypothetical protein	-none-	D23_1c0297	NA	
fig/66666666.60966.peg.290	CDS	264137	265633	2	+	1497	SHH503 protein	-none-	D23_1c0298	NA	
fig/66666666.60966.peg.294	CDS	266897	266760	-2	-	138	hypothetical protein	-none-	D23_1c0300	NA	
fig/66666666.60966.peg.295	CDS	267026	267370	2	+	345	COGs COG3339	-none-	D23_1c0301	Neut_0371	
fig/66666666.60966.peg.297	CDS	268862	267765	-2	-	1098	L-lactate dehydrogenase (EC 1.1.2.3)	Lactate utilization; Respiratory dehydrogenases 1	D23_1c0302	Neut_0372	
fig/66666666.60966.peg.298	CDS	269655	268972	-3	-	684	Iron-uptake factor PiuC	-none-	D23_1c0303	Neut_0373	
fig/66666666.60966.peg.299	CDS	271893	269683	-3	-	2211	TonB-dependent siderophore receptor protein of unknown function DUF81	Ton and Tol transport systems	D23_1c0304	Neut_0374	
fig/66666666.60966.peg.301	CDS	272682	273740	3	+	1059	hypothetical protein	-none-	D23_1c0306	Neut_0377	
fig/66666666.60966.peg.302	CDS	273758	274108	2	+	351	hypothetical protein	-none-	D23_1c0307	NA	
fig/66666666.60966.peg.303	CDS	274775	274182	-2	-	594	InterPro IPR001226	-none-	D23_1c0308	Neut_0379	
fig/66666666.60966.peg.304	CDS	274944	274792	-3	-	153	COGs COG0790	-none-	D23_1c0309	NA	
fig/66666666.60966.peg.305	CDS	276110	274986	-2	-	1125	dNTP triphosphohydrolase, broad substrate specificity, subgroup 2	Purine conversions; dNTP triphosphohydrolase protein family	D23_1c0310	Neut_0380	
fig/66666666.60966.peg.306	CDS	277212	276103	-3	-	1110	3-dehydroquinase synthase (EC 4.2.3.4)	Chlorismate Synthesis; Common Pathway For Synthesis of Aromatic Compounds	D23_1c0311	Neut_0381	

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 GbkId	C91 Alias
fig/66666666.60966.peg.308	CDS	277726	277896	1	+	171	hypothetical protein	(DAHP synthase to chorismate) -none- Chorismate Synthesis; Common Pathway For Synthesis of Aromatic Compounds (DAHP synthase to chorismate) -none- Phage replication	D23_1c0312	Neut_0382
fig/66666666.60966.peg.307	CDS	277692	277261	-3	-	432	Shikimate kinase I (EC 2.7.1.71)		D23_1c0312	Neut_0382
fig/66666666.60966.peg.309	CDS	279343	277982	-1	-	1362	Putative protease		D23_1c0313	Neut_0383
fig/66666666.60966.peg.310	CDS	279362	282934	2	+	3573	DNA polymerase III alpha subunit (EC 2.7.7.7)		D23_1c0314	Neut_0384
fig/66666666.60966.peg.311	CDS	283139	282948	-2	-	192	putative transmembrane protein	-none-	D23_1c0315	Neut_0385
fig/66666666.60966.peg.312	CDS	283290	284243	3	+	954	tRNA dimethylallyltransferase (EC 2.5.1.75)	tRNA modification Bacteria; tRNA processing	D23_1c0316	Neut_0386
fig/66666666.60966.peg.313	CDS	284258	284401	2	+	144	hypothetical protein	-none- Possible RNA degradation cluster -none- -none- RNA methylation	D23_1c0317	NA
fig/66666666.60966.peg.314	CDS	286041	284776	-3	-	1266	Two component, sigma54 specific, transcriptional regulator, Fis family		D23_1c0320	Neut_0387
fig/66666666.60966.peg.315	CDS	286328	286191	-2	-	138	hypothetical protein	-none-	D23_1c0321	NA
fig/66666666.60966.peg.316	CDS	288462	286330	-3	-	2133	Nitrogen regulation protein NtrY (EC 2.7.3.-)	Possible RNA degradation cluster	D23_1c0322	Neut_0388
fig/66666666.60966.peg.317	CDS	289077	288514	-3	-	564	Probable proline rich signal peptide protein	-none-	D23_1c0323	Neut_0389
fig/66666666.60966.peg.318	CDS	290401	289121	-1	-	1281	16S rRNA (cytosine(967)-C(5))-methyltransferase (EC 2.1.1.176) ## SSU rRNA m5C967	RNA methylation	D23_1c0324	Neut_0390
fig/66666666.60966.peg.319	CDS	291388	290414	-1	-	975	Methionyl-tRNA formyltransferase (EC 2.1.2.9)	Translation initiation factors bacterial	D23_1c0325	Neut_0391
fig/66666666.60966.peg.320	CDS	291933	291427	-3	-	507	Peptide deformylase (EC 3.5.1.88)	Bacterial RNA-metabolizing Zn-dependent hydrolases; Translation termination factors bacterial	D23_1c0326	Neut_0392
fig/66666666.60966.peg.321	CDS	292108	293136	1	+	1029	Uncharacterized protein with LysM domain, COG1652	-none-	D23_1c0327	Neut_0393
fig/66666666.60966.peg.322	CDS	293235	294356	3	+	1122	Rossmann fold nucleotide-binding	-none-	D23_1c0328	Neut_0394

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.									
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld C91 Alias
fig/66666666.60966.peg.323	CDS	294438	294896	3	+	459	protein Smf possibly involved in DNA uptake	-none-	D23_1c0329 Neut_0395
fig/66666666.60966.peg.324	CDS	295024	297522	1	+	2499	Protein of unknown function Smg DNA topoisomerase III, Burkholderia type (EC 5.99.1.2)	DNA topoisomerases, Type I, ATP-independent	D23_1c0330 Neut_0396
fig/66666666.60966.peg.325	CDS	297826	297575	-1	-	252	FIG00858730: hypothetical protein	-none-	D23_1c0331 Neut_0397
fig/66666666.60966.peg.326	CDS	298176	299477	3	+	1302	5-Enolpyruvylshikimate-3-phosphate synthase (EC 2.5.1.19)	Chorismate Synthesis; Common Pathway For Synthesis of Aromatic Compounds (DAHP synthase to chorismate)	D23_1c0333 Neut_0398
fig/66666666.60966.peg.327	CDS	299557	300228	1	+	672	Cytidylate Kinase (EC 2.7.4.14)	-none-	D23_1c0335 Neut_0399
fig/66666666.60966.peg.328	CDS	300339	302051	3	+	1713	SSU ribosomal protein Slp	-none-	D23_1c0336 Neut_0400
fig/66666666.60966.peg.329	CDS	302061	302378	3	+	318	Integration host factor beta subunit	DNA structural proteins, bacterial	D23_1c0337 Neut_0401
fig/66666666.60966.peg.330	CDS	302493	302368	-3	-	126	hypothetical protein	-none-	D23_1c0338 NA
fig/66666666.60966.peg.331	CDS	302902	303597	1	+	696	Orotidine 5'-phosphate decarboxylase (EC 4.1.1.23)	De Novo Pyrimidine Synthesis; Riboflavin synthesis cluster	D23_1c0339 Neut_0402
fig/66666666.60966.peg.332	CDS	304632	303592	-3	-	1041	Squalene synthase (EC 2.5.1.21)	Hopanes	D23_1c0340 Neut_0403
fig/66666666.60966.peg.333	CDS	305907	304654	-3	-	1254	Diaminopimelate decarboxylase (EC 4.1.1.20)	Lysine Biosynthesis DAP Pathway, GJO scratch	D23_1c0341 Neut_0404
fig/66666666.60966.peg.334	CDS	306026	305904	-2	-	123	hypothetical protein	-none-	D23_1c0342 NA
fig/66666666.60966.peg.335	CDS	306654	306052	-3	-	603	Probable lipoprotein	-none-	D23_1c0343 Neut_0405
fig/66666666.60966.peg.336	CDS	307556	306651	-2	-	906	ABC-type transport system involved in resistance to organic solvents, periplasmic component	-none-	D23_1c0344 Neut_0406
fig/66666666.60966.peg.337	CDS	308341	307583	-1	-	759	Inositol-1-monophosphatase (EC 3.1.3.25)	-none-	D23_1c0345 Neut_0407
fig/66666666.60966.peg.339	CDS	308500	309207	1	+	708	tRNA:Gln32/Urn32 methyltransferase	RNA methylation; RNA modification	D23_1c0346 Neut_0408
fig/66666666.60966.peg.340	CDS	309905	309291	-2	-	615	Glutathione S-transferase family protein	Bacteria Glutathione: Non-redox reactions; Scaffold proteins for [4Fe-4S]	D23_1c0347 Neut_0409

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld	C91 Alias
fig/66666666.60966.peg.341	CDS	310042	311418	1	+	1377	Adenylosuccinate lyase (EC 4.3.2.2)	cluster assembly (MRP family) De Novo Purine Biosynthesis; Purine conversions	D23_1c0348	Neut_0410
fig/66666666.60966.peg.342	CDS	311556	312146	3	+	591	Heat shock protein GrpE	GroEL GroES; Heat shock dnaK gene cluster extended; Protein chaperones	D23_1c0349	Neut_0411
fig/66666666.60966.peg.343	CDS	312210	314153	3	+	1944	Chaperone protein DnaK	GroEL GroES; Heat shock dnaK gene cluster extended; Protein chaperones	D23_1c0350	Neut_0412
fig/66666666.60966.peg.344	CDS	314344	315453	1	+	1110	Chaperone protein DnaJ	GroEL GroES; Heat shock dnaK gene cluster extended; Protein chaperones	D23_1c0351	Neut_0413
fig/66666666.60966.peg.345	CDS	318894	315550	-3	-	3345	Potassium efflux system KefA protein/Small-conductance mechanosensitive channel	Potassium homeostasis	D23_1c0352	Neut_0414
fig/66666666.60966.peg.346	CDS	319923	319357	-3	-	567	Transcriptional regulator, TetR family	-none-	D23_1c0353	Neut_0415
fig/66666666.60966.peg.347	CDS	321174	319990	-3	-	1185	InterPro IPR001327 COGs COG2072	-none-	D23_1c0354	Neut_0416
fig/66666666.60966.peg.348	CDS	321778	321236	-1	-	543	hypothetical protein	-none-	D23_1c0355	Neut_0417
fig/66666666.60966.peg.349	CDS	322196	322363	2	+	168	hypothetical protein	-none-	D23_1c0356	NA
fig/66666666.60966.peg.350	CDS	325140	322522	-3	-	2619	Membrane alanine aminopeptidase N (EC 3.4.11.2)	Aminopeptidases (EC 3.4.11.—)	D23_1c0357	Neut_0418
fig/66666666.60966.peg.351	CDS	325139	325255	2	+	117	hypothetical protein	-none-	D23_1c0358	NA
fig/66666666.60966.peg.352	CDS	326547	325213	-3	-	1335	Peptide methionine sulfoxide reductase MsrA (EC 1.8.4.11)/Thiol:disulfide oxidoreductase associated with MetSO reductase/Peptide Methionine sulfoxide reductase MsrB (EC 1.8.4.12)	Peptide methionine sulfoxide reductase; Peptide methionine sulfoxide reductase; Peptide methionine sulfoxide reductase	D23_1c0359	Neut_0419
fig/66666666.60966.peg.354	CDS	326909	329791	2	+	2883	Diguanylate cyclase/phosphodiesterase domain 2 (EAL)	-none-	D23_1c0360	Neut_0422
fig/66666666.60966.peg.355	CDS	331130	329874	-2	-	1257	FIG00858721: hypothetical protein	-none-	D23_1c0361	Neut_0423

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld	C91 Alias
fig/66666666.60966.peg.356	CDS	331369	332730	1	+	1362	O-acetylhomoserine sulphydrylase (EC 2.5.1.49)/O-succinylhomoserine sulphydrylase (EC 2.5.1.48)	Methionine Biosynthesis; Methionine Biosynthesis	D23_1c0363	Neut_0424
fig/66666666.60966.peg.357	CDS	334115	332718	-2	-	1398	NmrS protein involved in response to NO	Denitrification; Nitrosative stress; Oxidative stress	D23_1c0364	Neut_0425
fig/66666666.60966.peg.358	CDS	334992	334066	-3	-	927	Serine acetyltransferase (EC 2.3.1.30)	Cysteine Biosynthesis; Methionine Biosynthesis	D23_1c0365	Neut_0426
fig/66666666.60966.peg.360	CDS	335392	336399	1	+	1008	Glucokinase (EC 2.7.1.2)	Glycolysis and Gluconeogenesis	D23_1c0367	Neut_0427
fig/66666666.60966.peg.361	CDS	336414	337073	3	+	660	Probable transmembrane protein	-none-	D23_1c0368	Neut_0428
fig/66666666.60966.peg.362	CDS	337412	337101	-2	-	312	FIG00858769; hypothetical protein	-none-	D23_1c0369	Neut_0429
fig/66666666.60966.peg.363	CDS	338169	337483	-3	-	687	6-phosphogluconolactonase (EC 3.1.1.31), eukaryotic type	Pentose phosphate pathway	D23_1c0370	Neut_0430
fig/66666666.60966.peg.364	CDS	338807	338151	-2	-	657	hydrolase, haloacid dehalogenase-like family	-none-	D23_1c0371	Neut_0431
fig/66666666.60966.peg.365	CDS	339746	338814	-2	-	933	NAD-dependent epimerase/dehydratase	CBSS-296591.1.peg.2330	D23_1c0372	Neut_0432
fig/66666666.60966.peg.366	CDS	340674	339739	-3	-	936	D-3-phosphoglycerate dehydrogenase (EC 1.1.1.95)	Glycine and Serine Utilization; Pyridoxin (Vitamin B6) Biosynthesis; Serine Biosynthesis	D23_1c0373	Neut_0433
fig/66666666.60966.peg.367	CDS	341432	340671	-2	-	762	2,4-dihydroxyhept-2-ene-1,7-dioic acid aldolase (EC 4.1.2.--)	-none-	D23_1c0374	Neut_0434
fig/66666666.60966.peg.368	CDS	342202	341444	-1	-	759	3-deoxy-manno-octulosonate cytidyltransferase (EC 2.7.7.38)	KDO2-Lipid A biosynthesis cluster 2	D23_1c0375	Neut_0435
fig/66666666.60966.peg.370	CDS	342384	342509	3	+	126	hypothetical protein	-none-	D23_1c0376	NA
fig/66666666.60966.peg.371	CDS	342506	343585	2	+	1080	Glycosyl transferase, group 2 family protein	-none-	D23_1c0377	Neut_0436
fig/66666666.60966.peg.372	CDS	343660	344931	1	+	1272	O-antigen ligase	-none-	D23_1c0378	Neut_0437
fig/66666666.60966.peg.373	CDS	344931	345620	3	+	690	O-methyltransferase family protein [C1]	-none-	D23_1c0379	Neut_0438
fig/66666666.60966.peg.374	CDS	345673	345930	1	+	258	FIG00859064; hypothetical protein	-none-	D23_1c0380	Neut_0439

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld	C91 Alias
fig/66666666.60966.peg.375	CDS	345980	346498	2	+	519	Mit4354 like protein	-none-	D23_1c0381	Neut_0440
fig/66666666.60966.peg.376	CDS	346511	346858	2	+	348	Arsenate reductase (EC 1.20.4.1)	Anaerobic respiratory reductases; Transcription repair cluster	D23_1c0382	Neut_0441
fig/66666666.60966.peg.377	CDS	347305	346916	-1	-	390	LSU ribosomal protein L19p	-none-	D23_1c0383	Neut_0442
fig/66666666.60966.peg.378	CDS	348122	347277	-2	-	846	tRNA (Guanine37-N1)-methyltransferase (EC 2.1.1.31)	RNA methylation; Ribosome biogenesis bacterial; tRNA modification	D23_1c0384	Neut_0443
fig/66666666.60966.peg.379	CDS	348636	348130	-3	-	507	16S rRNA processing protein RimM	Bacteria	D23_1c0385	Neut_0444
fig/66666666.60966.peg.381	CDS	349157	350437	2	+	1281	Glycolate dehydrogenase (EC 1.1.99.14), iron-sulfur subunit GicF	Ribosome biogenesis bacterial	D23_1c0386	Neut_0446
fig/66666666.60966.peg.382	CDS	350492	351118	2	+	627	Uncharacterized hydroxylase PA0655	Glycolate, glyoxylate interconversions; Photorespiration (oxidative C2 cycle)	D23_1c0387	Neut_0447
fig/66666666.60966.peg.383	CDS	351152	351712	2	+	561	UPF0301 protein YqgE	Ubi	D23_1c0388	Neut_0448
fig/66666666.60966.peg.384	CDS	351705	352178	3	+	474	Putative Holliday junction resolvase (EC 3.1.-.-.-)	-none-	D23_1c0389	Neut_0449
fig/66666666.60966.peg.385	CDS	352165	352668	1	+	504	Uracil phosphoribosyltransferase (EC 2.4.2.9)/Pyrimidine operon regulatory protein PyrR	De Novo Pyrimidine Synthesis; De Novo Pyrimidine conversions	D23_1c0390	Neut_0450
fig/66666666.60966.peg.386	CDS	352856	353806	2	+	951	Aspartate carbamoyltransferase (EC 2.1.3.2)	De Novo Pyrimidine Synthesis	D23_1c0391	Neut_0451
fig/66666666.60966.peg.387	CDS	353822	355093	2	+	1272	Dihydroorotase (EC 3.5.2.3)	De Novo Pyrimidine Synthesis; Zinc regulated enzymes	D23_1c0392	Neut_0452
fig/66666666.60966.peg.388	CDS	355217	357322	2	+	2106	Oligopeptidase A (EC 3.4.24.70)	Protein degradation	D23_1c0393	Neut_0453
fig/66666666.60966.peg.389	CDS	357558	358709	3	+	1152	Carbamoyl-phosphate synthase small chain (EC 6.3.5.5)	De Novo Pyrimidine Synthesis; Macromolecular synthesis operon	D23_1c0394	Neut_0454
fig/66666666.60966.peg.390	CDS	358735	361932	1	+	3198	Carbamoyl-phosphate synthase large chain (EC 6.3.5.5)	De Novo Pyrimidine Synthesis; Macromolecular synthesis operon	D23_1c0395	Neut_0455
fig/66666666.60966.peg.391	CDS	362117	362593	2	+	477	Transcription elongation factor GreA	Transcription factors bacterial	D23_1c0396	Neut_0456

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 GbkId	C91 Alias
fig/66666666.60966.peg.392	CDS	363579	362608	-3	-	972	ErkY/bisY/cfsY/nhG family protein	-none-	D23_1c0397	Neut_0457
fig/66666666.60966.peg.393	CDS	364226	366052	2	+	1827	Long-chain-fatty-acid--CoA ligase (EC 6.2.1.3)	Biotin biosynthesis; Biotin synthesis cluster; Fatty acid metabolism cluster	D23_1c0398	Neut_0458
fig/66666666.60966.peg.394	CDS	366141	367064	3	+	924	FIG010773; NAD-dependent epimerase/dehydratase phosphopantetheine-binding	-none-	D23_1c0399	Neut_0459
fig/66666666.60966.peg.395	CDS	367176	367430	3	+	255	Aminotransferase class II, serine palmitoyltransferase like (EC 2.3.1.50)	-none-	D23_1c0400	Neut_0460
fig/66666666.60966.peg.396	CDS	367430	368617	2	+	1188	COG1496; Uncharacterized conserved protein	-none-	D23_1c0401	Neut_0461
fig/66666666.60966.peg.397	CDS	368669	369427	2	+	759	Zinc transporter, ZIP family	-none-	D23_1c0402	Neut_0462
fig/66666666.60966.peg.398	CDS	369615	370427	3	+	813	Dolichyl-phosphate beta-D-mannosyltransferase (EC: 2.4.1.83)	-none-	D23_1c0403	Neut_0463
fig/66666666.60966.peg.399	CDS	373049	370434	-2	-	2616	FIG004453; protein YceG like	-none-	D23_1c0404	Neut_0464
fig/66666666.60966.peg.400	CDS	374173	373157	-1	-	1017	CBSS-323097.3.peg.2594; Cluster containing Alanyl-tRNA synthetase; tRNA modification	CBSS-323097.3.peg.2594;	D23_1c0405	Neut_0465
fig/66666666.60966.peg.401	CDS	374277	374140	-3	-	138		Bacteria	D23_1c0406	NA
fig/66666666.60966.peg.402	CDS	375542	374301	-2	-	1242		-none- Fatty Acid Biosynthesis FASII	D23_1c0407	Neut_0466
fig/66666666.60966.peg.403	CDS	375822	375577	-3	-	246	hypothetical protein	Fatty Acid Biosynthesis FASII; Glycerolipid and Glycerophospholipid Metabolism in Bacteria	D23_1c0408	Neut_0467
fig/66666666.60966.peg.405	CDS	376731	375988	-3	-	744	3-oxoacyl-[acyl-carrier protein] reductase (EC 1.1.1.100)	Fatty Acid Biosynthesis FASII	D23_1c0409	Neut_0468
fig/66666666.60966.peg.406	CDS	377726	376788	-2	-	939	Malonyl CoA-acyl carrier protein transacylase (EC 2.3.1.39)	Fatty Acid Biosynthesis FASII	D23_1c0410	Neut_0469
fig/66666666.60966.peg.407	CDS	378692	377730	-2	-	963	3-oxoacyl-[acyl-carrier-protein] synthase, KASIII (EC 2.3.1.41)	Fatty Acid Biosynthesis FASII	D23_1c0411	Neut_0470

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.									
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkid C91 Alias
fig/66666666.60966.peg.408	CDS	379722	378703	-3	-	1020	Phosphate-acyl-ACP acyltransferase PlsX	Glycerolipid and Glycerophospholipid Metabolism in Bacteria	D23_1c0412 Neut_0471
fig/66666666.60966.peg.409	CDS	379982	379800	-2	-	183	LSU ribosomal protein L32p	-none-	D23_1c0414 Neut_0472
fig/66666666.60966.peg.410	CDS	380510	380007	-2	-	504	COG1399 protein, clustered with ribosomal protein L32p	-none-	D23_1c0415 Neut_0473
fig/66666666.60966.peg.411	CDS	380534	381175	2	+	642	FIG146278: MafYcef/YhdE family protein	-none-	D23_1c0416 Neut_0474
fig/66666666.60966.peg.412	CDS	381292	381684	1	+	393	FIG00858587: hypothetical protein	-none-	D23_1c0417 Neut_0475
fig/66666666.60966.peg.413	CDS	381793	383217	1	+	1425	Heavy metal RND efflux outer membrane protein, CzcC family	Cobalt-zinc-cadmium resistance	D23_1c0418 Neut_0476
fig/66666666.60966.peg.414	CDS	383214	384710	3	+	1497	Cobalt/zinc-cadmium efflux RND transporter, membrane fusion protein, CzcB family	Cobalt-zinc-cadmium resistance	D23_1c0419 Neut_0477
fig/66666666.60966.peg.415	CDS	384811	388020	1	+	3210	Cobalt-zinc-cadmium resistance protein CzcA; Cation efflux system protein CusA	Cobalt-zinc-cadmium resistance; Cobalt-zinc-cadmium resistance	D23_1c0421 Neut_0478
fig/66666666.60966.peg.416	CDS	388294	388731	1	+	438	FIG00858457: hypothetical protein	-none-	D23_1c0423 Neut_0479
fig/66666666.60966.peg.417	CDS	388756	389043	1	+	288	FIG00858508: hypothetical protein	-none-	D23_1c0424 Neut_0480
fig/66666666.60966.peg.418	CDS	389040	389948	3	+	909	FIG00858931: hypothetical protein	-none-	D23_1c0425 Neut_0481
fig/66666666.60966.peg.419	CDS	389941	391068	1	+	1128	hypothetical protein	-none-	D23_1c0426 Neut_0482
fig/66666666.60966.peg.420	CDS	392521	391079	-1	-	1443	Mg/Co/Ni transporter	Magnesium transport	D23_1c0427 Neut_0483
fig/66666666.60966.peg.424	CDS	394723	393761	-1	-	963	MgR/CBS domain	-none-	D23_1c0430 Neut_1746
fig/66666666.60966.peg.425	CDS	394947	394834	-3	-	114	hypothetical protein	-none-	D23_1c0431 NA
fig/66666666.60966.peg.426	CDS	394946	395251	2	+	306	hypothetical protein	-none-	D23_1c0432 Neut_0486
fig/66666666.60966.peg.427	CDS	395968	395309	-1	-	660	COG1272: Predicted membrane protein	-none-	D23_1c0433 Neut_0487
fig/66666666.60966.peg.428	CDS	396481	396179	-1	-	303	hemolysin III homolog	-none-	D23_1c0434 Neut_0488
fig/66666666.60966.peg.430	CDS	397189	396863	-1	-	327	hypothetical protein	-none-	D23_1c0435 Neut_0490
fig/66666666.60966.peg.433	CDS	397653	398393	3	+	741	cAMP-binding proteins-catabolite gene activator and regulatory subunit of cAMP-dependent protein kinases	cAMP signaling in bacteria	D23_1c0436 Neut_0491

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld	C91 Alias
fig/66666666.60966.peg.434	CDS	398690	398424	-2	-	267	Putative lipoprotein	-none-	D23_1c0437	Neut_0492
fig/66666666.60966.peg.435	CDS	399146	398973	-2	-	174	hypothetical protein	-none-	D23_1c0438	NA
fig/66666666.60966.peg.436	CDS	399498	399373	-3	-	126	hypothetical protein	-none-	D23_1c0439	NA
fig/66666666.60966.peg.437	CDS	400841	399609	-2	-	1233	hypothetical protein	-none-	D23_1c0440	Neut_0494
fig/66666666.60966.peg.438	CDS	401592	400858	-3	-	735	Monofunctional biosynthetic peptidoglycan transglycosylase (EC 2.4.2.—)	Peptidoglycan Biosynthesis	D23_1c0441	Neut_0495
fig/66666666.60966.peg.439	CDS	402422	401589	-2	-	834	Shikimate 5-dehydrogenase I alpha (EC 1.1.1.25)	Chorismate Synthesis; Cluster containing Alanyl-tRNA synthetase; Common Pathway For Synthesis of Aromatic Compounds (DAHP synthase to chorismate)	D23_1c0442	Neut_0496
fig/66666666.60966.peg.440	CDS	403340	402456	-2	-	885	TonB protein	-none-	D23_1c0443	Neut_0497
fig/66666666.60966.peg.441	CDS	405237	403378	-3	-	1860	Exoribonuclease II (EC 3.1.13.1)	RNA processing and degradation, bacterial	D23_1c0444	Neut_0498
fig/66666666.60966.peg.443	CDS	405594	406985	3	+	1392	Glutamyl-tRNA synthetase (EC 6.1.1.17)	Heme and Siroheme Biosynthesis; tRNA aminoacylation, Glu and Gln	D23_1c0445	Neut_0499
fig/66666666.60966.peg.444	CDS	407045	410752	2	+	3708	5-methyltetrahydrofolate--homocysteine methyltransferase (EC 2.1.1.13)	Methionine Biosynthesis	D23_1c0446	Neut_0500
fig/66666666.60966.peg.445	CDS	410924	412267	2	+	1344	NADP-specific glutamate dehydrogenase (EC 1.4.1.4)	Arginine and Ornithine Degradation; Glutamate dehydrogenases; Glutamine, Glutamate, Aspartate and Asparagine Biosynthesis; Proline Synthesis	D23_1c0447	Neut_0501
fig/66666666.60966.peg.446	CDS	412461	414368	3	+	1908	Soluble lytic murein transglycosylase precursor (EC 3.2.1.—)	Murein Hydrolases	D23_1c0449	Neut_0502
fig/66666666.60966.peg.447	CDS	414379	415617	1	+	1239	tRNA nucleotidyltransferase (EC 2.7.7.21) (EC 2.7.7.25)	A cluster relating to Tryptophanyl-tRNA synthetase; Polyadenylation bacterial; tRNA nucleotidyltransferase	D23_1c0450	Neut_0503

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld	C91 Alias
fig/66666666.60966.peg.448	CDS	417316	415592	-1	-	1725	Phospholipid-lipopolysaccharide ABC transporter	N-linked Glycosylation in Bacteria	D23_1c0451	Neut_0504
fig/66666666.60966.peg.449	CDS	418171	417335	-1	-	837	Diaminopimelate epimerase (EC 5.1.1.7)	CBS-84588.1.peg.1247; Lysine Biosynthesis DAP Pathway, GJO scratch	D23_1c0452	Neut_0505
fig/66666666.60966.peg.450	CDS	418345	418193	-1	-	153	hypothetical protein	-none-	D23_1c0453	NA
fig/66666666.60966.peg.451	CDS	418574	419011	2	+	438	Predicted secretion system X protein GspG-like 3	Predicted secretion system X	D23_1c0454	Neut_0506
fig/66666666.60966.peg.452	CDS	419030	420214	2	+	1185	Predicted secretion system X protein GspF-like	Predicted secretion system X	D23_1c0455	Neut_0507
fig/66666666.60966.peg.453	CDS	420211	421905	1	+	1695	Predicted secretion system X protein GspE-like	Predicted secretion system X	D23_1c0456	Neut_0508
fig/66666666.60966.peg.454	CDS	421910	422731	2	+	822	Predicted secretion system X FIG084745; hypothetical protein	Predicted secretion system X	D23_1c0457	Neut_0509
fig/66666666.60966.peg.455	CDS	422772	423260	3	+	489	Predicted secretion system X transmembrane protein 1	Predicted secretion system X	D23_1c0458	Neut_0510
fig/66666666.60966.peg.472	CDS	440367	440753	3	+	387	Mobile element protein	-none-	D23_1c0476	Neut_0884
fig/66666666.60966.peg.473	CDS	440716	441171	1	+	456	Mobile element protein	-none-	D23_1c0477	Neut_2502
fig/66666666.60966.peg.474	CDS	441829	441158	-1	-	672	Gluconate 2-dehydrogenase (EC 1.1.99.3), membrane-bound, gamma subunit	D-glucamate and ketogluconates metabolism	D23_1c0478	NA
fig/66666666.60966.peg.475	CDS	444093	441973	-3	-	2121	diguanylate cyclase/phosphodiesterase (GGDEF & EAL domains) with PAS/PAC sensor(s)	-none-	D23_1c0479	Neut_0525
fig/66666666.60966.peg.476	CDS	444457	444311	-1	-	147	hypothetical protein	-none-	D23_1c0480	NA
fig/66666666.60966.peg.478	CDS	444629	445810	2	+	1182	NAD(FAD)-utilizing dehydrogenases	-none-	D23_1c0481	Neut_0526
fig/66666666.60966.peg.479	CDS	446569	445952	-1	—	618	Methionine biosynthesis protein MetW	-none-	D23_1c0482	Neut_0527
fig/66666666.60966.peg.480	CDS	447733	446600	-1	-	1134	Homoserine O-acetyltransferase (EC 2.3.1.31)	Methionine Biosynthesis	D23_1c0483	Neut_0528
fig/66666666.60966.peg.481	CDS	449559	447832	-3	-	1728	Phosphoenolpyruvate protein phosphotransferase of PTS system (EC 2.7.3.9)	-none-	D23_1c0484	Neut_0529

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld	C91 Alias
fig 66666666.60966.peg.482	CDS	449825	449556	-2	-	270	Phosphocarrier protein, nitrogen regulation associated	-none-	D23_1c0485	Neut_0530
fig 66666666.60966.peg.483	CDS	450219	449815	-3	-	405	Sugar transport PTS system Ila component	-none-	D23_1c0486	Neut_0531
fig 66666666.60966.peg.484	CDS	450568	451605	1	+	1038	Phosphatidylglycerophosphatase B (EC 3.1.3.27)	Glycerolipid and Glycerophospholipid Metabolism in Bacteria	D23_1c0488	Neut_0532
fig 66666666.60966.peg.485	CDS	451971	451705	-3	-	267	HrgA protein	-none-	D23_1c0489	Neut_2454
fig 66666666.60966.peg.486	CDS	452384	453631	2	+	1248	Mobile element protein	-none-	D23_1c0490	Neut_0357
fig 66666666.60966.peg.487	CDS	455203	454049	-1	-	1155	hypothetical protein	-none-	D23_1c0491	NA
fig 66666666.60966.peg.488	CDS	455538	455371	-3	-	168	hypothetical protein	-none-	D23_1c0492	NA
fig 66666666.60966.peg.489	CDS	455581	456603	1	+	1023	Lipolytic enzyme, G-D-S-L	-none-	D23_1c0493	Neut_0534
fig 66666666.60966.peg.490	CDS	457214	456669	-2	-	546	N-acetylmuramoyl-L-alanine amidase (EC 3.5.1.28) AmpD	Recycling of Peptidoglycan Amino Acids	D23_1c0494	Neut_0535
fig 66666666.60966.peg.491	CDS	457304	457951	2	+	648	Thymidylate kinase (EC 2.7.4.9)	pyrimidine conversions	D23_1c0495	Neut_0536
fig 66666666.60966.peg.492	CDS	461332	458087	-1	-	3246	Type I restriction-modification system, restriction subunit R (EC 3.1.21.3)	Restriction-Modification System; Type I Restriction-Modification	D23_1c0496	Neut_0537
fig 66666666.60966.peg.493	CDS	462292	461345	-1	-	948	Putative DNA-binding protein in cluster with Type I restriction-modification system	Restriction-Modification System	D23_1c0497	NA
fig 66666666.60966.peg.494	CDS	462416	462285	-2	-	132	hypothetical protein	-none-	D23_1c0498	NA
fig 66666666.60966.peg.495	CDS	463405	462413	-1	-	993	hypothetical protein	-none-	D23_1c0499	NA
fig 66666666.60966.peg.496	CDS	464694	463405	-3	-	1290	Type I restriction-modification system, specificity subunit S (EC 3.1.21.3)	Restriction-Modification System; Type I Restriction-Modification	D23_1c0500	Neut_0540
fig 66666666.60966.peg.497	CDS	466246	464684	-1	-	1563	Type I restriction-modification system, DNA-methyltransferase subunit M (EC 2.1.1.72)	Restriction-Modification System; Type I Restriction-Modification	D23_1c0501	Neut_0541
fig 66666666.60966.peg.498	CDS	467880	466453	-3	-	1428	Na ⁺ /H ⁺ antiporter NhaC	-none-	D23_1c0502	Neut_0542
fig 66666666.60966.peg.499	CDS	468057	467896	-3	-	162	hypothetical protein	-none-	D23_1c0503	NA
fig 66666666.60966.peg.500	CDS	468126	469190	3	+	1065	DNA polymerase III delta prime subunit (EC 2.7.7.7)	-none-	D23_1c0504	Neut_0543
fig 66666666.60966.peg.501	CDS	469691	469194	-2	-	498	hypothetical protein	-none-	D23_1c0505	Neut_0544
fig 66666666.60966.peg.502	CDS	469703	469870	2	+	168	hypothetical protein	-none-	D23_1c0506	NA
fig 66666666.60966.peg.503	CDS	470025	471155	3	+	1131	Magnesium and cobalt transport protein CorA	Magnesium transport	D23_1c0508	Neut_0545
fig 66666666.60966.peg.504	CDS	471202	471447	1	+	246	Mobile element protein	-none-	D23_1c0509	Neut_0884
fig 66666666.60966.peg.505	CDS	471504	471617	3	+	114	hypothetical protein	-none-	D23_1c0510	Neut_0547

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.									
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld C91 Alias
fig/66666666.60966.peg.506	CDS	471862	473013	1	+	1152	conserved hypothetical protein	-none-	D23_1c0511 Neut_0548
fig/66666666.60966.peg.507	CDS	473412	473957	3	+	546	Uncharacterized protein conserved in bacteria	-none-	D23_1c0512 Neut_0550
fig/66666666.60966.peg.508	CDS	474111	474269	3	+	159	Mobile element protein	-none-	D23_1c0513 NA
fig/66666666.60966.peg.510	CDS	474450	474653	3	+	204	hypothetical protein	-none-	D23_1c0514 NA
fig/66666666.60966.peg.512	CDS	475553	476005	2	+	453	SSU ribosomal protein S7p (S5e)	<i>Mycobacterium</i> vinience operon involved in protein synthesis (SSU ribosomal proteins)	D23_1c0515 Neut_0554
fig/66666666.60966.peg.513	CDS	476121	478166	3	+	2046	Translation elongation factor G	<i>Mycobacterium</i> vinience operon involved in protein synthesis (SSU ribosomal proteins);	D23_1c0516 Neut_0555
fig/66666666.60966.peg.514	CDS	478196	479386	2	+	1191	Translation elongation factor Tu	elongation factor G family; Translation elongation factors bacterial; Universal GTPases <i>Mycobacterium</i> vinience operon involved in protein synthesis (SSU ribosomal proteins);	D23_1c0517 Neut_0556
fig/66666666.60966.peg.515	CDS	479569	479775	1	+	207	SSU ribosomal protein S10p (S20e)	 Translation elongation factors bacterial; Universal GTPases	D23_1c0518 Neut_0557
fig/66666666.60966.peg.516	CDS	479823	480476	3	+	654	LSU ribosomal protein L3p (L3e)	-none-	D23_1c0519 Neut_0558
fig/66666666.60966.peg.517	CDS	480494	481114	2	+	621	LSU ribosomal protein L4p (L1e)	-none-	D23_1c0520 Neut_0559
fig/66666666.60966.peg.518	CDS	481111	481446	1	+	336	LSU ribosomal protein L23p (L23Ae)	-none-	D23_1c0521 Neut_0560
fig/66666666.60966.peg.519	CDS	481446	482279	3	+	834	LSU ribosomal protein L2p (L8e)	-none-	D23_1c0522 Neut_0561
fig/66666666.60966.peg.520	CDS	482948	483595	2	+	648	SSU ribosomal protein S3p (S3e)	-none-	D23_1c0523 Neut_0564
fig/66666666.60966.peg.521	CDS	483680	484096	2	+	417	LSU ribosomal protein L16p (L10e)	-none-	D23_1c0524 Neut_0565
fig/66666666.60966.peg.524	CDS	485008	485331	1	+	324	LSU ribosomal protein L24p (L26e)	-none-	D23_1c0525 Neut_0569

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.									
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld C91 Alias
fig/66666666.60966.peg.525	CDS	485458	485886	1	+	429	LSU ribosomal protein L5p (L11e)	-none-	D23_1c0526 Neut_0570
fig/66666666.60966.peg.526	CDS	486881	487222	2	+	342	LSU ribosomal protein L6p (L9e)	-none-	D23_1c0527 Neut_0573
fig/66666666.60966.peg.527	CDS	487678	488151	1	+	474	SSU ribosomal protein S5p (S2e)	-none-	D23_1c0528 Neut_0575
fig/66666666.60966.peg.528	CDS	488796	490118	3	+	1323	Preprotein translocase secY subunit (TC 3.A.5.1.1)	-none-	D23_1c0530 Neut_0578
fig/66666666.60966.peg.530	CDS	491337	491963	3	+	627	SSU ribosomal protein S4p (S9e)	-none-	D23_1c0531 Neut_0582
fig/66666666.60966.peg.531	CDS	492065	492997	2	+	933	DNA-directed RNA polymerase alpha subunit (EC 2.7.7.6)	RNA polymerase bacterial	D23_1c0532 Neut_0583
fig/66666666.60966.peg.532	CDS	494213	493998	-2	-	216	Putative oligonucleotide cyclase/dehydratase or lipid transport protein YfgJ	Possible RNA degradation cluster	D23_1c0534 Neut_0586
fig/66666666.60966.peg.533	CDS	494546	494995	2	+	450	tmRNA-binding protein SmpB	Heat shock dnaK gene cluster extended; Translation termination factors bacterial	D23_1c0535 Neut_0587
fig/66666666.60966.peg.534	CDS	495005	496075	2	+	1071	Heme A synthase, cytochrome oxidase biogenesis protein Cox15-CtaA	Biogenesis of cytochrome c oxidases	D23_1c0536 Neut_0588
fig/66666666.60966.peg.535	CDS	496163	496582	2	+	420	Probable transmembrane protein	-none-	D23_1c0537 Neut_0589
fig/66666666.60966.peg.537	CDS	496945	498528	1	+	1584	DNA polymerase III subunits gamma and tau (EC 2.7.7.7)	DNA processing cluster	D23_1c0538 Neut_0590
fig/66666666.60966.peg.538	CDS	498545	498868	2	+	324	FIG000557: hypothetical protein co-occurring with RecR	DNA processing cluster	D23_1c0539 Neut_0591
fig/66666666.60966.peg.539	CDS	498922	499740	1	+	819	Pseudouridine synthase family protein	-none-	D23_1c0540 Neut_0592
fig/66666666.60966.peg.540	CDS	500690	499770	-2	-	921	Protein-N(5)-glutamine methyltransferase PimB, methylates LSU ribosomal protein L3p	-none-	D23_1c0541 Neut_0593
fig/66666666.60966.peg.541	CDS	500738	501241	2	+	504	tRNA-specific adenosine-34 deaminase (EC 3.5.4.—)	tRNA modification Bacteria; tRNA processing	D23_1c0542 Neut_0594
fig/66666666.60966.peg.542	CDS	501238	501717	1	+	480	Conserved domain protein	-none-	D23_1c0543 Neut_0595

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.									
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld C91 Alias
fig/66666666.60966.peg.543	CDS	501779	503389	2	+	1611	NAD-dependent malic enzyme (EC 1.1.1.38)	Pyruvate metabolism I: anaerobic reactions, PEP	D23_1c0544 Neut_0596
fig/66666666.60966.peg.544	CDS	504268	503438	-1	-	831	Phosphoserine phosphatase (EC 3.1.3.3)	Glycine and Serine Utilization; Serine Biosynthesis; Serine Biosynthesis	D23_1c0545 Neut_0597
fig/66666666.60966.peg.545	CDS	505831	504356	-1	-	1476	FIG00858790: hypothetical protein	-none-	D23_1c0546 Neut_0598
fig/66666666.60966.peg.547	CDS	506088	507581	3	+	1494	Cytosol aminopeptidase PepA (EC 3.4.11.1)	Aminopeptidases (EC 3.4.11.--); Dehydrogenase complexes	D23_1c0547 Neut_0599
fig/66666666.60966.peg.548	CDS	507615	508043	3	+	429	DNA polymerase III chi subunit (EC 2.7.7.7)	-none-	D23_1c0548 Neut_0600
fig/66666666.60966.peg.549	CDS	508116	508517	3	+	402	FIG00859089: hypothetical protein	-none-	D23_1c0549 Neut_0601
fig/66666666.60966.peg.550	CDS	508581	511334	3	+	2754	VaiY-HRNA synthetase (EC 6.1.1.9)	tRNA aminoacylation, Val	D23_1c0550 Neut_0602
fig/66666666.60966.peg.551	CDS	511430	512500	2	+	1071	Uroporphyrinogen III decarboxylase (EC 4.1.1.37)	Heme and Siroheme Biosynthesis	D23_1c0551 Neut_0603
fig/66666666.60966.peg.552	CDS	513466	512660	-1	-	807	MaebI	-none-	D23_1c0552 Neut_0604
fig/66666666.60966.peg.553	CDS	514503	513616	-3	-	888	Succinyl-CoA ligase [ADP-forming] alpha chain (EC 6.2.1.5)	TCA Cycle	D23_1c0553 Neut_0605
fig/66666666.60966.peg.554	CDS	515705	514533	-2	-	1173	Succinyl-CoA ligase [ADP-forming] beta chain (EC 6.2.1.5)	TCA Cycle	D23_1c0554 Neut_0606
fig/66666666.60966.peg.555	CDS	516828	515878	-3	-	951	Malyl-CoA lyase (EC 4.1.3.24)	Photorespiration (oxidative C2 cycle)	D23_1c0555 Neut_0607
fig/66666666.60966.peg.557	CDS	518554	517079	-1	-	1476	Glycogen synthase, ADP-glucose transglucosylase (EC 2.4.1.21)	Glycogen metabolism	D23_1c0557 Neut_0608
fig/66666666.60966.peg.558	CDS	520242	518608	-3	-	1635	Glucose-6-phosphate isomerase (EC 5.3.1.9)	Glycolysis and Gluconeogenesis	D23_1c0558 Neut_0609
fig/66666666.60966.peg.559	CDS	521449	520271	-1	-	1179	3-ketoacyl-CoA thiolase (EC 2.3.1.16) @ Acetyl-CoA acetyltransferase (EC 2.3.1.9)	Acetyl-CoA fermentation to Butyrate; Biotin biosynthesis; Biotin synthesis cluster; Butanol Biosynthesis; Butyrate metabolism cluster; Fatty acid metabolism cluster; Isoprenoid	D23_1c0559 Neut_0610

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)				
Function										
Subsystem										
D23 Gbkld										
C91 Alias										
fig/66666666.60966.peg.560	CDS	522790	521459	-1	-	1332	ATP-dependent hsl protease ATP-binding subunit HslU	Biosynthesis; Polyhydroxybutyrate metabolism; Polyhydroxybutyrate metabolism	D23_1c0560	Neut_0611
fig/66666666.60966.peg.561	CDS	523346	522825	-2	-	522	ATP-dependent protease HslV (EC 3.4.25.—)	Proteasome bacterial; Proteolysis in bacteria, ATP-dependent	D23_1c0561	Neut_0612
fig/66666666.60966.peg.563	CDS	523758	523555	-3	-	204	DNA-directed RNA polymerase omega subunit (EC 2.7.7.6)	Proteasome bacterial; Proteolysis in bacteria, ATP-dependent	D23_1c0562	Neut_0613
fig/66666666.60966.peg.564	CDS	524414	523809	-2	-	606	Guanylate kinase (EC 2.7.4.8)	RNA polymerase bacterial	D23_1c0563	Neut_0614
fig/66666666.60966.peg.565	CDS	525166	524684	-1	-	483	Dihydroneopterin triphosphate pyrophosphohydrolase type 2 (nudB)	CBSS-323097.3.peg.2594; Purine conversions	D23_1c0565	Neut_0615
fig/66666666.60966.peg.566	CDS	526961	525180	-2	-	1782	Aspartyl-tRNA synthetase (EC 6.1.1.12) @ Aspartyl-tRNA(Asn) synthetase (EC 6.1.1.23)	tRNA aminoacylation, Asp and Asn; tRNA aminoacylation, Asp and Asn	D23_1c0566	Neut_0616
fig/66666666.60966.peg.567	CDS	527054	527206	2	+	153	hypothetical protein	-none-	D23_1c0567	NA
fig/66666666.60966.peg.568	CDS	528907	527456	-1	-	1452	Mannose-1-phosphate guanylyltransferase (GDP) (EC 2.7.7.22)/ Mannose-6-phosphate isomerase (EC 5.3.1.8)	Mannose Metabolism; Mannose Metabolism	D23_1c0569	Neut_0618
fig/66666666.60966.peg.569	CDS	530083	528971	-1	-	1113	UDP-N-acetylglucosamine 2-epimerase (EC 5.1.3.14)	CMP-N-acetylneuraminate Biosynthesis; Sialic Acid Metabolism	D23_1c0570	Neut_0619
fig/66666666.60966.peg.570	CDS	530171	530284	2	+	114	hypothetical protein	-none-	D23_1c0571	NA
fig/66666666.60966.peg.571	CDS	530407	530535	1	+	129	hypothetical protein	-none-	D23_1c0572	NA
fig/66666666.60966.peg.572	CDS	530637	531041	3	+	405	Truncated hemoglobins	-none-	D23_1c0573	Neut_0620
fig/66666666.60966.peg.573	CDS	531034	532257	1	+	1224	NarS protein involved in response to NO	Denitrification; Nitrosative stress; Oxidative stress	D23_1c0574	Neut_0621
fig/66666666.60966.peg.574	CDS	532298	532738	2	+	441	putative membrane protein	-none-	D23_1c0575	Neut_0622
fig/66666666.60966.peg.575	CDS	532841	533326	2	+	486	FIG001943; hypothetical protein YajQ	Broadly distributed proteins not in subsystems	D23_1c0576	Neut_0623
fig/66666666.60966.peg.576	CDS	534972	533485	-3	-	1488	FIG00859034; hypothetical protein	-none-	D23_1c0577	Neut_0624

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.									
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	C91 Alias
fig/66666666.60966.peg.577	CDS	535028	535240	2	+	213	hypothetical protein	-none-	D23_1c0578 NA
fig/66666666.60966.peg.578	CDS	536092	535289	-1	-	804	FIG00858513:	-none-	D23_1c0579 Neut_0626
fig/66666666.60966.peg.579	CDS	537487	536616	-2	-	882	hypothetical protein	-none-	D23_1c0581 NA
fig/66666666.60966.peg.580	CDS	538547	537726	-2	-	822	hypothetical protein	-none-	D23_1c0582 NA
fig/66666666.60966.peg.581	CDS	539856	538789	-3	-	1068	Conserved domain protein	-none-	D23_1c0583 NA
fig/66666666.60966.peg.582	CDS	540712	539849	-1	-	864	Conserved domain protein	-none-	D23_1c0584 NA
fig/66666666.60966.peg.583	CDS	541704	540841	-3	-	864	possible long-chain N-acyl amino acid synthase	-none-	D23_1c0585 Neut_0638
fig/66666666.60966.peg.584	CDS	541934	541812	-2	-	123	hypothetical protein	-none-	D23_1c0586 NA
fig/66666666.60966.peg.586	CDS	542270	542467	2	+	198	conserved hypothetical protein	-none-	D23_1c0587 Neut_0639
fig/66666666.60966.peg.587	CDS	542451	542618	3	+	168	hypothetical protein	-none-	D23_1c0588 Neut_0640
fig/66666666.60966.peg.588	CDS	542602	542724	1	+	123	hypothetical protein	-none-	D23_1c0589 Neut_0641
fig/66666666.60966.peg.590	CDS	543111	544673	3	+	1563	Putative inner membrane protein	-none-	D23_1c0590 Neut_0642
fig/66666666.60966.peg.591	CDS	544721	544834	2	+	114	hypothetical protein	-none-	D23_1c0591 NA
fig/66666666.60966.peg.592	CDS	545193	546098	3	+	906	Membrane-bound lytic murein transglycosylase D precursor (EC 3.2.1.—)	CBSS-228410.1.peg.134; CBSS-342610.3.peg.1536; Murein Hydrolases DNA Repair Base	D23_1c0592 Neut_0643
fig/66666666.60966.peg.593	CDS	546933	546274	-3	-	660	Endonuclease III (EC 4.2.99.18)	Excision	D23_1c0593 Neut_0644
fig/66666666.60966.peg.594	CDS	547586	546930	-2	-	657	Electron transport complex protein RnIB	-none-	D23_1c0594 Neut_0645
fig/66666666.60966.peg.595	CDS	548604	547576	-3	-	1029	Dihydroorotate dehydrogenase (EC 1.3.3.1)	De Novo Pyrimidine Synthesis	D23_1c0595 Neut_0646
fig/66666666.60966.peg.596	CDS	549246	548605	-3	-	642	Arginine-tRNA-protein transferase (EC 2.3.2.8)	Protein degradation	D23_1c0596 Neut_0647
fig/66666666.60966.peg.597	CDS	550041	549343	-3	-	699	Leucyl/phenylalanyl-tRNA-protein transferase (EC 2.3.2.6)	Protein degradation	D23_1c0597 Neut_0648
fig/66666666.60966.peg.598	CDS	550783	550328	-1	-	456	Mobile element protein	-none-	D23_1c0598 Neut_2502
fig/66666666.60966.peg.599	CDS	551132	550746	-2	-	387	Mobile element protein	-none-	D23_1c0599 Neut_0884
fig/66666666.60966.peg.600	CDS	551404	551517	1	+	114	hypothetical protein	-none-	D23_1c0600 NA
fig/66666666.60966.peg.601	CDS	551625	552500	3	+	876	FIG00859053: hypothetical protein	-none-	D23_1c0601 Neut_0650
fig/66666666.60966.peg.602	CDS	554066	552684	-2	-	1383	UDP-N-acetylmutamate-L-alanyl-gamma-D-glutamyl-meso-diaminopimelate ligase (EC 6.3.2.—)	Peptidoglycan biosynthesis--gio; Recycling of Peptidoglycan Amino Acids	D23_1c0602 Neut_0651

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.									
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld
fig 66666666.60966.peg.603	CDS	554191	555363	1	+	1173	NADH dehydrogenase (EC 1.6.99.3)	Respiratory dehydrogenases 1; Rboflavin synthesis cluster	D23_1c0603
fig 66666666.60966.peg.604	CDS	556325	555387	-2	-	939	Mutator mutT protein (7,8-dihydro-8-oxoguanine-triphosphatase) (EC 3.6.1.-)/Thiamin-phosphate pyrophosphorylase-like protein	Nudix proteins (nucleoside triphosphate hydrolases); Nudix proteins (nucleoside triphosphate hydrolases)	D23_1c0604
fig 66666666.60966.peg.605	CDS	557210	556338	-2	-	873	putative ATP/GTP-binding protein	-none-	D23_1c0605
fig 66666666.60966.peg.606	CDS	558441	557212	-3	-	1230	Glutamate N-acetyltransferase (EC 2.3.1.35)/N-acetylglutamate synthase (EC 2.3.1.1)	Arginine Biosynthesis-gio; Arginine Biosynthesis-gio; Arginine Biosynthesis extended; Arginine Biosynthesis extended	D23_1c0607
fig 66666666.60966.peg.607	CDS	558602	559015	2	+	414	FIG136845: Rhodanese-related sulfotransferase	Glutaredoxin 3 containing cluster	D23_1c0608
fig 66666666.60966.peg.608	CDS	559045	559302	1	+	258	Glutaredoxin 3 (Grx3)	Glutaredoxin 3 containing cluster; Glutaredoxins; Glutathione: Redox cycle	D23_1c0609
fig 66666666.60966.peg.609	CDS	559377	559862	3	+	486	Protein export cytoplasm chaperone protein (SecB, maintains protein to be exported in unfolded state)	Glutaredoxin 3 containing cluster	D23_1c0610
fig 66666666.60966.peg.610	CDS	559866	560345	3	+	480	FIG00859406: hypothetical protein	-none-	D23_1c0611
fig 66666666.60966.peg.611	CDS	560342	561331	2	+	990	Glycerol-3-phosphate dehydrogenase [NAD(P)+] (EC 1.1.1.94)	Glutaredoxin 3 containing cluster; Glycerol and Glycerol-3-phosphate Uptake and Utilization; Glycerolipid and Glycerophospholipid Metabolism in Bacteria	D23_1c0612
fig 66666666.60966.peg.612	CDS	561519	561791	3	+	273	DNA-binding protein HU-beta	DNA structural proteins, bacterial	D23_1c0613

Neut_0652

Neut_0653

Neut_0654

Neut_0655

Neut_0656

Neut_0657

Neut_0658

Neut_0659

Neut_0660

Neut_0661

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.									
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	C91 Alias
fig 66666666.60966.peg.614	CDS	562230	564023	3	+	1794	Peptidyl-prolyl cis-trans isomerase PpiD (EC 5.2.1.8)	Peptidyl-prolyl cis-trans isomerase	D23_1c0616 Neut_0662
fig 66666666.60966.peg.615	CDS	564844	564050	-1	-	795	Enoyl-[acyl-carrier-protein] reductase [NADH] (EC 1.3.1.9)	Fatty Acid Biosynthesis FASII	D23_1c0617 Neut_0663
fig 66666666.60966.peg.616	CDS	565168	564959	-1	-	210	hypothetical protein	-none-	D23_1c0618 NA
fig 66666666.60966.peg.617	CDS	565170	567587	3	+	2418	Transcription accessory protein (S1 RNA-binding domain)	CBSS-243765.1, peg.198; Transcription factors bacterial	D23_1c0619 Neut_0664
fig 66666666.60966.peg.618	CDS	567598	567975	1	+	378	hypothetical protein	-none-	D23_1c0620 Neut_0682
fig 66666666.60966.peg.619	CDS	567977	568255	2	+	279	hypothetical protein	-none-	D23_1c0621 Neut_0682
fig 66666666.60966.peg.621	CDS	568500	568871	3	+	372	hypothetical protein	-none-	D23_1c0622 Neut_0683
fig 66666666.60966.peg.622	CDS	569150	568989	-2	-	162	hypothetical protein	-none-	D23_1c0623 Neut_0684
fig 66666666.60966.peg.623	CDS	570485	569238	-2	-	1248	Mobile element protein	-none-	D23_1c0624 Neut_0357
fig 66666666.60966.peg.624	CDS	571236	570577	-3	-	660	N-hydroxyarylamine O-acetyltransferase (EC 2.3.1.118)	-none-	D23_1c0626 Neut_0684
fig 66666666.60966.peg.625	CDS	572209	571295	-1	-	915	Permease of the drug/metabolite transporter (DMT) superfamily	Quenosine-Archaecoline Biosynthesis	D23_1c0627 Neut_0685
fig 66666666.60966.peg.626	CDS	573622	572339	-1	-	1284	TRAP dicarboxylate transporter, DetM subunit, unknown substrate 6	TRAP Transporter unknown substrate 6	D23_1c0628 Neut_0686
fig 66666666.60966.peg.627	CDS	574259	573705	-2	-	555	TRAP dicarboxylate transporter, DetQ subunit, unknown substrate 6	TRAP Transporter unknown substrate 6	D23_1c0629 Neut_0687
fig 66666666.60966.peg.628	CDS	575698	574334	-1	-	1365	TldE protein, part of TldE/TldD proteolytic complex	Putative TldE-TldD proteolytic complex	D23_1c0630 Neut_0688
fig 66666666.60966.peg.629	CDS	575950	576474	1	+	525	FIG138315; Putative alpha helix protein	Putative TldE-TldD proteolytic complex	D23_1c0632 Neut_0689
fig 66666666.60966.peg.630	CDS	576475	576609	1	+	135	hypothetical protein	proteolytic complex	D23_1c0633 NA
fig 66666666.60966.peg.631	CDS	576740	577006	2	+	267	FIG00859002; hypothetical protein	-none-	D23_1c0635 Neut_0690
fig 66666666.60966.peg.632	CDS	577046	577804	2	+	759	Exodeoxyribonuclease III (EC 3.1.11.2)	DNA repair, bacterial	D23_1c0636 Neut_0691
fig 66666666.60966.peg.633	CDS	577801	579195	1	+	1395	AmpG permease	Recycling of Peptidoglycan Amino Acids	D23_1c0637 Neut_0692
fig 66666666.60966.peg.636	CDS	580217	579867	-2	-	351	Cytochrome O ubiquinol oxidase subunit IV (EC 1.10.3.—)	Terminal cytochrome O ubiquinol oxidase; Terminal cytochrome oxidases	D23_1c0638 Neut_0694

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld	C91 Alias
fig/66666666.60966.peg.637	CDS	580885	580214	-1	-	672	Cytochrome O ubiquinol oxidase subunit III (EC 1.10.3.—)	Terminal cytochrome O ubiquinol oxidase; Terminal cytochrome oxidases	D23_1c0639	Neut_0695
fig/66666666.60966.peg.638	CDS	582993	580882	-3	-	2112	Cytochrome O ubiquinol oxidase subunit 1 (EC 1.10.3.—)	Terminal cytochrome O ubiquinol oxidase; Terminal cytochrome oxidases	D23_1c0640	Neut_0696
fig/66666666.60966.peg.639	CDS	583998	582997	-3	-	1002	Cytochrome O ubiquinol oxidase subunit II (EC 1.10.3.—)	Terminal cytochrome O ubiquinol oxidase; Terminal cytochrome oxidases	D23_1c0641	Neut_0697
fig/66666666.60966.peg.640	CDS	585484	584237	-1	-	1248	Mobile element protein	cytochrome oxidases	D23_1c0642	Neut_0357
fig/66666666.60966.peg.641	CDS	585502	585633	1	+	132	patatin family protein	-none-	D23_1c0643	Neut_1317
fig/66666666.60966.peg.642	CDS	585643	586530	1	+	888	UTP--glucose-1-phosphate uridylyltransferase (EC 2.7.7.9)	-none-	D23_1c0644	Neut_0698
fig/66666666.60966.peg.643	CDS	586705	587817	1	+	1113	FIG00859666; hypothetical protein	-none-	D23_1c0645	Neut_0699
fig/66666666.60966.peg.644	CDS	587837	589201	2	+	1365	Succinate-semialdehyde dehydrogenase [NAD] (EC 1.2.1.24); Succinate-semialdehyde dehydrogenase [NADP+] (EC 1.2.1.16)	-none-	D23_1c0646	Neut_0700
fig/66666666.60966.peg.645	CDS	589224	590942	3	+	1719	InterPro IPR001440 COGs COG0457	-none-	D23_1c0647	Neut_0701
fig/66666666.60966.peg.646	CDS	592871	591033	-2	-	1839	Long-chain-fatty-acid--CoA ligase (EC 6.2.1.3)	Biotin biosynthesis; Biotin synthesis cluster; Fatty acid metabolism cluster	D23_1c0648	Neut_0702
fig/66666666.60966.peg.647	CDS	595204	592868	-1	-	2337	Butyryl-CoA dehydrogenase (EC 1.3.99.2)	-none-	D23_1c0649	Neut_0703
fig/66666666.60966.peg.648	CDS	595489	595223	-1	-	267	hypothetical protein	-none-	D23_1c0650	Neut_0704
fig/66666666.60966.peg.649	CDS	596053	595673	-1	-	381	Putative membrane protein	-none-	D23_1c0651	Neut_0705
fig/66666666.60966.peg.650	CDS	596962	596099	-1	-	864	Pirin	-none-	D23_1c0652	Neut_0706
fig/66666666.60966.peg.651	CDS	597098	598030	2	+	933	Transcriptional regulator; LysR family	-none-	D23_1c0653	Neut_0707
fig/66666666.60966.peg.652	CDS	598202	599338	2	+	1137	hypothetical protein	-none-	D23_1c0655	Neut_0708
fig/66666666.60966.peg.653	CDS	599418	599657	3	+	240	Flavodoxin reductases (ferredoxin-NADPH reductases) family 1	Anaerobic respiratory reductases	D23_1c0656	Neut_0709
fig/66666666.60966.peg.654	CDS	599689	600114	1	+	426	Flavodoxin reductases (ferredoxin-NADPH reductases) family 1	Anaerobic respiratory reductases	D23_1c0657	Neut_0709

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld	C91 Alias
fig/66666666.60966.peg.655 fig/66666666.60966.peg.656	CDS	601243	600251	-1	-	993	Multicopper oxidase	Copper homeostasis -none-	D23_1c0658	Neut_0710
	CDS	602188	601454	-1	-	735	FIG00859807:		D23_1c0659	Neut_0711
fig/66666666.60966.peg.657	CDS	602534	602346	-2	-	189	hypothetical protein	-none-	D23_1c0660	Neut_0712
fig/66666666.60966.peg.658	CDS	602837	602700	-2	-	138	Putative NAD(P)-dependent oxidoreductase EC-YbbO	-none-	D23_1c0661	Neut_0990
fig/66666666.60966.peg.659	CDS	603148	602882	-1	-	267	ABC-type multidrug transport system, permease component	CBSS-196164.1.peg.1690	D23_1c0662	Neut_0713
fig/66666666.60966.peg.660	CDS	603362	605239	2	+	1878	1,4-alpha-glucan branching enzyme (EC 2.4.1.18)	-none-	D23_1c0663	Neut_0714
fig/66666666.60966.peg.661	CDS	605497	605619	1	+	123	Glutathione peroxidase (EC 1.11.1.9)	Glutathione: Redox cycle	D23_1c0664	Neut_0715
fig/66666666.60966.peg.662	CDS	607768	605882	-1	-	1887	TonB-dependent hemin, ferrichrome receptor	Ton and Tol transport systems	D23_1c0665	Neut_0716
fig/66666666.60966.peg.663	CDS	609583	607862	-1	-	1722	hypothetical protein	-none-	D23_1c0666	Neut_0717
fig/66666666.60966.peg.664	CDS	610957	609656	-1	-	1302	Membrane protein involved in colicin uptake	-none-	D23_1c0667	Neut_0718
fig/66666666.60966.peg.665	CDS	611184	612413	3	+	1230	FIG00858430: hypothetical protein	-none-	D23_1c0668	Neut_0719
fig/66666666.60966.peg.666	CDS	612533	615022	2	+	2490	TonB-dependent receptor	Ton and Tol transport systems	D23_1c0669	Neut_0720
fig/66666666.60966.peg.667	CDS	615143	615012	-2	-	132	hypothetical protein	-none-	D23_1c0670	NA
fig/66666666.60966.peg.668	CDS	615337	617457	1	+	2121	TonB-dependent receptor	Ton and Tol transport systems	D23_1c0672	Neut_0721
fig/66666666.60966.peg.669	CDS	617524	618762	1	+	1239	putative signal peptide protein	-none-	D23_1c0673	Neut_0722
fig/66666666.60966.peg.670	CDS	618762	619262	3	+	501	InterPro IPR000063 COGs COG0526	-none-	D23_1c0674	Neut_0723
fig/66666666.60966.peg.671	CDS	619459	621978	1	+	2520	Enoyl-CoA hydratase (EC 4.2.1.17)3,2-trans-enoyl-CoA isomerase (EC 5.3.3.8)/3-hydroxyacyl-CoA dehydrogenase (EC 1.1.1.35)	Acetyl-CoA fermentation to Butyrate; Acetyl-CoA fermentation to Butyrate; Butanol Biosynthesis; Butyrate metabolism cluster; Butyrate metabolism cluster; Fatty acid metabolism cluster; Fatty acid metabolism cluster; Polyhydroxybutyrate metabolism;	D23_1c0675	Neut_0724

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.									
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld C91 Alias
fig/66666666.60966.peg.672	CDS	622043	623245	2	+	1203	3-ketoacyl-CoA thiolase (EC 2.3.1.16) @ Acetyl-CoA acetyltransferase (EC 2.3.1.9)	 Polyhydroxybutyrate metabolism Acetyl-CoA fermentation to Butyrate; Biotin biosynthesis; Biotin synthesis cluster; Butanol Biosynthesis; Butyrate metabolism cluster; Fatty acid metabolism cluster; Isoprenoid Biosynthesis; Polyhydroxybutyrate metabolism; Polyhydroxybutyrate metabolism -none-	D23_1c0676 Neut_0725
fig/66666666.60966.peg.673	CDS	623639	623301	-2	-	339	FIG00859796; hypothetical protein		D23_1c0677 Neut_0726
fig/66666666.60966.peg.674	CDS	623711	623827	2	+	117	hypothetical protein		D23_1c0678 NA
fig/66666666.60966.peg.675	CDS	624186	624479	3	+	294	Mobile element protein	-none-	D23_1c0680 Neut_1719
fig/66666666.60966.peg.676	CDS	624578	625456	2	+	879	Mobile element protein	-none-	D23_1c0681 Neut_1720
fig/66666666.60966.peg.677	CDS	626009	625605	-2	-	405	hypothetical protein	-none-	D23_1c0682 NA
fig/66666666.60966.peg.678	CDS	626270	628669	2	+	2400	Predicted hydrolase of the metallo-beta-lactamase superfamily, clustered with KDO2-Lipid A biosynthesis genes	-none-	D23_1c0684 Neut_0728
fig/66666666.60966.peg.679	CDS	628866	629183	3	+	318	Flagellar transcriptional activator FlhD	Flagellum	D23_1c0685 Neut_0729
fig/66666666.60966.peg.680	CDS	629210	629785	2	+	576	Flagellar transcriptional activator FlhC	Flagellum	D23_1c0686 Neut_0730
fig/66666666.60966.peg.681	CDS	630020	631549	2	+	1530	Proposed peptidoglycan lipid II flippase MurJ	Peptidoglycan lipid II flippase	D23_1c0687 Neut_0731
fig/66666666.60966.peg.683	CDS	633437	632277	-2	-	1161	Outer membrane protein NlpB, lipoprotein component of the protein assembly complex (forms a complex with YaeI, YfiO, and YfiL); Lipoprotein-34 precursor	Lipopolysaccharide assembly	D23_1c0688 Neut_0732
fig/66666666.60966.peg.684	CDS	634289	633447	-2	-	843	Dihydrodipicolinate synthase (EC 4.2.1.52)	-none-	D23_1c0689 Neut_0733

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld	C91 Alias
fig 66666666.60966.peg.685	CDS	637007	634416	-2	-	2592	CipB protein Proteolysis in bacteria, ATP-dependent	Protein chaperones;	D23_1c0690	Neut_0734
fig 66666666.60966.peg.686	CDS	637484	637326	-2	-	159	hypothetical protein	-none-	D23_1c0693	NA
fig 66666666.60966.peg.687	CDS	638994	637501	-3	-	1494	Ferredoxin-dependent glutamate synthase (EC 1.4.7.1)	 Glutamine, Glutamate, Aspartate and Asparagine Biosynthesis	D23_1c0694	Neut_0735
fig 66666666.60966.peg.688	CDS	639034	639930	1	+	897	Quinolinate phosphoribosyltransferase [decarboxylating] (EC 2.4.2.19)	<i>Mycobacterium</i> virulence operon possibly involved in quinolate biosynthesis; NAD and NADP cofactor biosynthesis global	D23_1c0695	NA
fig 66666666.60966.peg.689	CDS	641903	640635	-2	-	1269	Flagellar hook-length control protein FlkK	Flagellum	D23_1c0696	Neut_0740
fig 66666666.60966.peg.690	CDS	642413	641961	-2	-	453	Flagellar protein FljJ	Flagellum	D23_1c0697	Neut_0741
fig 66666666.60966.peg.691	CDS	643836	642430	-3	-	1407	Flagellum-specific ATP synthase FljI	Flagellar motility; Flagellum	D23_1c0698	Neut_0742
fig 66666666.60966.peg.692	CDS	644577	643858	-3	-	720	Flagellar assembly protein FljH	Flagellum	D23_1c0699	Neut_0743
fig 66666666.60966.peg.693	CDS	645764	644769	-2	-	996	Flagellar motor switch protein FljG	Flagellum	D23_1c0700	Neut_0744
fig 66666666.60966.peg.694	CDS	647397	645754	-3	-	1644	Flagellar M-ring protein FljF	Flagellum	D23_1c0701	Neut_0745
fig 66666666.60966.peg.695	CDS	647548	647402	-1	-	147	hypothetical protein	-none-	D23_1c0702	NA
fig 66666666.60966.peg.696	CDS	647628	648872	3	+	1245	Flagellar sensor histidine kinase FlcS	Flagellum	D23_1c0703	Neut_0746
fig 66666666.60966.peg.697	CDS	648876	650189	3	+	1314	InterPro IPR001789; IPR002078; IPR002197; IPR003593	-none-	D23_1c0704	Neut_0747
fig 66666666.60966.peg.698	CDS	650217	650546	3	+	330	COGs COG2204 Flagellar hook-basal body complex protein FlcE	Flagellum; Flagellum in <i>Campylobacter</i>	D23_1c0705	Neut_0748
fig 66666666.60966.peg.699	CDS	650581	651390	1	+	810	FIG00858443: hypothetical protein	-none-	D23_1c0706	Neut_0749
fig 66666666.60966.peg.700	CDS	651752	651411	-2	-	342	Flagellar biosynthesis protein FlhB	Flagellar motility; Flagellum	D23_1c0707	Neut_0750
fig 66666666.60966.peg.701	CDS	652764	651739	-3	-	1026	FIG00726091: hypothetical protein	-none-	D23_1c0708	Neut_0751
fig 66666666.60966.peg.702	CDS	652766	652948	2	+	183	hypothetical protein	-none-	D23_1c0709	NA
fig 66666666.60966.peg.703	CDS	653125	653517	1	+	393	hypothetical protein	-none-	D23_1c0710	Neut_2449
fig 66666666.60966.peg.704	CDS	653664	653810	3	+	147	Mobile element protein	-none-	D23_1c0711	Neut_1756
fig 66666666.60966.peg.705	CDS	654316	653858	-1	-	459	Cytochrome c family protein	-none-	D23_1c0712	Neut_0754

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld	C91 Alias
fig/66666666.60966.peg.706	CDS	654869	654345	-2	-	525	CopG protein	Copper homeostasis	D23_1c0713	Neut_0755
	CDS	655268	656209	2	+	942	tRNA(Cytosine32)-2-thiocytidine synthetase	CBSS-326442.4.peg.1852; tRNA modification	D23_1c0714	Neut_0756
fig/66666666.60966.peg.708	CDS	657180	656236	-3	-	945	Lipoate synthase	Bacteria Lipoic acid metabolism; Lipoic acid synthesis cluster; Possible RNA degradation cluster	D23_1c0715	Neut_0757
fig/66666666.60966.peg.709	CDS	657847	657170	-1	-	678	Octanoate-[acyl-carrier-protein]-protein-N-octanoyltransferase	Lipoic acid metabolism; Lipoic acid synthesis cluster	D23_1c0716	Neut_0758
fig/66666666.60966.peg.710	CDS	658198	657935	-1	-	264	Proposed lipoate regulatory protein YbeD	Lipoic acid metabolism; Lipoic acid synthesis cluster	D23_1c0717	Neut_0759
fig/66666666.60966.peg.711	CDS	659068	658208	-1	-	861	D-alanine aminotransferase (EC 2.6.1.21)	Pyruvate Alanine Serine Interconversions	D23_1c0718	Neut_0760
fig/66666666.60966.peg.712	CDS	660251	659088	-2	-	1164	D-alanyl-D-alanine carboxypeptidase (EC 3.4.16.4)	CBSS-84588.1.peg.1247; Metallocoarboxypeptidases (EC 3.4.17.—); Murein Hydrolases; Peptidoglycan Biosynthesis	D23_1c0719	Neut_0761
fig/66666666.60966.peg.713	CDS	660353	660787	2	+	435	LSU ribosomal protein L13p (L13Ae)	-none-	D23_1c0720	Neut_0762
fig/66666666.60966.peg.714	CDS	660799	661191	1	+	393	SSU ribosomal protein S9p (S16c)	-none-	D23_1c0721	Neut_0763
fig/66666666.60966.peg.715	CDS	661324	662352	1	+	1029	N-acetyl-gamma-glutamyl-phosphate reductase (EC 1.2.1.38)	Arginine Biosynthesis--glo; Arginine Biosynthesis extended	D23_1c0722	Neut_0764
fig/66666666.60966.peg.716	CDS	662454	663221	3	+	768	putative integral membrane protein	-none-	D23_1c0723	Neut_0765
fig/66666666.60966.peg.717	CDS	663202	663627	1	+	426	Integral membrane protein CcmA involved in cell shape determination	-none-	D23_1c0724	Neut_0766
fig/66666666.60966.peg.718	CDS	665316	663655	-3	-	1662	DNA repair protein RecN	DNA repair, bacterial	D23_1c0725	Neut_0767
fig/66666666.60966.peg.719	CDS	666357	665326	-3	-	1032	NAD kinase (EC 2.7.1.23)	NAD and NADP cofactor biosynthesis global	D23_1c0726	Neut_0768
fig/66666666.60966.peg.720	CDS	666433	667449	1	+	1017	Heat-inducible transcription repressor HrcA	GroEL GroES; Heat shock dnaK gene cluster extended	D23_1c0727	Neut_0769
fig/66666666.60966.peg.721	CDS	667469	668566	2	+	1098	Ferrochelatase, protoheme ferro-lyase (EC 4.99.1.1)	Heme and Siroheme Biosynthesis	D23_1c0728	Neut_0770

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld	C91 Alias
fig/66666666.60966.peg.722	CDS	668678	669469	2	+	792	Zn-dependent protease with chaperone function PA4632	-none-	D23_1c0729	Neut_0771
fig/66666666.60966.peg.723	CDS	669589	670461	1	+	873	Phosphoribulokinase (EC 2.7.1.19)	Calvin-Benson cycle	D23_1c0730	Neut_0772
fig/66666666.60966.peg.724	CDS	670525	672771	1	+	2247	ATP-dependent DNA helicase UvrD/PcrA	DNA repair, bacterial UvrD and related helicases	D23_1c0731	Neut_0773
fig/66666666.60966.peg.725	CDS	673569	672778	-3	-	792	Possible transmembrane protein	-none-	D23_1c0732	Neut_0774
fig/66666666.60966.peg.726	CDS	674610	673660	-3	-	951	Homoserine kinase (EC 2.7.1.39)	CBSS-269482.1.peg.1294; Methionine Biosynthesis; Threonine and Homoserine Biosynthesis	D23_1c0733	Neut_0775
fig/66666666.60966.peg.727	CDS	674585	674716	2	+	132	hypothetical protein	-none-	D23_1c0734	NA
fig/66666666.60966.peg.728	CDS	675030	674704	-3	-	327	FIG00858562; hypothetical protein		D23_1c0735	Neut_0776
fig/66666666.60966.peg.729	CDS	675190	674996	-1	-	195	hypothetical protein	-none-	D23_1c0736	NA
fig/66666666.60966.peg.730	CDS	675866	675147	-2	-	720	FIG00859241; hypothetical protein	-none-	D23_1c0737	Neut_0777
fig/66666666.60966.peg.731	CDS	675882	678602	3	+	2721	DNA polymerase I (EC 2.7.7.7)	DNA Repair Base Excision	D23_1c0738	Neut_0778
fig/66666666.60966.peg.732	CDS	678725	679912	2	+	1188	Fatty acid desaturase (EC 1.14.19.1); Delta-9 fatty acid desaturase (EC 1.14.19.1)	-none-	D23_1c0739	Neut_0779
fig/66666666.60966.peg.733	CDS	680149	679994	-1	-	156	LSU ribosomal protein L33p @ LSU ribosomal protein L33p, zinc-independent	-none-	D23_1c0740	Neut_0780
fig/66666666.60966.peg.734	CDS	680390	680190	-2	-	201	LSU ribosomal protein L28p	-none-	D23_1c0741	Neut_0781
fig/66666666.60966.peg.735	CDS	681169	680495	-1	-	675	DNA repair protein RadC	Bacterial cell division cluster; DNA repair, bacterial	D23_1c0742	Neut_0782
fig/66666666.60966.peg.736	CDS	681339	682508	3	+	1170	Phosphopantothenoylcysteine decarboxylase (EC 4.1.1.36)/ Phosphopantothenoylcysteine synthetase (EC 6.3.2.5)	Coenzyme A Biosynthesis; Coenzyme A Biosynthesis	D23_1c0743	Neut_0783
fig/66666666.60966.peg.737	CDS	682518	682967	3	+	450	Deoxyuridine 5'-triphosphate nucleotidohydrolase (EC 3.6.1.23)	Housecleaning nucleoside triphosphate pyrophosphatases; Nudix proteins	D23_1c0744	Neut_0784

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 GbkId	C91 Alias
fig/66666666.60966.peg.738	CDS	682961	683386	2	+	426	exported protein	(nucleoside triphosphate hydrolases)	D23_1c0745	Neut_0785
fig/66666666.60966.peg.739	CDS	685816	683759	-1	-	2058	Pyrophosphate-energized proton pump (EC 3.6.1.1)	-none- Phosphate metabolism	D23_1c0746	Neut_0786
fig/66666666.60966.peg.740	CDS	687229	685970	-1	-	1260	6-phosphofructokinase (EC 2.7.1.11)	Glycolysis and Gluconeogenesis	D23_1c0747	Neut_0787
fig/66666666.60966.peg.741	CDS	687933	687400	-3	-	534	Adenylate kinase (EC 2.7.4.3)	Purine conversions	D23_1c0748	NA
fig/66666666.60966.peg.742	CDS	688328	689359	2	+	1032	RecA protein	DNA repair, bacterial; DNA repair system including RecA, MutS and a hypothetical protein; RecA and RecX	D23_1c0749	Neut_0789
fig/66666666.60966.peg.743	CDS	689362	689802	1	+	441	Regulatory protein RecX	DNA repair system including RecA, MutS and a hypothetical protein; RecA and RecX	D23_1c0750	Neut_0790
fig/66666666.60966.peg.744	CDS	689820	692411	3	+	2592	Alanyl-tRNA synthetase (EC 6.1.1.7)	Cluster containing Alanyl-tRNA synthetase; tRNA	D23_1c0751	Neut_0791
fig/66666666.60966.peg.745	CDS	692450	693403	2	+	954	Thioredoxin reductase (EC 1.8.1.9)	aminoacylation, Ala Thioredoxin-disulfide reductase; pyrimidine conversions	D23_1c0752	Neut_0792
fig/66666666.60966.peg.746	CDS	693409	693991	1	+	591	Smr domain	-none-	D23_1c0753	Neut_0793
fig/66666666.60966.peg.747	CDS	694203	694349	3	+	147	Carbonic anhydrase (EC 4.2.1.1)	Zinc regulated enzymes	D23_1c0754	Neut_0794
fig/66666666.60966.peg.750	CDS	695056	695208	1	+	153	hypothetical protein	-none-	D23_1c0755	Neut_1255
fig/66666666.60966.peg.751	CDS	695216	695383	2	+	168	hypothetical protein	-none-	D23_1c0756	Neut_2449
fig/66666666.60966.peg.752	CDS	695522	695986	2	+	465	Mobile element protein	-none-	D23_1c0758	Neut_1256
fig/66666666.60966.peg.753	CDS	696281	696072	-2	-	210	Chemotaxis regulator-transmits chemoreceptor signals to flagellar motor components CheY	Flagellar motility	D23_1c0759	Neut_0796
fig/66666666.60966.peg.754	CDS	696491	696619	2	+	129	hypothetical protein	-none-	D23_1c0760	Neut_0797
fig/66666666.60966.peg.755	CDS	696639	696812	3	+	174	Mobile element protein	-none-	D23_1c0761	Neut_0797
fig/66666666.60966.peg.756	CDS	696806	696934	2	+	129	Mobile element protein	-none-	D23_1c0761	NA
fig/66666666.60966.peg.757	CDS	697179	698426	3	+	1248	Mobile element protein	-none-	D23_1c0762	Neut_0357
fig/66666666.60966.peg.758	CDS	698760	700454	3	+	1695	NADH dehydrogenase, subunit 5	Respiratory Complex I	D23_1c0763	Neut_0799
fig/66666666.60966.peg.759	CDS	700473	701258	3	+	786	hypothetical protein	-none-	D23_1c0765	Neut_0800
fig/66666666.60966.peg.760	CDS	701243	704371	2	+	3129	Hypothetical	CO2 uptake,	D23_1c0766	Neut_0801

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 GbkId	C91 Alias
fig/66666666.60966.peg.761	CDS	704368	704694	1	+	327	transmembrane protein coupled to NADH-ubiquinone oxidoreductase chain 5 homolog	carboxysome; Respiratory Complex I	D23_1c0767	Neut_0802
fig/66666666.60966.peg.762	CDS	704850	705059	3	+	210	Nitrogen regulatory protein P-II	-none-	D23_1c0768	Neut_0800
fig/66666666.60966.peg.763	CDS	705044	708184	2	+	3141	Hypothetical transmembrane protein coupled to NADH-ubiquinone oxidoreductase chain 5 homolog	carboxysome; Respiratory Complex I	D23_1c0769	Neut_0801
fig/66666666.60966.peg.764	CDS	708181	708507	1	+	327	Nitrogen regulatory protein P-II	Ammonia assimilation	D23_1c0770	Neut_0802
fig/66666666.60966.peg.765	CDS	709429	708509	-1	-	921	RuBisCO operon transcriptional regulator CbbR	CO2 uptake, carboxysome	D23_1c0771	Neut_0803
fig/66666666.60966.peg.766	CDS	709628	711049	2	+	1422	Ribulose biphosphate carboxylase large chain (EC 4.1.1.39)	CO2 uptake, carboxysome; Calvin-Benson cycle; Photorespiration (oxidative C2 cycle)	D23_1c0772	Neut_0804
fig/66666666.60966.peg.767	CDS	711134	711460	2	+	327	Ribulose biphosphate carboxylase small chain (EC 4.1.1.39)	CO2 uptake, carboxysome; Calvin-Benson cycle; Photorespiration (oxidative C2 cycle)	D23_1c0773	Neut_0805
fig/66666666.60966.peg.768	CDS	711860	711645	-2	-	216	hypothetical protein	-none-	D23_1c0774	Neut_0806
fig/66666666.60966.peg.769	CDS	711868	714264	1	+	2397	carboxysome shell protein CsoS2	CO2 uptake, carboxysome	D23_1c0774	Neut_0806
fig/66666666.60966.peg.770	CDS	714275	715804	2	+	1530	carboxysome shell protein CsoS3	CO2 uptake, carboxysome	D23_1c0775	Neut_0807
fig/66666666.60966.peg.771	CDS	715825	716082	1	+	258	putative carboxysome peptide A	carboxysome	D23_1c0776	Neut_0808
fig/66666666.60966.peg.772	CDS	716082	716330	3	+	249	putative carboxysome peptide B	CO2 uptake, carboxysome	D23_1c0777	Neut_0809
fig/66666666.60966.peg.773	CDS	716439	716735	3	+	297	carboxysome shell protein CsoS1	CO2 uptake, carboxysome	D23_1c0778	Neut_0810
fig/66666666.60966.peg.774	CDS	716777	717124	2	+	348	carboxysome shell protein CsoS1	carboxysome	D23_1c0779	Neut_0811
fig/66666666.60966.peg.775	CDS	717145	717567	1	+	423	bacterioferritin possible associated with carboxysome	CO2 uptake, carboxysome; -none-	D23_1c0780	Neut_0812

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld	C91 Alias
fig/66666666.60966.peg.776	CDS	717576	717842	3	+	267	Possible pterin-4 alpha-carbinolamine dehydratase-like protein	CO2 uptake, carboxysome	D23_1c0781	Neut_0813
fig/66666666.60966.peg.777	CDS	717911	718483	2	+	573	Chromosome (plasmid) partitioning protein ParA	Bacterial Cell Division; Bacterial Cytoskeleton; Cell Division Subsystem including YidCD; RNA modification and chromosome partitioning cluster	D23_1c0782	Neut_0814
fig/66666666.60966.peg.778	CDS	718480	718674	1	+	195	hypothetical protein	-none-	D23_1c0783	NA
fig/66666666.60966.peg.779	CDS	718681	719631	1	+	951	Rubisco activation protein CbbQ	CO2 uptake, carboxysome	D23_1c0784	Neut_0815
fig/66666666.60966.peg.780	CDS	719654	722014	2	+	2361	Rubisco activation protein CbbO	CO2 uptake, carboxysome	D23_1c0785	Neut_0816
fig/66666666.60966.peg.781	CDS	722027	722659	2	+	633	FIG00852745; hypothetical protein	-none-	D23_1c0786	Neut_0817
fig/66666666.60966.peg.782	CDS	722704	723528	1	+	825	FIG00853400; hypothetical protein	-none-	D23_1c0787	Neut_0818
fig/66666666.60966.peg.783	CDS	723822	723676	-3	-	147	hypothetical protein	-none-	D23_1c0788	NA
fig/66666666.60966.peg.784	CDS	723900	724055	3	+	156	Mobile element protein	-none-	D23_1c0790	NA
fig/66666666.60966.peg.785	CDS	724125	724304	3	+	180	hypothetical protein	-none-	D23_1c0791	NA
fig/66666666.60966.peg.786	CDS	724415	724621	2	+	207	Rubisco activation protein CbbO	CO2 uptake, carboxysome	D23_1c0792	Neut_0816
fig/66666666.60966.peg.787	CDS	724700	724975	2	+	276	Rubisco activation protein CbbO Nitric oxide reductase activation protein NorD	carboxysome Denitrification; Denitrifying reductase gene clusters	D23_1c0793	Neut_0816
fig/66666666.60966.peg.788	CDS	724985	725104	2	+	120	hypothetical protein	-none-	D23_1c0794	Neut_0816
fig/66666666.60966.peg.789	CDS	725079	725336	3	+	258	Nitric oxide reductase activation protein NorD	Denitrification; Denitrifying reductase gene clusters	D23_1c0794	Neut_0816
fig/66666666.60966.peg.790	CDS	725493	725657	3	+	165	hypothetical protein	-none-	D23_1c0795	Neut_0821
fig/66666666.60966.peg.791	CDS	727182	725782	-3	-	1401	FIG00861154; hypothetical protein	-none-	D23_1c0796	Neut_1550
fig/66666666.60966.peg.792	CDS	727830	727411	-3	-	420	FIG00859219; hypothetical protein	-none-	D23_1c0797	Neut_0823
fig/66666666.60966.peg.793	CDS	729449	728481	-2	-	969	hypothetical protein Octaprenyl diphosphate synthase (EC 2.5.1.90) Dimethylallyltransferase (EC 2.5.1.1)(2E,6E)-farnesyl diphosphate synthase (EC 2.5.1.10) Geranylgeranyl diphosphate synthase (EC 2.5.1.29)	Isoprenoid Biosynthesis; Isoprenoid Biosynthesis; Isoprenoid Biosynthesis; Interconversions; Isoprenoids for Quinones; Isoprenoids for	D23_1c0800	Neut_0825

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.									
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld C91 Alias
fig/66666666, 60966, peg.795	CDS	729633	730886	3	+	1254	Glutamyl-tRNA reductase (EC 1.2.1.70)	Quinones; Isoprenoids for Quinones; Isoprenoids for Quinones; Polyprenyl Diphosphate Biosynthesis; Polyprenyl Diphosphate Biosynthesis; Polyprenyl Diphosphate Biosynthesis; Polyprenyl Diphosphate Biosynthesis A Gammaproteobacteria Cluster Relating to Translation; Heme and Siroheme Biosynthesis A Gammaproteobacteria Cluster Relating to Translation; CBSS-216600.3, peg.802; Translation termination factors bacterial	Neut_0826
fig/66666666, 60966, peg.796	CDS	730883	731962	2	+	1080	Peptide chain release factor 1	A Gammaproteobacteria Cluster Relating to Translation; CBSS-216600.3, peg.802; Translation termination factors bacterial	Neut_0827
fig/66666666, 60966, peg.797	CDS	731959	732840	1	+	882	Protein-N(5)-glutamine methyltransferase PrnC, methylates polypeptide chain release factors RF1 and RF2	A Gammaproteobacteria Cluster Relating to Translation; CBSS-216600.3, peg.802; Translation termination factors bacterial	Neut_0828
fig/66666666, 60966, peg.798	CDS	732995	733303	2	+	309	Glutaredoxin-related protein	Glutaredoxins	Neut_0829
fig/66666666, 60966, peg.799	CDS	733743	733324	-3	-	420	Purative membrane protein	-none-	Neut_0830
fig/66666666, 60966, peg.800	CDS	734018	734161	2	+	144	hypothetical protein	-none-	NA
fig/66666666, 60966, peg.801	CDS	734194	734826	1	+	633	Glutathione S-transferase (EC 2.5.1.18)	Glutathione: Non-redox reactions	Neut_0831
fig/66666666, 60966, peg.802	CDS	735097	735381	1	+	285	conserved hypothetical protein	-none-	Neut_0832
fig/66666666, 60966, peg.803	CDS	735730	736488	1	+	759	hypothetical protein	-none-	Neut_0833
fig/66666666, 60966, peg.804	CDS	737600	736485	-2	-	1116	COGs COG1502	-none-	Neut_0834
fig/66666666, 60966, peg.805	CDS	738943	737585	-1	-	1359	hypothetical protein	-none-	Neut_0835
fig/66666666, 60966, peg.806	CDS	739386	739240	-3	-	147	hypothetical protein	-none-	NA
fig/66666666, 60966, peg.807	CDS	740790	739585	-3	-	1206	DnaI domain protein	-none-	Neut_0836

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.									
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	C91 Alias
fig/66666666.60966.peg.809	CDS	742340	741159	-2	-	1182	Putative aminotransferase	-none-	D23_1c0815 Neut_0837
fig/66666666.60966.peg.810	CDS	744454	742337	-1	-	2118	Conserved domain protein	-none-	D23_1c0816 Neut_0838
fig/66666666.60966.peg.811	CDS	744823	744647	-1	-	177	hypothetical protein	-none-	D23_1c0817 NA
fig/66666666.60966.peg.812	CDS	745953	745228	-3	-	726	hypothetical protein	-none-	D23_1c0818 Neut_0840
fig/66666666.60966.peg.814	CDS	748473	746350	-3	-	2124	InterPro IPR000209 COGs COG1404	-none-	D23_1c0820 Neut_0841
fig/66666666.60966.peg.815	CDS	748862	749245	2	+	384	ApaG protein	EC49-61	D23_1c0822 Neut_0842
fig/66666666.60966.peg.816	CDS	749258	749992	2	+	735	Tetrapyrrole methylase family protein	-none-	D23_1c0823 Neut_0843
fig/66666666.60966.peg.817	CDS	751297	750077	-1	-	1221	Proton/glutamate symport protein @ Sodium/glutamate symport protein	Glutamate and Aspartate uptake in Bacteria	D23_1c0824 Neut_0844
fig/66666666.60966.peg.818	CDS	752156	751611	-2	-	546	Cytochrome c-type biogenesis protein ResA	Biogenesis of c-type cytochromes	D23_1c0825 Neut_0845
fig/66666666.60966.peg.819	CDS	754247	752304	-2	-	1944	Cytochrome c-type biogenesis protein DsbD, protein-disulfide reductase (EC 1.8.1.8)	Biogenesis of c-type cytochromes; Periplasmic disulfide interchange	D23_1c0826 Neut_0846
fig/66666666.60966.peg.820	CDS	754631	754260	-2	-	372	Periplasmic divalent cation tolerance protein CutA	Copper homeostasis: copper tolerance	D23_1c0827 Neut_0847
fig/66666666.60966.peg.821	CDS	754715	754909	2	+	195	FIG00859483: hypothetical protein	-none-	D23_1c0828 Neut_0848
fig/66666666.60966.peg.822	CDS	754952	755695	2	+	744	FIG00859295: hypothetical protein	-none-	D23_1c0829 Neut_0849
fig/66666666.60966.peg.823	CDS	756608	755733	-2	-	876	Staphylococcus nuclease (SNase) domain	-none-	D23_1c0830 Neut_0850
fig/66666666.60966.peg.824	CDS	756650	757549	2	+	900	Methionine ABC transporter ATP-binding protein	Methionine Biosynthesis; Methionine Degradation	D23_1c0831 Neut_0851
fig/66666666.60966.peg.825	CDS	757674	758384	3	+	711	Uncharacterized ABC transporter, permease component YrbE	Lipopolysaccharide assembly	D23_1c0832 Neut_0852
fig/66666666.60966.peg.826	CDS	758396	758863	2	+	468	Uncharacterized ABC transporter, periplasmic component YrbD	Lipopolysaccharide assembly	D23_1c0833 Neut_0853
fig/66666666.60966.peg.827	CDS	758879	759496	2	+	618	Uncharacterized ABC transporter, auxiliary component YrbC	Lipopolysaccharide assembly	D23_1c0834 Neut_0854
fig/66666666.60966.peg.828	CDS	759503	759817	2	+	315	STAS domain	-none-	D23_1c0835 Neut_0855
fig/66666666.60966.peg.829	CDS	759879	760793	3	+	915	ABC-type multidrug transport system, ATPase component	CBSS-196164.1.peg.1690	D23_1c0836 Neut_0856

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.											
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld	C91 Alias	
fig/66666666.60966.peg.830	CDS	760790	761545	2	+	756	ABC-type multidrug transport system, permease component	CBSS-196164.1.peg.1690	D23_1c0837	Neut_0857	
fig/66666666.60966.peg.831	CDS	761590	761844	1	+	255	YrbA protein		Broadly distributed proteins not in subsystems	D23_1c0838	Neut_0858
fig/66666666.60966.peg.832	CDS	763192	761900	-1	-	1293	Dihydroliipoamide succinyltransferase component (E2) of 2-oxoglutarate dehydrogenase complex (EC 2.3.1.61)	Dehydrogenase complexes; TCA Cycle	D23_1c0839	Neut_0859	
fig/66666666.60966.peg.833	CDS	766072	763214	-1	-	2859	2-oxoglutarate dehydrogenase E1 component (EC 1.2.4.2)	Dehydrogenase complexes; TCA Cycle	D23_1c0840	Neut_0860	
fig/66666666.60966.peg.834	CDS	767529	766234	-3	-	1296	Citrate synthase (si) (EC 2.3.3.1)	TCA Cycle	D23_1c0841	Neut_0861	
fig/66666666.60966.peg.835	CDS	767808	767575	-3	-	234	YgY COG2938	-none- 5-FCL-like protein; Succinate dehydrogenase; TCA Cycle	D23_1c0842	Neut_0862	
fig/66666666.60966.peg.836	CDS	768500	767805	-2	-	696	Succinate dehydrogenase iron-sulfur protein (EC 1.3.99.1)		D23_1c0843	Neut_0863	
fig/66666666.60966.peg.837	CDS	770059	768629	-1	-	1431	Threonine synthase (EC 4.2.3.1)	Threonine and Homoserine	D23_1c0844	Neut_0864	
fig/66666666.60966.peg.839	CDS	771500	770184	-2	-	1317	Homoserine dehydrogenase (EC 1.1.1.3)	Biosynthesis Methionine	D23_1c0845	Neut_0865	
fig/66666666.60966.peg.840	CDS	772833	771607	-3	-	1227	Aspartate aminotransferase (EC 2.6.1.1)	Biosynthesis CBSS-216591.1.peg.168; Glutamate, Aspartate and Asparagine	D23_1c0846	Neut_0866	
fig/66666666.60966.peg.841	CDS	773025	773147	3	+	123	hypothetical protein	Biosynthesis; Threonine and Homoserine	D23_1c0847	NA	
fig/66666666.60966.peg.842	CDS	773125	773508	1	+	384	Membrane protein		-none-	D23_1c0848	Neut_0867
fig/66666666.60966.peg.843	CDS	773605	775980	1	+	2376	Phosphoenolpyruvate synthase (EC 2.7.9.2)		A Hypothetical that Clusters with PEP Synthase; Glycolysis and Gluconeogenesis; Pyruvate metabolism I; anaplerotic reactions, PEP	D23_1c0849	Neut_0868

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld	C91 Alias
fig/66666666.60966.peg.844	CDS	775985	776812	2	+	828	FIG137360: hypothetical protein	A Hypothetical that Clusters with PEP Synthase	D23_1c0850	Neut_0869
fig/66666666.60966.peg.845	CDS	777372	776854	-3	-	519	NLP/P60	-none-	D23_1c0851	Neut_0870
fig/66666666.60966.peg.846	CDS	779204	777507	-2	-	1698	Glutaminyl-RNA synthetase (EC 6.1.1.18)	tRNA aminoacylation, Glu and Gln	D23_1c0852	Neut_0871
fig/66666666.60966.peg.847	CDS	779394	779245	-3	-	150	hypothetical protein	-none-	D23_1c0853	NA
fig/66666666.60966.peg.848	CDS	779393	780991	2	+	1599	Lysyl-tRNA synthetase (class II) (EC 6.1.1.6)	tRNA aminoacylation, Lys	D23_1c0854	Neut_0872
fig/66666666.60966.peg.849	CDS	781093	781749	1	+	657	FIG00858849: hypothetical protein	-none-	D23_1c0855	Neut_0873
fig/66666666.60966.peg.850	CDS	781999	782385	1	+	387	hypothetical protein	-none-	D23_1c0856	NA
fig/66666666.60966.peg.851	CDS	782415	782747	3	+	333	Mobile element protein	-none-	D23_1c0857	NA
fig/66666666.60966.peg.852	CDS	782952	783485	3	+	534	hypothetical protein	-none-	D23_1c0858	Neut_0875
fig/66666666.60966.peg.853	CDS	783597	783767	3	+	171	hypothetical protein	-none-	D23_1c0859	NA
fig/66666666.60966.peg.854	CDS	784671	784048	-3	-	624	Trp repressor binding protein	-none-	D23_1c0860	Neut_0876
fig/66666666.60966.peg.855	CDS	785041	784757	-1	-	285	Mobile element protein	-none-	D23_1c0861	NA
fig/66666666.60966.peg.857	CDS	787300	785558	-1	-	1743	Beta-glucosidase (EC 3.2.1.21)	-none-	D23_1c0862	Neut_0879
fig/66666666.60966.peg.858	CDS	788430	787327	-3	-	1104	Mobile element protein	-none-	D23_1c0863	Neut_1278
fig/66666666.60966.peg.859	CDS	789557	789045	-2	-	513	Mobile element protein	-none-	D23_1c0864	Neut_1624
fig/66666666.60966.peg.860	CDS	789894	789589	-3	-	306	Mobile element protein	-none-	D23_1c0865	Neut_1371
fig/66666666.60966.peg.861	CDS	790136	789951	-2	-	186	Mobile element protein	-none-	D23_1c0866	Neut_2500
fig/66666666.60966.peg.863	CDS	792236	790869	-2	-	1368	ATP-dependent RNA helicase RhlE	ATP-dependent RNA helicases, bacterial	D23_1c0869	Neut_0889
fig/66666666.60966.peg.865	CDS	794169	792502	-3	-	1668	ATPase components of ABC transporters with duplicated ATPase domains	-none-	D23_1c0871	Neut_0890
fig/66666666.60966.peg.867	CDS	794389	794568	1	+	180	hypothetical protein	-none-	D23_1c0872	NA
fig/66666666.60966.peg.868	CDS	795019	795636	1	+	618	FIG123464: Polysaccharide export protein	Cell wall related cluster	D23_1c0873	Neut_0891
fig/66666666.60966.peg.869	CDS	795679	797223	1	+	1545	Lipopolysaccharide biosynthesis chain length determinant protein	Cell wall related cluster	D23_1c0874	Neut_0892
fig/66666666.60966.peg.870	CDS	797305	798243	1	+	939	Protein-tyrosine kinase (EC 2.7.1.112)	Cell wall related cluster	D23_1c0875	Neut_0893
fig/66666666.60966.peg.871	CDS	798243	799823	3	+	1581	Glycine-rich cell wall structural protein	Cell wall related cluster	D23_1c0876	Neut_0894
fig/66666666.60966.peg.872	CDS	799837	800670	1	+	834	precursor	Cell wall related cluster	D23_1c0877	Neut_0895
fig/66666666.60966.peg.873	CDS	800676	801515	3	+	840	FIG004655: Polysaccharide deacetylase	Cell wall related cluster	D23_1c0878	Neut_0896

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 GbkId	C91 Alias
fig/66666666.60966.peg.875	CDS	801827	802591	2	+	765	FIG070318: hypothetical protein	Cell wall related cluster	D23_1c0879	Neut_0897
fig/66666666.60966.peg.876	CDS	802597	803808	1	+	1212	FIG137776: Glycosyltransferase	Cell wall related cluster	D23_1c0880	Neut_0898
fig/66666666.60966.peg.877	CDS	803877	805457	3	+	1581	Eight transmembrane protein EpsH/EpsI	Cell wall related cluster; Cell wall related cluster	D23_1c0881	Neut_0899
fig/66666666.60966.peg.878	CDS	805493	806635	2	+	1143	FIG040338: Glycosyl transferase	Cell wall related cluster	D23_1c0882	Neut_0900
fig/66666666.60966.peg.879	CDS	806677	808611	1	+	1935	Asparagine synthetase [glutamine-hydrolyzing] (EC 6.3.5.4) AsnH	Cell wall related cluster; Glutamine, Aspartate and Asparagine Biosynthesis	D23_1c0884	Neut_0901
fig/66666666.60966.peg.880	CDS	808662	809654	3	+	993	FIG00859061: hypothetical protein	-none-	D23_1c0885	Neut_0902
fig/66666666.60966.peg.881	CDS	809654	810592	2	+	939	FIG00859041: hypothetical protein	-none-	D23_1c0886	Neut_0903
fig/66666666.60966.peg.882	CDS	810623	811849	2	+	1227	hypothetical protein	-none-	D23_1c0887	Neut_0903
fig/66666666.60966.peg.883	CDS	812814	811852	-3	-	963	Mobile element protein	-none-	D23_1c0888	Neut_1746
fig/66666666.60966.peg.884	CDS	813957	813094	-3	-	864	Mobile element protein	-none-	D23_1c0889	Neut_2192
fig/66666666.60966.peg.885	CDS	814250	813954	-2	-	297	hypothetical protein	-none-	D23_1c0890	Neut_2193
fig/66666666.60966.peg.887	CDS	814624	815706	1	+	1083	glycosyltransferase	-none-	D23_1c0891	NA
fig/66666666.60966.peg.888	CDS	817016	816339	-2	-	678	hypothetical protein	-none-	D23_1c0892	NA
fig/66666666.60966.peg.889	CDS	818253	817114	-3	-	1140	hypothetical protein	-none-	D23_1c0893	Neut_0906
fig/66666666.60966.peg.890	CDS	819313	818282	-1	-	1032	hypothetical protein	-none-	D23_1c0894	Neut_2116
fig/66666666.60966.peg.891	CDS	820446	819313	-3	-	1134	hypothetical protein	-none-	D23_1c0895	Neut_0905
fig/66666666.60966.peg.892	CDS	821935	820481	-1	-	1455	hypothetical protein	-none-	D23_1c0896	Neut_0909
fig/66666666.60966.peg.893	CDS	823840	822074	-1	-	1767	Asparagine synthetase [glutamine-hydrolyzing] (EC 6.3.5.4)	Cyanophycin Metabolism; Glutamate and Aspartate uptake in Bacteria; Glutamine, Aspartate and Asparagine Biosynthesis	D23_1c0897	Neut_0910
fig/66666666.60966.peg.895	CDS	824361	825170	3	+	810	hypothetical protein	-none-	D23_1c0898	Neut_0911
fig/66666666.60966.peg.896	CDS	825186	826454	3	+	1269	hypothetical protein	-none-	D23_1c0899	Neut_0912
fig/66666666.60966.peg.897	CDS	827356	826457	-1	-	900	hypothetical protein	-none-	D23_1c0900	Neut_0913
fig/66666666.60966.peg.898	CDS	827907	827404	-3	-	504	Low molecular weight protein tyrosine phosphatase (EC 3.1.3.48)	LMP TP YfkJ cluster; Protein deglycation	D23_1c0901	Neut_0914
fig/66666666.60966.peg.899	CDS	828032	829777	2	+	1746	ABC transporter, fused permease and ATPase domains	-none-	D23_1c0902	Neut_0915

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld	C91 Alias
fig/66666666.60966.peg.900	CDS	830559	829798	-3	-	762	Sulfur carrier protein adenylyltransferase ThiF	Thiamin biosynthesis	D23_1c0903	Neut_0916
fig/66666666.60966.peg.901	CDS	832018	830588	-1	-	1431	Carboxyl-terminal protease (EC 3.4.21.102)	Phosphoglycerate mutase protein family	D23_1c0904	Neut_0917
fig/66666666.60966.peg.902	CDS	833380	832100	-1	-	1281	Lipoprotein NlpD	Stationary phase repair cluster	D23_1c0905	Neut_0918
fig/66666666.60966.peg.903	CDS	834129	833380	-3	-	750	Phosphoglycerate mutase (EC 5.4.2.1)	Glycolysis and Gluconeogenesis; Phosphoglycerate mutase protein family	D23_1c0906	Neut_0919
fig/66666666.60966.peg.904	CDS	834328	835086	1	+	759	Triosephosphate isomerase (EC 5.3.1.1)	CBSS-331978.3.peg.2915; Calvin-Benson cycle; Glycolysis and Gluconeogenesis	D23_1c0907	Neut_0920
fig/66666666.60966.peg.905	CDS	835105	835470	1	+	366	Preprotein translocase subunit SecE (TC 3.A.5.1.1)	Gluconeogenesis CBSS-331978.3.peg.2915	D23_1c0908	Neut_0921
fig/66666666.60966.peg.906	CDS	835677	836045	3	+	369	NADH ubiquinone oxidoreductase chain A (EC 1.6.5.3)	NADH ubiquinone oxidoreductase; Respiratory Complex I	D23_1c0910	Neut_0922
fig/66666666.60966.peg.907	CDS	836049	836525	3	+	477	NADH-ubiquinone oxidoreductase chain B (EC 1.6.5.3)	NADH ubiquinone oxidoreductase; Respiratory Complex I	D23_1c0911	Neut_0923
fig/66666666.60966.peg.908	CDS	836535	837155	3	+	621	NADH-ubiquinone oxidoreductase chain C (EC 1.6.5.3)	NADH ubiquinone oxidoreductase; Respiratory Complex I	D23_1c0912	Neut_0924
fig/66666666.60966.peg.909	CDS	837213	838466	3	+	1254	NADH-ubiquinone oxidoreductase chain D (EC 1.6.5.3)	NADH ubiquinone oxidoreductase; Respiratory Complex I	D23_1c0913	Neut_0925
fig/66666666.60966.peg.910	CDS	838475	838951	2	+	477	NADH-ubiquinone oxidoreductase chain E (EC 1.6.5.3)	NADH ubiquinone oxidoreductase; Respiratory Complex I	D23_1c0914	Neut_0926
fig/66666666.60966.peg.911	CDS	838948	840225	1	+	1278	NADH-ubiquinone oxidoreductase chain F (EC 1.6.5.3)	NADH ubiquinone oxidoreductase; Respiratory Complex I	D23_1c0915	Neut_0927
fig/66666666.60966.peg.912	CDS	840287	842692	2	+	2406	NADH-ubiquinone oxidoreductase chain G (EC 1.6.5.3)	NADH ubiquinone oxidoreductase; Respiratory Complex I	D23_1c0917	Neut_0928

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld	C91 Alias
fig/66666666.60966.peg.913	CDS	842717	843814	2	+	1098	NADH-ubiquinone oxidoreductase chain H (EC 1.6.5.3)	NADH ubiquinone oxidoreductase; Respiratory Complex I	D23_1c0918	Neut_0929
fig/66666666.60966.peg.914	CDS	843832	844320	1	+	489	NADH-ubiquinone oxidoreductase chain I (EC 1.6.5.3)	NADH ubiquinone oxidoreductase; Respiratory Complex I	D23_1c0919	Neut_0930
fig/66666666.60966.peg.915	CDS	844339	844944	1	+	606	NADH-ubiquinone oxidoreductase chain J (EC 1.6.5.3)	NADH ubiquinone oxidoreductase; Respiratory Complex I	D23_1c0920	Neut_0931
fig/66666666.60966.peg.916	CDS	845000	845305	2	+	306	NADH-ubiquinone oxidoreductase chain K (EC 1.6.5.3)	NADH ubiquinone oxidoreductase; Respiratory Complex I	D23_1c0921	Neut_0932
fig/66666666.60966.peg.917	CDS	845366	847312	2	+	1947	NADH-ubiquinone oxidoreductase chain L (EC 1.6.5.3)	NADH ubiquinone oxidoreductase; Respiratory Complex I	D23_1c0922	Neut_0933
fig/66666666.60966.peg.918	CDS	847401	848882	3	+	1482	NADH-ubiquinone oxidoreductase chain M (EC 1.6.5.3)	NADH ubiquinone oxidoreductase; Respiratory Complex I	D23_1c0923	Neut_0934
fig/66666666.60966.peg.919	CDS	848945	850390	2	+	1446	NADH-ubiquinone oxidoreductase chain N (EC 1.6.5.3)	NADH ubiquinone oxidoreductase; Respiratory Complex I	D23_1c0924	Neut_0935
fig/66666666.60966.peg.920	CDS	851362	850412	-1	-	951	L-sorbose dehydrogenase	-none-	D23_1c0925	Neut_0936
fig/66666666.60966.peg.921	CDS	851619	853541	3	+	1923	Chaperone protein HtpG	Protein chaperones	D23_1c0926	Neut_0937
fig/66666666.60966.peg.922	CDS	853878	854918	3	+	1041	WD40 domain protein beta Propeller	-none-	D23_1c0927	Neut_0938
fig/66666666.60966.peg.924	CDS	855176	855598	2	+	423	Mobile element protein	-none-	D23_1c0929	Neut_0939
fig/66666666.60966.peg.925	CDS	855805	858447	1	+	2643	Hopanoid-associated RND transporter, HpnN	Hopanones	D23_1c0930	Neut_0940
fig/66666666.60966.peg.926	CDS	859079	858483	-2	-	597	DedA protein	Colicin V and Bacteriocin Production Cluster; DedA family of inner membrane proteins; Uptake of selenate and selenite	D23_1c0931	Neut_0941
fig/66666666.60966.peg.928	CDS	861652	859424	-1	-	2229	Probable transmembrane protein	-none-	D23_1c0933	Neut_0942
fig/66666666.60966.peg.929	CDS	861629	861754	2	+	126	hypothetical protein	-none-	D23_1c0934	NA
fig/66666666.60966.peg.930	CDS	861857	861720	-2	-	138	hypothetical protein	-none-	D23_1c0935	NA

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld	C91 Alias
fig 66666666.60966.peg.931	CDS	861928	862341	1	+	414	Protoporphyrinogen IX oxidase, novel form, HemJ (EC 1.3.---)	Heme and Siroheme Biosynthesis	D23_1c0936	Neut_0943
fig 66666666.60966.peg.932	CDS	862893	862363	-3	-	531	Peptide deformylase (EC 3.5.1.88)	Bacterial RNA-metabolizing Zn-dependent hydrolases; Translation termination factors bacterial	D23_1c0937	Neut_0944
fig 66666666.60966.peg.933	CDS	863632	862886	-1	-	747	5'-methylthioadenosine phosphorylase (EC 2.4.2.28)	bacterial -none-	D23_1c0938	Neut_0945
fig 66666666.60966.peg.934	CDS	865800	863752	-3	-	2049	DNA ligase (EC 6.5.1.2)	DNA Repair Base	D23_1c0939	Neut_0946
fig 66666666.60966.peg.935	CDS	865959	866654	3	+	696	hypothetical protein	Excision	D23_1c0940	Neut_0947
fig 66666666.60966.peg.936	CDS	867392	867021	-2	-	372	Flagellar biosynthesis protein Flit	-none-Flagellum	D23_1c0941	Neut_0948
fig 66666666.60966.peg.937	CDS	867859	867389	-1	-	471	Flagellar biosynthesis protein Flis	Flagellum	D23_1c0942	Neut_0949
fig 66666666.60966.peg.938	CDS	869359	867917	-1	-	1443	Flagellar hook-associated protein FlitD	Flagellum	D23_1c0943	Neut_0950
fig 66666666.60966.peg.939	CDS	870186	869716	-3	-	471	hypothetical protein	-none-	D23_1c0944	Neut_0951
fig 66666666.60966.peg.940	CDS	870616	870891	1	+	276	hypothetical protein	-none-	D23_1c0945	Neut_0952
fig 66666666.60966.peg.941	CDS	871905	870955	-3	-	951	Glutathione synthetase (EC 6.3.2.3)	Glutathione; Biosynthesis and gamma-glutamyl cycle; Heat shock dnaK	D23_1c0946	Neut_0953
fig 66666666.60966.peg.942	CDS	873212	871902	-2	-	1311	Glutamate--cysteine ligase (EC 6.3.2.2), divergent, of Alpha- and Beta-proteobacteria	gene cluster extended Glutathione; Biosynthesis and gamma-glutamyl cycle	D23_1c0947	Neut_0954
fig 66666666.60966.peg.944	CDS	873411	873722	3	+	312	LSU ribosomal protein L21p	CBSS-176279.3.peg.868	D23_1c0949	Neut_0955
fig 66666666.60966.peg.945	CDS	873734	873991	2	+	258	LSU ribosomal protein L27p	CBSS-176279.3.peg.868	D23_1c0950	Neut_0956
fig 66666666.60966.peg.946	CDS	873991	874104	1	+	114	hypothetical protein	-none-	D23_1c0951	NA
fig 66666666.60966.peg.947	CDS	874121	875152	2	+	1032	GTP-binding protein Obg	CBSS-176279.3.peg.868; Universal GTPases	D23_1c0952	Neut_0957
fig 66666666.60966.peg.948	CDS	875158	876279	1	+	1122	Glutamate 5-kinase (EC 2.7.2.11)/RNA-binding C-terminal domain PUA	Proline Synthesis; Proline Synthesis	D23_1c0953	Neut_0958
fig 66666666.60966.peg.949	CDS	876330	877331	3	+	1002	InterPro IPR000379 COGs COG0429	-none-	D23_1c0954	Neut_0959

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.									
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkid C91 Alias
fig/66666666.60966.peg.950	CDS	877428	878744	3	+	1317	Phosphate regulon sensor protein PhoR (SphS) (EC 2.7.13.3)	High affinity phosphate transporter and control of PHO regulon; PhoR-PhoB two-component regulatory system; Phosphate metabolism	D23_1c0955 Neut_0960
fig/66666666.60966.peg.951	CDS	879101	879343	2	+	243	RNA-binding protein Hfq	Hfl operon; Polyadenylation bacterial; Possible RNA degradation cluster	D23_1c0957 Neut_0961
fig/66666666.60966.peg.952	CDS	879345	880478	3	+	1134	GTP-binding protein HflX	Hfl operon; Possible RNA degradation cluster; Universal GTPases	D23_1c0958 Neut_0962
fig/66666666.60966.peg.953	CDS	880535	881725	2	+	1191	HflK protein	Hfl operon; Scaffold proteins for [4Fe-4S] cluster assembly (MRP family)	D23_1c0959 Neut_0963
fig/66666666.60966.peg.954	CDS	881725	882603	1	+	879	HflC protein	Hfl operon; Scaffold proteins for [4Fe-4S] cluster assembly (MRP family)	D23_1c0960 Neut_0964
fig/66666666.60966.peg.955	CDS	882732	882917	3	+	186	Putative inner membrane protein YjeT (clustered with HflC)	Hfl operon	D23_1c0961 Neut_0965
fig/66666666.60966.peg.956	CDS	882992	884164	2	+	1173	ATP phosphoribosyltransferase regulatory subunit (EC 2.4.2.17)	Histidine Biosynthesis	D23_1c0962 Neut_0966
fig/66666666.60966.peg.957	CDS	884298	885596	3	+	1299	Adenylosuccinate synthetase (EC 6.3.4.4)	Purine conversions	D23_1c0963 Neut_0967
fig/66666666.60966.peg.958	CDS	885982	885665	-1	-	318	FIG00858510: hypothetical protein	-none-	D23_1c0964 Neut_0968
fig/66666666.60966.peg.959	CDS	886361	886164	-2	-	198	FIG00859475: hypothetical protein	-none-	D23_1c0966 Neut_0969
fig/66666666.60966.peg.961	CDS	889010	886635	-2	-	2376	ATP-dependent protease La (EC 3.4.21.53) Type I	Proteasome bacterial; Proteolysis in bacteria, ATP-dependent	D23_1c0967 Neut_0970
fig/66666666.60966.peg.962	CDS	889609	889160	-1	-	450	CBS domain protein	-none-	D23_1c0968 Neut_0971
fig/66666666.60966.peg.963	CDS	890160	889657	-3	-	504	hypothetical protein	-none-	D23_1c0969 Neut_0972
fig/66666666.60966.peg.964	CDS	890193	890345	3	+	153	hypothetical protein	-none-	D23_1c0970 NA
fig/66666666.60966.peg.965	CDS	890326	890442	1	+	117	hypothetical protein	-none-	D23_1c0971 NA
fig/66666666.60966.peg.966	CDS	890503	891663	1	+	1161	ABC-transporter permease protein	-none-	D23_1c0972 Neut_0973
fig/66666666.60966.peg.967	CDS	891663	892352	3	+	690	ABC transporter, ATP-binding protein	-none-	D23_1c0973 Neut_0974
fig/66666666.60966.peg.968	CDS	892533	893492	3	+	960	Membrane protein	-none-	D23_1c0974 Neut_0975

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld	C91 Alias
fig/66666666.60966.peg.969	CDS	895517	893706	-2	-	1812	Sodium/hydrogen exchanger family protein	-none-	D23_1c0976	Neut_0976
fig/66666666.60966.peg.970	CDS	895888	896574	1	+	687	FIG00859851: hypothetical protein	-none-	D23_1c0978	Neut_0977
fig/66666666.60966.peg.971	CDS	897991	897029	-1	-	963	Mobile element protein	-none-	D23_1c0979	Neut_0978
fig/66666666.60966.peg.972	CDS	899539	898112	-1	-	1428	Pyruvate kinase (EC 2.7.1.40)	Glycerate metabolism; Glycolysis and Gluconeogenesis; Pyruvate metabolism I: anaplerotic reactions, PEP	D23_1c0980	Neut_0979
fig/66666666.60966.peg.973	CDS	899838	899563	-3	-	276	hypothetical protein	-none-	D23_1c0981	Neut_0980
fig/66666666.60966.peg.974	CDS	900234	900398	3	+	165	hypothetical protein	-none-	D23_1c0982	NA
fig/66666666.60966.peg.975	CDS	903871	900686	-1	-	3186	TonB-dependent receptor	Ton and Tol transport systems	D23_1c0983	Neut_1076
fig/66666666.60966.peg.976	CDS	904856	903873	-2	-	984	hypothetical protein	-none-	D23_1c0984	Neut_1077
fig/66666666.60966.peg.977	CDS	906168	904861	-3	-	1308	putative helicase	-none-	D23_1c0985	Neut_1078
fig/66666666.60966.peg.978	CDS	908497	906392	-1	-	2106	sucrose synthase	-none-	D23_1c0988	Neut_1079
fig/66666666.60966.peg.979	CDS	910925	908787	-2	-	2139	Sucrose phosphate synthase	-none-	D23_1c0989	Neut_1080
fig/66666666.60966.peg.980	CDS	911865	910936	-3	-	930	Fructokinase (EC 2.7.1.4)	-none-	D23_1c0990	Neut_1081
fig/66666666.60966.peg.981	CDS	914147	912060	-2	-	2088	Exonuclease ABC subunit B	DNA repair, UvrABC system	D23_1c0991	Neut_1082
fig/66666666.60966.peg.982	CDS	914226	915419	3	+	1194	Aspartate aminotransferase (EC 2.6.1.1)	CBSS-216591.1.peg.168; Glutamine, Glutamate, Aspartate and Asparagine Biosynthesis; Threonine and Homoserine Biosynthesis	D23_1c0992	Neut_1083
fig/66666666.60966.peg.983	CDS	915551	915688	2	+	138	hypothetical protein	-none-	D23_1c0993	NA
fig/66666666.60966.peg.984	CDS	915891	916277	3	+	387	Mobile element protein	-none-	D23_1c0994	Neut_0884
fig/66666666.60966.peg.985	CDS	916240	916695	1	+	456	Mobile element protein	-none-	D23_1c0995	Neut_2502
fig/66666666.60966.peg.986	CDS	916764	917726	3	+	963	Mobile element protein	-none-	D23_1c0996	Neut_1862
fig/66666666.60966.peg.987	CDS	918535	918419	-1	-	117	hypothetical protein	-none-	D23_1c0997	NA
fig/66666666.60966.peg.988	CDS	919358	918504	-2	-	855	DNA-directed RNA polymerase alpha subunit (EC 2.7.7.6)	RNA polymerase bacterial	D23_1c0998	NA
fig/66666666.60966.peg.989	CDS	919525	919409	-1	-	117	hypothetical protein	-none-	D23_1c0999	Neut_1085
fig/66666666.60966.peg.990	CDS	919587	919787	3	+	201	Glutathione synthetase (EC 6.3.2.3)	Glutathione: Biosynthesis and gamma-glutamyl cycle; Heat shock dnaK gene cluster extended	D23_1c0999	Neut_1085

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld	C91 Alias
fig/6666666.60966.peg.991	CDS	919919	920287	2	+	369	FIG00858546; hypothetical protein	-none-	D23_1c1000	Neut_1086
fig/6666666.60966.peg.992	CDS	920496	921119	3	+	624	bacteriocin resistance protein, putative	-none-	D23_1c1001	Neut_1087
fig/6666666.60966.peg.993	CDS	921168	921644	3	+	477	FIG00859915; hypothetical protein	-none-	D23_1c1002	Neut_1088
fig/6666666.60966.peg.994	CDS	921659	923050	2	+	1392	FIG00858837; hypothetical protein	-none-	D23_1c1003	Neut_1089
fig/6666666.60966.peg.995	CDS	923533	923105	-1	-	429	Zn-dependent hydrolases, including glyoxylases	-none-	D23_1c1004	Neut_1090
fig/6666666.60966.peg.996	CDS	923845	924936	1	+	1092	Diaminohydroxyphosphoribosylaminopyrimidine deaminase (EC 3.5.4.26)/5-amino-6-(5-phosphoribosylamino)uracil reductase (EC 1.1.1.193)	Riboflavin, FMN and FAD metabolism; Riboflavin, FMN and FAD metabolism; FAD metabolism; Riboflavin, FMN and FAD metabolism in plants; Riboflavin, FMN and FAD metabolism in plants; Riboflavin synthesis cluster; Riboflavin synthesis cluster	D23_1c1005	Neut_1091
fig/6666666.60966.peg.997	CDS	925042	926262	1	+	1221	Glycosyl transferase, group 1	-none-	D23_1c1006	Neut_1092
fig/6666666.60966.peg.998	CDS	926363	927685	2	+	1323	UDP-glucose dehydrogenase (EC 1.1.1.22)	-none-	D23_1c1007	Neut_1093
fig/6666666.60966.peg.999	CDS	928127	928318	2	+	192	hypothetical protein	-none-	D23_1c1009	Neut_1095
fig/6666666.60966.peg.1000	CDS	928500	928315	-3	-	186	hypothetical protein	-none-	D23_1c1010	NA
fig/6666666.60966.peg.1002	CDS	929085	930362	3	+	1278	InterPro IPR001296 COGs COG0438	-none-	D23_1c1011	Neut_1096
fig/6666666.60966.peg.1003	CDS	930467	931729	2	+	1263	Coenzyme F390 synthetase	-none-	D23_1c1012	Neut_1097
fig/6666666.60966.peg.1004	CDS	932351	931731	-2	-	621	exopolysaccharide synthesis protein ExoD-related protein	-none-	D23_1c1013	Neut_1098
fig/6666666.60966.peg.1005	CDS	932631	933110	3	+	480	FIG00859304; hypothetical protein	-none-	D23_1c1014	Neut_1099
fig/6666666.60966.peg.1006	CDS	933200	934357	2	+	1158	FIG010505; hypothetical protein	-none-	D23_1c1015	Neut_1100
fig/6666666.60966.peg.1007	CDS	934409	935224	2	+	816	FIG00858774; hypothetical protein	-none-	D23_1c1016	Neut_1101
fig/6666666.60966.peg.1008	CDS	935243	936316	2	+	1074	Probable transmembrane protein	-none-	D23_1c1017	Neut_1102
fig/6666666.60966.peg.1009	CDS	936361	937383	1	+	1023	FIG000906: Predicted Permease	-none-	D23_1c1018	Neut_1103

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.									
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld C91 Alias
fig/66666666.60966.peg.1010	CDS	937391	938608	2	+	1218	CDP-alcitol phosphatidytransferase	-none-	D23_1c1019 Neut_1104
fig/66666666.60966.peg.1011	CDS	938637	939593	3	+	957	FIG00480695: hypothetical protein	-none-	D23_1c1020 Neut_1105
fig/66666666.60966.peg.1012	CDS	939607	940575	1	+	969	FIG00859274: hypothetical protein	-none-	D23_1c1021 Neut_1106
fig/66666666.60966.peg.1013	CDS	940636	941283	1	+	648	hypothetical protein FIG00859276: hypothetical protein	-none-	D23_1c1022 Neut_1107
fig/66666666.60966.peg.1014	CDS	941749	941294	-1	-	456	Zn-ribbon-containing, possibly RNA-binding protein and truncated derivatives	DNA replication cluster 1	D23_1c1023 Neut_1108
fig/66666666.60966.peg.1016	CDS	942037	942186	1	+	150	hypothetical protein	-none-	D23_1c1024 NA D23_1c1025 Neut_1109
fig/66666666.60966.peg.1017	CDS	942242	944971	2	+	2730	Protein export cytoplasm protein SecA ATPase RNA helicase (TC 3.A.5.1.1)	-none-	
fig/66666666.60966.peg.1018	CDS	945151	945756	1	+	606	Ubiquinol-cytochrome C reductase iron-sulfur subunit (EC 1.10.2.2)	Ubiquinone Menquinone-cytochrome c reductase complexes	D23_1c1026 Neut_1110
fig/66666666.60966.peg.1019	CDS	945758	947005	2	+	1248	Ubiquinol-cytochrome c reductase, cytochrome B subunit (EC 1.10.2.2)	Ubiquinone Menquinone-cytochrome c reductase complexes	D23_1c1027 Neut_1111
fig/66666666.60966.peg.1020	CDS	947002	947706	1	+	705	ubiquinol cytochrome C oxidoreductase, cytochrome C1 subunit	Ubiquinone Menquinone-cytochrome c reductase complexes	D23_1c1028 Neut_1112
fig/66666666.60966.peg.1021	CDS	947758	948357	1	+	600	Stringent starvation protein A	Carbon Starvation	D23_1c1029 Neut_1113
fig/66666666.60966.peg.1023	CDS	949015	949215	1	+	201	hypothetical protein	-none-	D23_1c1031 Neut_2449
fig/66666666.60966.peg.1024	CDS	949203	949682	3	+	480	Mobile element protein	-none-	D23_1c1032 Neut_2417
fig/66666666.60966.peg.1025	CDS	950361	950651	3	+	291	Mobile element protein	-none-	D23_1c1034 Neut_2190
fig/66666666.60966.peg.1026	CDS	951943	950696	-1	-	1248	Mobile element protein	-none-	D23_1c1035 Neut_0357
fig/66666666.60966.peg.1027	CDS	952378	952208	-1	-	171	hypothetical protein	-none-	D23_1c1036 NA
fig/66666666.60966.peg.1028	CDS	953114	952689	-2	-	426	C4-type zinc finger protein, DksA/TrfA family	Zinc regulated enzymes	D23_1c1037 Neut_1117
fig/66666666.60966.peg.1029	CDS	955194	953479	-3	-	1716	Adenylate cyclase (EC 4.6.1.1)	cAMP signaling in bacteria	D23_1c1038 Neut_1118
fig/66666666.60966.peg.1030	CDS	956008	955157	-1	-	852	HD domain	-none-	D23_1c1039 Neut_1119
fig/66666666.60966.peg.1032	CDS	956563	957111	1	+	549	InterPro IPR000345	-none-	D23_1c1042 Neut_1126
fig/66666666.60966.peg.1033	CDS	958663	957200	-1	-	1464	3-polyprenyl-4-hydroxybenzoate carboxy-lyase (EC 4.1.1.—)	Ubiquinone Biosynthesis; Ubiquinone Biosynthesis-gio	D23_1c1043 Neut_1127

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.									
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld C91 Alias
fig/66666666.60966.peg.1034	CDS	959616	958756	-3	-	861	4-hydroxybenzoate polyprenyltransferase (EC 2.5.1.39)	Ubiquinone Biosynthesis; Ubiquinone Biosynthesis-glo	D23_1c1045 Neut_1128
fig/66666666.60966.peg.1035	CDS	960159	959695	-3	-	465	Chorismate-pyruvate lyase (EC 4.1.3.40)	Ubiquinone Biosynthesis; Ubiquinone Biosynthesis-glo -none-	D23_1c1046 Neut_1129
fig/66666666.60966.peg.1036	CDS	960706	960299	-1	-	408	twitching motility protein PilG	Hyperosmotic potassium uptake; Potassium homeostasis; Potassium homeostasis	D23_1c1047 Neut_1130
fig/66666666.60966.peg.1037	CDS	962283	960874	-3	-	1410	Potassium uptake protein TrkH	Hyperosmotic potassium uptake; Potassium homeostasis; Potassium homeostasis	D23_1c1048 Neut_1131
fig/66666666.60966.peg.1038	CDS	963803	962355	-2	-	1449	Trk system potassium uptake protein TrkA	Bacterial RNA-metabolizing Zn-dependent hydrolases; Hyperosmotic potassium uptake; Possible RNA degradation cluster; Potassium homeostasis; Potassium homeostasis -none-	D23_1c1049 Neut_1132
fig/66666666.60966.peg.1039	CDS	965344	964070	-1	-	1275	FIG00858490: hypothetical protein	Carbon Starvation	D23_1c1050 Neut_1133
fig/66666666.60966.peg.1040	CDS	965355	965495	3	+	141	hypothetical protein	Carbon Starvation	D23_1c1051 NA
fig/66666666.60966.peg.1041	CDS	965968	965477	-1	-	492	Starvation lipoprotein Slp paralog	Carbon Starvation	D23_1c1052 Neut_1134
fig/66666666.60966.peg.1042	CDS	966371	967021	2	+	651	Septum site-determining protein MinC	Bacterial Cell Division; Bacterial Cytoskeleton; Bacterial cell division cluster; Septum site-determining cluster Min	D23_1c1054 Neut_1135
fig/66666666.60966.peg.1043	CDS	967047	967856	3	+	810	Septum site-determining protein MinD	Bacterial Cell Division; Bacterial Cytoskeleton; Bacterial cell division cluster; Septum site-determining cluster Min	D23_1c1055 Neut_1136
fig/66666666.60966.peg.1044	CDS	967856	968152	2	+	297	Cell division topological specificity factor MinE	Bacterial Cell Division; Bacterial Cytoskeleton;	D23_1c1056 Neut_1137

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.									
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkid C91 Alias
fig/66666666.60966.peg.1045	CDS	968284	968895	1	+	612	Outer membrane lipoprotein LoIB	 Bacterial cell division cluster; Septum site-determining cluster Min A Gammaproteobacteria Cluster Relating to Translation; Lipopolysaccharide assembly; Lipoprotein sorting system	D23_1c1057 Neut_1138
fig/66666666.60966.peg.1046	CDS	968920	969756	1	+	837	4-diphosphocytidyl-2-C-methyl-D-erythritol kinase (EC 2.7.1.148)	A Gammaproteobacteria Cluster Relating to Translation; Isoprenoid Biosynthesis; Nonmevalonate Branch of Isoprenoid Biosynthesis	D23_1c1058 Neut_1139
fig/66666666.60966.peg.1047	CDS	969932	970882	2	+	951	Ribose-phosphate pyrophosphokinase (EC 2.7.6.1)	A Gammaproteobacteria Cluster Relating to Translation; De Novo Purine Biosynthesis; Pentose phosphate pathway; Transcription repair cluster	D23_1c1060 Neut_1140
fig/66666666.60966.peg.1048	CDS	970936	971544	1	+	609	LSU ribosomal protein L25p	Transcription repair cluster	D23_1c1061 Neut_1141
fig/66666666.60966.peg.1049	CDS	971667	972236	3	+	570	Peptidyl-tRNA hydrolase (EC 3.1.1.29)	Sporulation-associated proteins with broader functions; Transcription repair cluster; Translation termination factors bacterial	D23_1c1062 Neut_1142
fig/66666666.60966.peg.1050	CDS	972291	973382	3	+	1092	GTP-binding and nucleic acid-binding protein YchF	Universal GTPases	D23_1c1063 Neut_1143
fig/66666666.60966.peg.1051	CDS	973589	973461	-2	-	129	hypothetical protein	-none-	D23_1c1064 NA
fig/66666666.60966.peg.1052	CDS	973891	975303	1	+	1413	3-isopropylmalate dehydratase large subunit (EC 4.2.1.33)	Branched-Chain Amino Acid Biosynthesis; Leucine	D23_1c1066 Neut_1144
fig/66666666.60966.peg.1053	CDS	975336	975464	3	+	129	FIG00858504; hypothetical protein	Biosynthesis -none-	D23_1c1067 Neut_1145

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.									
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld
fig/66666666.60966.peg.1054	CDS	975470	976108	2	+	639	3-isopropylmalate dehydratase small subunit (EC 4.2.1.33)	Branched-Chain Amino Acid Biosynthesis; Leucine Biosynthesis	D23_1c1068
fig/66666666.60966.peg.1055	CDS	976133	977203	2	+	1071	3-isopropylmalate dehydrogenase (EC 1.1.1.85)	Branched-Chain Amino Acid Biosynthesis; Leucine Biosynthesis	D23_1c1069
fig/66666666.60966.peg.1056	CDS	977333	978457	2	+	1125	Aspartate-semialdehyde dehydrogenase (EC 1.2.1.11)	Lysine Biosynthesis DAP Pathway, GJO scratch; Threonine and Homoserine Biosynthesis	D23_1c1070
fig/66666666.60966.peg.1057	CDS	978574	980970	1	+	2397	hypothetical protein	-none-	D23_1c1071
fig/66666666.60966.peg.1058	CDS	981084	981917	3	+	834	tRNA pseudouridine synthase A (EC 4.2.1.70)	Colicin V and Bacteriocin Production Cluster; RNA pseudouridine synthases; tRNA modification	D23_1c1072
fig/66666666.60966.peg.1059	CDS	981929	982555	2	+	627	Phosphoribosylanthranilate isomerase (EC 5.3.1.24)	 tRNA processing	D23_1c1073
fig/66666666.60966.peg.1060	CDS	982542	983741	3	+	1200	Tryptophan synthase beta chain (EC 4.2.1.20)	Auxin biosynthesis; Chorismate; Intermediate for synthesis of Tryptophan, PABA antibiotics, PABA, 3-hydroxyanthranilate and more; Tryptophan synthesis	D23_1c1074
fig/66666666.60966.peg.1061	CDS	983794	984618	1	+	825	Tryptophan synthase alpha chain (EC 4.2.1.20)	Auxin biosynthesis; Chorismate; Intermediate for synthesis of Tryptophan, PABA antibiotics, PABA, 3-hydroxyanthranilate and more; Tryptophan synthesis	D23_1c1075

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld	C91 Alias
fig/66666666.60966.peg.1062	CDS	984623	985513	2	+	891	Acetyl-coenzyme A carboxyl transferase beta chain (EC 6.4.1.2)	Colicin V and Bacteriocin Production Cluster; Fatty Acid Biosynthesis FASII	D23_1c1076	Neut_1154
fig/66666666.60966.peg.1063	CDS	985642	986925	1	+	1284	Dihydrofolate synthase (EC 6.3.2.12) Folypolyglutamate synthase (EC 6.3.2.17)	Colicin V and Bacteriocin Production Cluster; Colicin V and Bacteriocin Production Cluster; Folate Biosynthesis; Folate Biosynthesis	D23_1c1077	Neut_1155
fig/66666666.60966.peg.1064	CDS	986945	987616	2	+	672	DedD protein	Colicin V and Bacteriocin Production Cluster	D23_1c1078	Neut_1156
fig/66666666.60966.peg.1065	CDS	987613	988107	1	+	495	Colicin V production protein	Colicin V and Bacteriocin Production Cluster	D23_1c1079	Neut_1157
fig/66666666.60966.peg.1066	CDS	988214	989731	2	+	1518	Amidophosphoribosyltransferase (EC 2.4.2.14)	Colicin V and Bacteriocin Production Cluster; De Novo Purine Biosynthesis	D23_1c1080	Neut_1158
fig/66666666.60966.peg.1067	CDS	989748	990923	3	+	1176	O-acetylhomoserine sulphydrylase (EC 2.5.1.49)/O-succinylhomoserine sulphydrylase (EC 2.5.1.48)	Methionine Biosynthesis; Methionine Biosynthesis	D23_1c1081	Neut_1159
fig/66666666.60966.peg.1068	CDS	991043	992554	2	+	1512	Threonine dehydratase biosynthetic (EC 4.3.1.19)	Branched-Chain Amino Acid Biosynthesis	D23_1c1082	Neut_1160
fig/66666666.60966.peg.1069	CDS	992670	992536	-3	-	135	hypothetical protein	-none-	D23_1c1083	NA
fig/66666666.60966.peg.1070	CDS	993016	994950	1	+	1935	twitching motility protein PilJ	-none-	D23_1c1084	Neut_1161
fig/66666666.60966.peg.1071	CDS	995179	995054	-1	-	126	hypothetical protein	-none-	D23_1c1085	Neut_1162
fig/66666666.60966.peg.1072	CDS	995174	1000177	2	+	5004	Signal transduction histidine kinase CheA (EC 2.7.3.—)	Flagellar motility	D23_1c1085	Neut_1162
fig/66666666.60966.peg.1073	CDS	1000248	1002095	3	+	1848	ParB-like nuclease domain	-none-	D23_1c1086	Neut_1163
fig/66666666.60966.peg.1074	CDS	1003717	1002203	-1	-	1515	Ubiquinone biosynthesis monooxygenase UbiB	Ubiquinone Biosynthesis; Ubiquinone	D23_1c1087	Neut_1164
fig/66666666.60966.peg.1075	CDS	1004439	1003807	-3	-	633	Protein YigP (COG3165) clustered with ubiquinone biosynthetic genes	Biosynthesis-gio Ubiquinone Biosynthesis; Ubiquinone	D23_1c1088	Neut_1165
fig/66666666.60966.peg.1076	CDS	1004959	1004528	-1	-	432	FIG00858586; hypothetical protein	 Ubiquinone Biosynthesis-gio	D23_1c1089	Neut_1166
fig/66666666.60966.peg.1077	CDS	1005165	1004962	-3	-	204	hypothetical protein	-none-	D23_1c1090	NA

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld	C91 Alias
fig/66666666.60966.peg.1078	CDS	1005185	1007323	2	+	2139	Signal transduction histidine kinase CheA (EC 2.7.3.—)	Flagellar motility	D23_1c1091	Neut_1167
fig/66666666.60966.peg.1079	CDS	1007395	1007916	1	+	522	Positive regulator of CheA protein activity (CheW)	-none-	D23_1c1092	Neut_1168
fig/66666666.60966.peg.1080	CDS	1008003	1010258	3	+	2256	Methyl-accepting chemotaxis protein I (serine chemoreceptor protein)	-none-	D23_1c1093	Neut_1169
fig/66666666.60966.peg.1082	CDS	1010399	1012744	2	+	2346	Methyl-accepting chemotaxis protein I (serine chemoreceptor protein)	-none-	D23_1c1094	Neut_1170
fig/66666666.60966.peg.1083	CDS	1012932	1013807	3	+	876	Chemotaxis protein methyltransferase CheR (EC 2.1.1.80)	-none-	D23_1c1095	Neut_1171
fig/66666666.60966.peg.1084	CDS	1013887	1014471	1	+	585	Chemotaxis protein CheD	-none-	D23_1c1096	Neut_1172
fig/66666666.60966.peg.1085	CDS	1014502	1015569	1	+	1068	Chemotaxis response regulator protein-glutamate methyltransferase CheB (EC 3.1.1.61)	-none-	D23_1c1097	Neut_1173
fig/66666666.60966.peg.1087	CDS	1017858	1016338	-3	-	1521	Ferredoxin reductase	Anaerobic respiratory reductases	D23_1c1099	Neut_1175
fig/66666666.60966.peg.1088	CDS	1018141	1019145	1	+	1005	Fructose-1,6-bisphosphatase, type I (EC 3.1.3.11)	Calvin-Benson cycle; Glycolysis and Gluconeogenesis	D23_1c1100	Neut_1176
fig/66666666.60966.peg.1089	CDS	1020142	1019183	-1	-	960	Glutathione S-transferase, omega (EC 2.5.1.18)	Glutathione: Non-redox reactions	D23_1c1101	Neut_1177
fig/66666666.60966.peg.1090	CDS	1020604	1020146	-1	-	459	Membrane protein, distant similarity to thiosulphate:quinone oxidoreductase DoxD	-none-	D23_1c1102	Neut_1178
fig/66666666.60966.peg.1091	CDS	1020710	1020862	2	+	153	hypothetical protein	-none-	D23_1c1103	NA
fig/66666666.60966.peg.1092	CDS	1022088	1021138	-3	-	951	COGs COG0726	-none-	D23_1c1104	Neut_1179
fig/66666666.60966.peg.1093	CDS	1022664	1022212	-3	-	453	Phosphohistidine phosphatase SixA	-none-	D23_1c1105	Neut_1180
fig/66666666.60966.peg.1094	CDS	1023442	1022753	-1	-	690	COGs COG1814	-none-	D23_1c1106	Neut_1181
fig/66666666.60966.peg.1095	CDS	1024149	1023445	-3	-	705	probable carboxylesterase	-none-	D23_1c1107	Neut_1182
fig/66666666.60966.peg.1096	CDS	1024228	1025334	1	+	1107	InterPro IPR002931 COGs COG1305	-none-	D23_1c1108	Neut_1183
fig/66666666.60966.peg.1097	CDS	1027262	1025541	-2	-	1722	Sulfite reductase [NADPH] hemoprotein	Cysteine Biosynthesis; Inorganic Sulfur	D23_1c1109	Neut_1184

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld	C91 Alias
fig/66666666.60966.peg.1098	CDS	1029106	1027271	-1	-	1836	beta-component (EC 1.8.1.2) Sulfite reductase [NADPH] flavoprotein alpha-component (EC 1.8.1.2)	Assimilation Cysteine Biosynthesis; Inorganic Sulfur Assimilation	D23_1c1110	Neut_1185
fig/66666666.60966.peg.1100	CDS	1030580	1029651	-2	-	930	Cys regulon transcriptional activator	Cysteine Biosynthesis; LysR-family proteins in <i>Escherichia coli</i>	D23_1c1111	Neut_1186
fig/66666666.60966.peg.1101	CDS	1030815	1031513	3	+	699	CysB Phosphoadenylyl-sulfate reductase [thioredoxin] (EC 1.8.4.8)/Adenylyl-sulfate reductase [thioredoxin] (EC 1.8.4.10)	Cysteine Biosynthesis; Inorganic Sulfur Assimilation; Inorganic Sulfur Assimilation	D23_1c1112	Neut_1187
fig/66666666.60966.peg.1102	CDS	1031599	1032447	1	+	849	Sulfate adenylyltransferase subunit 2 (EC 2.7.7.4)	Cysteine Biosynthesis; Inorganic Sulfur Assimilation	D23_1c1113	Neut_1188
fig/66666666.60966.peg.1103	CDS	1032508	1033791	1	+	1284	Sulfate adenylyltransferase subunit 1 (EC 2.7.7.4)	Cysteine Biosynthesis; Inorganic Sulfur Assimilation	D23_1c1114	Neut_1189
fig/66666666.60966.peg.1105	CDS	1033967	1037446	2	+	3480	Glutamate synthase [NADPH] small chain (EC 1.4.1.13)	Ammonia assimilation; Glutamine, Glutamate, Aspartate and Asparagine Biosynthesis	D23_1c1115	Neut_1190
fig/66666666.60966.peg.1106	CDS	1037596	1038723	1	+	1128	NAD(P) transhydrogenase alpha subunit (EC 1.6.1.2)	Phosphate metabolism	D23_1c1116	Neut_1191
fig/66666666.60966.peg.1107	CDS	1038778	1039086	1	+	309	NAD(P) transhydrogenase alpha subunit (EC 1.6.1.2)	Phosphate metabolism	D23_1c1117	Neut_1192
fig/66666666.60966.peg.1108	CDS	1039087	1040466	1	+	1380	NAD(P) transhydrogenase subunit beta (EC 1.6.1.2)	Phosphate metabolism	D23_1c1118	Neut_1193
fig/66666666.60966.peg.1109	CDS	1041147	1040488	-3	-	660	FIG00858826; hypothetical protein	-none-	D23_1c1119	Neut_1194
fig/66666666.60966.peg.1110	CDS	1042046	1041582	-2	-	465	Bacterioferritin	-none-	D23_1c1120	Neut_1195
fig/66666666.60966.peg.1112	CDS	1043119	1043253	1	+	135	hypothetical protein	-none-	D23_1c1121	NA
fig/66666666.60966.peg.1113	CDS	1043368	1043511	1	+	144	hypothetical protein	-none-	D23_1c1122	NA
fig/66666666.60966.peg.1115	CDS	1043900	1043736	-2	-	165	hypothetical protein	-none-	D23_1c1123	NA
fig/66666666.60966.peg.1116	CDS	1043900	1044847	2	+	948	2,3-bisphosphoglycerate-independent phosphoglycerate mutase	-none-	D23_1c1124	Neut_1198

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld	C91 Alias
fig/66666666.60966.peg.1117	CDS	1044958	1045530	1	+	573	InterPro IPR000014;IPR001633 COGs COG2200	-none-	D23_1c1125	Neut_1199
fig/66666666.60966.peg.1118	CDS	1045566	1045859	3	+	294	Mobile element protein	-none-	D23_1c1126	Neut_1719
fig/66666666.60966.peg.1119	CDS	1045958	1046836	2	+	879	Mobile element protein	-none-	D23_1c1127	Neut_1720
fig/66666666.60966.peg.1120	CDS	1046891	1047784	2	+	894	InterPro IPR000014;IPR001633 COGs COG2200	-none-	D23_1c1128	Neut_1199
fig/66666666.60966.peg.1121	CDS	1048688	1047801	-2	-	888	Phosphoribosylaminoimidazole-succinocarboxamide synthase (EC 6.3.2.6)	De Novo Purine Biosynthesis	D23_1c1129	Neut_1200
fig/66666666.60966.peg.1122	CDS	1049856	1048726	-3	-	1131	Phosphoribosylaminoimidazole carboxylase ATPase subunit (EC 4.1.1.21)	De Novo Purine Biosynthesis	D23_1c1130	Neut_1201
fig/66666666.60966.peg.1123	CDS	1050317	1049847	-2	-	471	Phosphoribosylaminoimidazole carboxylase catalytic subunit (EC 4.1.1.21)	De Novo Purine Biosynthesis	D23_1c1131	Neut_1202
fig/66666666.60966.peg.1124	CDS	1050521	1050955	2	+	435	Biopolymer transport protein ExbD/TolR	Ton and Tol transport systems	D23_1c1133	Neut_1203
fig/66666666.60966.peg.1125	CDS	1050942	1051664	3	+	723	Superoxide dismutase [Fe] (EC 1.15.1.1)	Oxidative stress; Protection from Reactive Oxygen Species	D23_1c1134	Neut_1204
fig/66666666.60966.peg.1126	CDS	1051674	1052324	3	+	651	ATP phosphoribosyltransferase (EC 2.4.2.17)	Histidine Biosynthesis; Riboflavin synthesis cluster	D23_1c1135	Neut_1205
fig/66666666.60966.peg.1127	CDS	1052409	1053644	3	+	1236	Histidinol dehydrogenase (EC 1.1.1.23)	Histidine Biosynthesis	D23_1c1136	Neut_1206
fig/66666666.60966.peg.1128	CDS	1053708	1054883	3	+	1176	2-octaprenyl-6-methoxyphenol hydroxylase (EC 1.14.13.--)	CBSS-87626.3.peg.3639; Ubiquinone Biosynthesis; Ubiquinone Biosynthesis-gio	D23_1c1137	Neut_1207
fig/66666666.60966.peg.1129	CDS	1055047	1056060	1	+	1014	tRNA dihydrouridine synthase B (EC 1.--.--)	Possible RNA degradation cluster; tRNA modification	D23_1c1138	Neut_1208
fig/66666666.60966.peg.1130	CDS	1056057	1056299	3	+	243	DNA-binding protein Fis	Bacteria DNA structural proteins, bacterial	D23_1c1139	Neut_1209
fig/66666666.60966.peg.1131	CDS	1056309	1057871	3	+	1563	IMP cyclodiolase (EC 3.5.4.10)/ Phosphoribosylaminoimidazolecarboxamide formyltransferase (EC 2.1.2.3)	5-FCL-like protein; De Novo Purine Biosynthesis; De Novo Purine Biosynthesis	D23_1c1140	Neut_1210

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld	C91 Alias
fig/66666666.60966.peg.1132	CDS	1057947	1059233	3	+	1287	Phosphoribosylamine-glycine ligase (EC 6.3.4.13)	De Novo Purine Biosynthesis	D23_1c1141	Neut_1211
fig/66666666.60966.peg.1133	CDS	1059226	1060497	1	+	1272	FIG00858582: hypothetical protein	-none-	D23_1c1142	Neut_1213
fig/66666666.60966.peg.1134	CDS	1061025	1060582	-3	-	444	TPR repeat precursor	-none-	D23_1c1143	Neut_1214
fig/66666666.60966.peg.1135	CDS	1061496	1061041	-3	-	456	Mobile element protein	-none-	D23_1c1144	Neut_2502
fig/66666666.60966.peg.1136	CDS	1061839	1061459	-1	-	381	Mobile element protein	-none-	D23_1c1145	Neut_0884
fig/66666666.60966.peg.1137	CDS	1062885	1061836	-3	-	1050	TPR repeat precursor	-none-	D23_1c1146	Neut_0701
fig/66666666.60966.peg.1138	CDS	1064906	1062960	-2	-	1947	DinG family ATP-dependent helicase	DNA repair, bacterial DinG and relatives	D23_1c1148	Neut_1215
fig/66666666.60966.peg.1139	CDS	1065051	1064932	-3	-	120	YoaA	-none-	D23_1c1149	NA
fig/66666666.60966.peg.1140	CDS	1065126	1067303	3	+	2178	Outer membrane protein Imp, required for envelope biogenesis/ Organic solvent tolerance protein precursor	EC49-61; ECSIG4-SIG7; Lipopolysaccharide assembly	D23_1c1150	Neut_1216
fig/66666666.60966.peg.1141	CDS	1067300	1068643	2	+	1344	Survival protein SurA precursor (Peptidyl-prolyl cis-trans isomerase SurA) (EC 5.2.1.8)	EC49-61; ECSIG4-SIG7; Lipopolysaccharide assembly; Peptidyl-prolyl cis-trans isomerase; Periplasmic Stress Response	D23_1c1151	Neut_1217
fig/66666666.60966.peg.1142	CDS	1068713	1069750	2	+	1038	4-hydroxythreonine-4-phosphate dehydrogenase (EC 1.1.1.262)	EC49-61; ECSIG4-SIG7; Pyridoxin (Vitamin B6) Biosynthesis	D23_1c1152	Neut_1218
fig/66666666.60966.peg.1143	CDS	1069754	1070524	2	+	771	SSU rRNA (adenine(1518)-N(6)/adenine(1519)-N(6))	EC49-61; ECSIG4-SIG7; RNA methylation; Ribosome biogenesis bacterial	D23_1c1153	Neut_1219
fig/66666666.60966.peg.1144	CDS	1071020	1070517	-2	-	504	dimethyltransferase (EC 2.1.1.182) ## SSU rRNA m6.m6-A1518-1519	DNA repair, bacterial	D23_1c1154	Neut_1220
fig/66666666.60966.peg.1145	CDS	1071346	1071080	-1	-	267	Methylated-DNA--protein-cysteine methyltransferase (EC 2.1.1.63)	DNA repair, bacterial; DNA repair, bacterial	D23_1c1155	Neut_1221

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.									
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld C91 Alias
fig/66666666.60966.peg.1146	CDS	1072448	1071324	-2	-	1125	ADA regulatory protein/ methylated-DNA-- protein-cysteine methyltransferase (EC 2.1.1.63)	DNA repair, bacterial; DNA repair, bacterial	D23_1c1156 Neut_1221
fig/66666666.60966.peg.1149	CDS	1073291	1073809	2	+	519	Helix-turn-helix motif	-none-	
fig/66666666.60966.peg.1150	CDS	1075783	1074353	-1	-	1431	Siroheme synthase/ Precorin-2 oxidase (EC 1.3.1.76)	Heme and Siroheme Biosynthesis; Heme and Siroheme Biosynthesis; Heme and Siroheme Biosynthesis	Neut_1223 Neut_1002
							Sirohydrochlorin ferrochelatase (EC 4.99.1.4)		
							Uroporphyrinogen-III methyltransferase (EC 2.1.1.107)		
fig/66666666.60966.peg.1151	CDS	1076939	1075926	-2	-	1014	Phosphate ABC transporter, periplasmic phosphate-binding protein PstS (TC 3.A.1.7.1)	High affinity phosphate transporter and control of PHO regulon; PhoR-PhoB two-component regulatory system; Phosphate metabolism	D23_1c1160 Neut_1001
fig/66666666.60966.peg.1152	CDS	1078435	1077059	-1	-	1377	Phosphoglucosamine mutase (EC 5.4.2.10)	Sialic Acid Metabolism; UDP-N-acetylmutaminate from Fructose-6-phosphate Biosynthesis	D23_1c1161 Neut_1000
fig/66666666.60966.peg.1153	CDS	1079505	1078603	-3	-	903	Dihydropterate synthase (EC 2.5.1.15)	Folate Biosynthesis	D23_1c1162 Neut_0999
fig/66666666.60966.peg.1154	CDS	1081457	1079529	-2	-	1929	Cell division protein FisH (EC 3.4.24.--)	Bacterial Cell Division	D23_1c1163 Neut_0998
fig/66666666.60966.peg.1155	CDS	1082161	1081541	-1	-	621	Cell division protein FtsJ/ Ribosomal RNA large subunit methyltransferase E (EC 2.1.1.--) ## LSU rRNA Um2552	Bacterial Cell Division; RNA methylation	D23_1c1164 Neut_0997
fig/66666666.60966.peg.1156	CDS	1082195	1082575	2	+	381	FIG004454; RNA binding protein	-none-	D23_1c1165 Neut_0996
fig/66666666.60966.peg.1157	CDS	1082674	1083084	1	+	411	CBS domain	-none-	D23_1c1166 Neut_0995
fig/66666666.60966.peg.1158	CDS	1083081	1083464	3	+	384	tRNA nucleotidyltransferase, A-adding (EC 2.7.7.25)	A cluster relating to Tryptophanyl-tRNA synthetase; Polyadenylation bacterial; tRNA nucleotidyltransferase	D23_1c1167 Neut_0994

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 GbkId	C91 Alias
fig/66666666.60966.peg.1159	CDS	1084218	1083646	-3	-	573	DNA-3-methyladenine glycosylase II (EC 3.2.2.21)	DNA Repair Base Excision	D23_1c1168	Neut_0993
fig/66666666.60966.peg.1160	CDS	1084717	1084241	-1	-	477	Glutathione peroxidase (EC 1.11.1.9)	Glutathione: Redox cycle	D23_1c1169	Neut_0992
fig/66666666.60966.peg.1161	CDS	1084748	1085341	2	+	594	Carbonic anhydrase, gamma class (EC 4.2.1.1)	Zinc regulated enzymes	D23_1c1170	Neut_0991
fig/66666666.60966.peg.1162	CDS	1085402	1086241	2	+	840	Putative NAD(P)-dependent oxidoreductase EC-YbbO	-none-	D23_1c1171	Neut_0990
fig/66666666.60966.peg.1163	CDS	1086401	1087885	2	+	1485	alpha amylase, catalytic region	-none-	D23_1c1172	Neut_0989
fig/66666666.60966.peg.1164	CDS	1090985	1087899	-2	-	3087	RND multidrug efflux transporter, Acriflavin resistance protein	Multidrug Resistance Efflux Pumps	D23_1c1173	Neut_0988
fig/66666666.60966.peg.1165	CDS	1092079	1090982	-1	-	1098	Membrane fusion protein of RND family multidrug efflux pump	Multidrug Resistance Efflux Pumps	D23_1c1174	Neut_0987
fig/66666666.60966.peg.1168	CDS	1092982	1092851	-1	-	132	hypothetical protein	-none-	D23_1c1175	NA
fig/66666666.60966.peg.1169	CDS	1094042	1093413	-2	-	630	Nicotinamidase family protein YcaC	NAD and NADP cofactor biosynthesis global	D23_1c1176	NA
fig/66666666.60966.peg.1170	CDS	1094408	1094223	-2	-	186	Mobile element protein	-none-	D23_1c1178	Neut_1984
fig/66666666.60966.peg.1171	CDS	1094750	1094514	-2	-	237	Mobile element protein	-none-	D23_1c1179	Neut_1353
fig/66666666.60966.peg.1172	CDS	1095334	1094984	-1	-	351	hypothetical protein	-none-	D23_1c1180	NA
fig/66666666.60966.peg.1173	CDS	1095854	1095486	-2	-	369	Mobile element protein	-none-	D23_1c1181	Neut_2502
fig/66666666.60966.peg.1175	CDS	1096049	1095933	-2	-	117	Mobile element protein	-none-	D23_1c1182	Neut_0884
fig/66666666.60966.peg.1176	CDS	1096623	1096288	-3	-	336	putative (AJ245540) NrdJ [Wolinetella succinogenes]	-none-	D23_1c1183	Neut_1054
fig/66666666.60966.peg.1177	CDS	1098408	1096669	-3	-	1740	FIG00859793: hypothetical protein	-none-	D23_1c1184	Neut_1053
fig/66666666.60966.peg.1178	CDS	1098651	1100255	3	+	1605	Multicopper oxidase	Copper homeostasis	D23_1c1185	Neut_1052
fig/66666666.60966.peg.1179	CDS	1101152	1100349	-2	-	804	Inositol-1-monophosphatase (EC 3.1.3.25)	-none-	D23_1c1186	Neut_1051
fig/66666666.60966.peg.1180	CDS	1101712	1101155	-1	-	558	Alkyl hydroperoxide reductase and/or thiol-specific antioxidant family (AhpC/TSA) protein	-none-	D23_1c1187	Neut_1050
fig/66666666.60966.peg.1181	CDS	1101810	1102571	3	+	762	Ribosomal RNA small subunit methyltransferase E (EC 2.1.1.—)	Heat shock dnaK gene cluster extended; RNA methylation	D23_1c1188	Neut_1049

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 GbkId	C91 Alias
fig/66666666.60966.peg.1182	CDS	1102595	1103929	2	+	1335	N-acetylglutamate synthase (EC 2.3.1.1)	Arginine Biosynthesis--glo; Arginine Biosynthesis extended -none-	D23_1c1189	Neut_1048
fig/66666666.60966.peg.1183	CDS	1104924	1104208	-3	-	717	FIG002842: hypothetical protein		D23_1c1190	Neut_1046
fig/66666666.60966.peg.1184	CDS	1105647	1105036	-3	-	612	Dephospho-CoA kinase (EC 2.7.1.24)	Coenzyme A Biosynthesis	D23_1c1191	Neut_1045
fig/66666666.60966.peg.1185	CDS	1106511	1105651	-3	-	861	Leader peptidase (Prepilin peptidase) (EC 3.4.23.43)/N-methyltransferase (EC 2.1.1.--)	-none-	D23_1c1192	Neut_1044
fig/66666666.60966.peg.1186	CDS	1107775	1106555	-1	-	1221	Type IV fimbrial assembly protein PilC	-none-	D23_1c1193	Neut_1043
fig/66666666.60966.peg.1187	CDS	1108692	1107835	-3	-	858	Nucleoside-diphosphate-sugar epimerases	CBSS-296591.1.peg.2330	D23_1c1194	Neut_1042
fig/66666666.60966.peg.1189	CDS	1108961	1109818	2	+	858	2-polyprenylphenol hydroxylase and related flavodoxin oxidoreductases/CDP-6-deoxy--delta-3,4-glucoseen reductase-like	-none-	D23_1c1195	Neut_1041
fig/66666666.60966.peg.1190	CDS	1111006	1109825	-1	-	1182	Homolog of <i>E. coli</i> HemY protein	-none-	D23_1c1196	Neut_1040
fig/66666666.60966.peg.1191	CDS	1112031	1111003	-3	-	1029	Uroporphyrinogen-III methyltransferase (EC 2.1.1.107)	Heme and Siroheme Biosynthesis	D23_1c1197	Neut_1039
fig/66666666.60966.peg.1192	CDS	1112828	1112046	-2	-	783	Uroporphyrinogen-III synthase (EC 4.2.1.75)	Heme and Siroheme Biosynthesis	D23_1c1198	Neut_1038
fig/66666666.60966.peg.1193	CDS	1113852	1112848	-3	-	1005	Porphobilinogen deaminase (EC 2.5.1.61)	Heme and Siroheme Biosynthesis	D23_1c1199	Neut_1037
fig/66666666.60966.peg.1194	CDS	1113898	1116699	1	+	2802	Phosphoenolpyruvate carboxylase (EC 4.1.1.31)	Pyruvate metabolism I: anaerobic reactions, PEP	D23_1c1200	Neut_1036
fig/66666666.60966.peg.1195	CDS	1116922	1118904	1	+	1983	FIG00858706: hypothetical protein	-none-	D23_1c1201	Neut_1035
fig/66666666.60966.peg.1196	CDS	1120329	1118899	-3	-	1431	probable integral membrane protein	-none-	D23_1c1202	Neut_1034
fig/66666666.60966.peg.1197	CDS	1121723	1120329	-2	-	1395	NMA1898	-none-	D23_1c1203	Neut_1033
fig/66666666.60966.peg.1198	CDS	1121909	1121763	-2	-	147	hypothetical protein	-none-	D23_1c1204	NA
fig/66666666.60966.peg.1199	CDS	1121985	1122698	3	+	714	COGs COG0518	-none-	D23_1c1205	Neut_1032
fig/66666666.60966.peg.1200	CDS	1123590	1122748	-3	-	843	RNA polymerase sigma factor RpoH	Heat shock Cell division Proteases and a Methyltransferase;	D23_1c1206	Neut_1031

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.									
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkid C91 Alias
fig/66666666.60966.peg.1201	CDS	1124005	1124823	1	+	819	Cytochrome c family protein	 Heat shock dnaK gene cluster extended;	
fig/66666666.60966.peg.1203	CDS	1125802	1127289	1	+	1488	FIG00858881: hypothetical protein	 Transcription initiation, bacterial sigma factors	NA
fig/66666666.60966.peg.1204	CDS	1127318	1128325	2	+	1008	Sulfate-binding protein	-none-	D23_1c1207
fig/66666666.60966.peg.1205	CDS	1128452	1128580	2	+	129	Sbp	Inorganic Sulfur Assimilation	D23_1c1209
fig/66666666.60966.peg.1206	CDS	1128661	1129494	1	+	834	hypothetical protein	-none-	D23_1c1210
							Sulfate transport system permease	Cysteine Biosynthesis; Inorganic Sulfur	D23_1c1211
							protein CysI	Assimilation	D23_1c1212
fig/66666666.60966.peg.1207	CDS	1129502	1130371	2	+	870	Sulfate transport system permease	Cysteine Biosynthesis; Inorganic Sulfur	Neut_1025
fig/66666666.60966.peg.1208	CDS	1130383	1131471	1	+	1089	protein CysW	Assimilation	Neut_1024
							Sulfate and thiosulfate import ATP-binding protein CysA (EC 3.6.3.25)	Cysteine Biosynthesis; Inorganic Sulfur	
fig/66666666.60966.peg.1209	CDS	1131827	1131510	-2	-	318	possible lipase	Assimilation; Uptake of selenate and selenite	Neut_1023
fig/66666666.60966.peg.1210	CDS	1133403	1131859	-3	-	1545	Aminopeptidase PepA-related protein	-none-	Neut_1022
fig/66666666.60966.peg.1211	CDS	1133455	1134249	1	+	795	Thymidylate synthase (EC 2.1.1.45)	Folate Biosynthesis; pyrimidine conversions	Neut_1021
fig/66666666.60966.peg.1212	CDS	1134246	1134776	3	+	531	Dihydrofolate reductase (EC 1.5.1.3)	5-FCL-like protein; EC49-61; Folate Biosynthesis	Neut_1020
fig/66666666.60966.peg.1213	CDS	1136053	1134809	-1	-	1245	Response regulator	-none-	Neut_1019
fig/66666666.60966.peg.1215	CDS	1136321	1136785	2	+	465	Mobile element protein	-none-	Neut_0357
fig/66666666.60966.peg.1216	CDS	1136808	1137512	3	+	705	Mobile element protein	-none-	Neut_1318
fig/66666666.60966.peg.1217	CDS	1137576	1138703	3	+	1128	Catalase (EC 1.11.1.6)	Oxidative stress; Photorespiration (oxidative C2cycle); Protection from Reactive Oxygen Species	NA
fig/66666666.60966.peg.1218	CDS	1139453	1138773	-2	-	681	Mobile element protein	-none-	Neut_1318
fig/66666666.60966.peg.1219	CDS	1139703	1139837	3	+	135	Mobile element protein	-none-	Neut_2500
fig/66666666.60966.peg.1220	CDS	1139894	1140238	2	+	345	Mobile element protein	-none-	Neut_1375
fig/66666666.60966.peg.1221	CDS	1141157	1140195	-2	-	963	Mobile element protein	-none-	Neut_1278
fig/66666666.60966.peg.1222	CDS	1141250	1141762	2	+	513	Mobile element protein	-none-	Neut_1624
fig/66666666.60966.peg.1223	CDS	1142310	1141846	-3	-	465	Mobile element protein	-none-	Neut_0357
fig/66666666.60966.peg.1224	CDS	1142882	1142409	-2	-	474	Mobile element protein	-none-	Neut_0883
fig/66666666.60966.peg.1225	CDS	1144087	1143419	-1	-	669	hypothetical protein	-none-	NA

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld	C91 Alias
fig/66666666.60966.peg.1226	CDS	1144329	1144078	-3	-	252	ABC-type antimicrobial peptide transport system, permease component	-none-	D23_1c1231	NA
fig/66666666.60966.peg.1227	CDS	1144623	1145231	3	+	609	Transcriptional regulator, TetR family	-none-	D23_1c1232	Neut_1011
fig/66666666.60966.peg.1228	CDS	1145317	1146351	1	+	1035	Predicted membrane fusion protein (MFP) component of efflux pump, membrane anchor protein YbhG	ATP-dependent efflux pump transporter Ybh	D23_1c1233	Neut_1010
fig/66666666.60966.peg.1229	CDS	1146341	1148125	2	+	1785	ABC transporter multidrug efflux pump, fused ATP-binding domains	ATP-dependent efflux pump transporter Ybh	D23_1c1234	Neut_1009
fig/66666666.60966.peg.1230	CDS	1148122	1149270	1	+	1149	ABC transport system, permease component YbhS	ATP-dependent efflux pump transporter Ybh	D23_1c1235	Neut_1008
fig/66666666.60966.peg.1231	CDS	1149276	1150400	3	+	1125	ABC transport system, permease component YbhR	ATP-dependent efflux pump transporter Ybh	D23_1c1236	Neut_1007
fig/66666666.60966.peg.1232	CDS	1150415	1150546	2	+	132	hypothetical protein	-none-	D23_1c1237	NA
fig/66666666.60966.peg.1233	CDS	1150719	1150507	-3	-	213	hypothetical protein	-none-	D23_1c1238	NA
fig/66666666.60966.peg.1234	CDS	1150690	1150875	1	+	186	hypothetical protein	-none-	D23_1c1239	Neut_1006
fig/66666666.60966.peg.1235	CDS	1150882	1151043	1	+	162	hypothetical protein	-none-	D23_1c1240	NA
fig/66666666.60966.peg.1236	CDS	1151054	1151923	2	+	870	hypothetical protein	-none-	D23_1c1241	Neut_1006
fig/66666666.60966.peg.1237	CDS	1151916	1152278	3	+	363	hypothetical protein	-none-	D23_1c1242	NA
fig/66666666.60966.peg.1239	CDS	1152362	1152703	2	+	342	hypothetical protein	-none-	D23_1c1243	Neut_1005
fig/66666666.60966.peg.1240	CDS	1153760	1152756	-2	-	1005	Mobile element protein	-none-	D23_1c1244	Neut_1862
fig/66666666.60966.peg.1241	CDS	1153821	1154177	3	+	357	Putative transport system permease protein	-none-	D23_1c1245	Neut_1004
fig/66666666.60966.peg.1242	CDS	1154289	1154441	3	+	153	FIG00626672: hypothetical protein	-none-	D23_1c1246	Neut_1003
fig/66666666.60966.peg.1243	CDS	1154607	1154476	-3	-	132	hypothetical protein	-none-	D23_1c1247	Neut_1226
fig/66666666.60966.peg.1244	CDS	1156287	1154728	-3	-	1560	amino acid transporter	-none-	D23_1c1248	Neut_1227
fig/66666666.60966.peg.1245	CDS	1157495	1156611	-2	-	885	DNA recombination-dependent growth factor C	DNA repair, bacterial	D23_1c1250	Neut_1228
fig/66666666.60966.peg.1246	CDS	1158541	1157738	-1	-	804	Probable component of the lipoprotein assembly complex (forms a complex with YaeT, YfgL, and NlpB)	Lipopolysaccharide assembly	D23_1c1252	Neut_1229
fig/66666666.60966.peg.1247	CDS	1158543	1159571	3	+	1029	Ribosomal large subunit pseudouridine synthase D (EC 4.2.1.70)	RNA pseudouridine syntheses; Ribosome biogenesis bacterial	D23_1c1253	Neut_1230

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 GbkId	C91 Alias
fig/66666666.60966.peg.1248	CDS	1159795	1161201	1	+	1407	Glutamine synthetase type I (EC 6.3.1.2)	Ammonia assimilation; Glutamine, Glutamate, Aspartate and Asparagine Biosynthesis; Glutamine synthetases; Peptidoglycan Biosynthesis	D23_1c1254	Neut_1231
fig/66666666.60966.peg.1249	CDS	1161356	1161826	2	+	471	FIG00858905; hypothetical protein	-none-	D23_1c1256	Neut_1232
fig/66666666.60966.peg.1250	CDS	1161926	1161813	-2	-	114	hypothetical protein	-none-	D23_1c1257	NA
fig/66666666.60966.peg.1251	CDS	1162096	1162263	1	+	168	hypothetical protein	-none-	D23_1c1258	NA
fig/66666666.60966.peg.1252	CDS	1162260	1162415	3	+	156	hypothetical protein	-none-	D23_1c1259	NA
fig/66666666.60966.peg.1253	CDS	1162412	1163191	2	+	780	Putative sulfate permease	Inorganic Sulfur Assimilation	D23_1c1260	Neut_1235
fig/66666666.60966.peg.1254	CDS	1163827	1163267	-1	-	561	Iron-sulfur cluster assembly scaffold protein IscU/NitU-like for SUF system, SufE3	-none-	D23_1c1261	Neut_1236
fig/66666666.60966.peg.1255	CDS	1164319	1163837	-1	-	483	Putative iron-sulfur cluster assembly scaffold protein for SUF system, SufE2	-none-	D23_1c1262	Neut_1237
fig/66666666.60966.peg.1256	CDS	1165584	1164316	-3	-	1269	Cysteine desulfurase (EC 2.8.1.7), SufS subfamily	Alanine biosynthesis; mmn5U34 biosynthesis bacteria; tRNA modification	D23_1c1263	Neut_1238
fig/66666666.60966.peg.1257	CDS	1166895	1165588	-3	-	1308	Iron-sulfur cluster assembly protein SufD	Bacteria CBSS-196164.1.peg.1690; tRNA modification	D23_1c1264	Neut_1239
fig/66666666.60966.peg.1258	CDS	1167683	1166892	-2	-	792	Iron-sulfur cluster assembly ATPase protein SufC	Bacteria CBSS-196164.1.peg.1690; tRNA modification	D23_1c1265	Neut_1240
fig/66666666.60966.peg.1259	CDS	1169116	1167680	-1	-	1437	Iron-sulfur cluster assembly protein SufB	CBSS-196164.1.peg.1690; tRNA modification	D23_1c1266	Neut_1241
fig/66666666.60966.peg.1260	CDS	1169465	1169136	-2	-	330	Iron binding protein IscA for iron-sulfur cluster assembly	Bacteria Alanine biosynthesis; tRNA modification	D23_1c1267	Neut_1242
fig/66666666.60966.peg.1261	CDS	1169958	1169491	-3	-	468	Iron-sulfur cluster regulator IscR	Bacteria Alanine biosynthesis; Rrf2 family transcriptional regulators	D23_1c1268	Neut_1243

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld	C91 Alias
fig/66666666.60966.peg.1262	CDS	1170306	1171838	3	+	1533	2-isopropylmalate synthase (EC 2.3.3.13)	Branched-Chain Amino Acid Biosynthesis; Leucine	D23_1c1270	Neut_1244
fig/66666666.60966.peg.1263	CDS	1171964	1172467	2	+	504	Cytochrome c-type biogenesis protein ResA FIG016425; Soluble lytic murein transglycosylase and related regulatory proteins (some contain LysM/invasin domains)	Biosynthesis	D23_1c1271	Neut_1245
fig/66666666.60966.peg.1264	CDS	1173081	1172488	-3	-	594	Prolyl-tRNA synthetase (EC 6.1.1.15), bacterial type	Biogenesis of c-type cytochromes -none-	D23_1c1272	Neut_1246
fig/66666666.60966.peg.1265	CDS	1174775	1173069	-2	-	1707	Adenosine (5')-pentaphospho-adenosine pyrophosphohydrolase (EC 3.6.1.-)	tRNA aminoacylation, Pro	D23_1c1273	Neut_1247
fig/66666666.60966.peg.1266	CDS	1175004	1175567	3	+	564	Cytochrome c551 peroxidase (EC 1.11.1.5)	CBSS-364106.7.peg.3204; Nudix proteins (nucleoside triphosphate hydrolases); Phosphoglycerate mutase protein family	D23_1c1274	Neut_1248
fig/66666666.60966.peg.1267	CDS	1175673	1176677	3	+	1005	Sensor histidine kinase PrtB (RegB) (EC 2.7.3.-)	Protection from Reactive Oxygen Species	D23_1c1275	Neut_1249
fig/66666666.60966.peg.1269	CDS	1176892	1178190	1	+	1299	Dna binding response regulator PrtA (RegA)	Global Two-component Regulator PrtBA in Proteobacteria	D23_1c1276	Neut_1250
fig/66666666.60966.peg.1270	CDS	1178205	1178750	3	+	546	Ferrichrome-iron receptor	Global Two-component Regulator PrtBA in Proteobacteria	D23_1c1277	Neut_1251
fig/66666666.60966.peg.1271	CDS	1181021	1178853	-2	-	2169	hypotheical protein	-none-	D23_1c1278	Neut_1252
fig/66666666.60966.peg.1272	CDS	1181355	1181558	3	+	204	Mobile element protein	-none-	D23_1c1279	NA
fig/66666666.60966.peg.1273	CDS	1181609	1181776	2	+	168	hypotheical protein	-none-	D23_1c1280	NA
fig/66666666.60966.peg.1274	CDS	1181748	1181981	3	+	234	Mobile element protein	-none-	D23_1c1281	NA
fig/66666666.60966.peg.1275	CDS	1182136	1181990	-1	-	147	hypotheical protein	-none-	D23_1c1282	NA
fig/66666666.60966.peg.1277	CDS	1182799	1182386	-1	-	414	hypotheical protein	-none-	D23_1c1283	Neut_1254
fig/66666666.60966.peg.1278	CDS	1182892	1183920	1	+	1029	Mobile element protein	-none-	D23_1c1284	Neut_1746
fig/66666666.60966.peg.1280	CDS	1185348	1184569	-3	-	780	CDP-diacylglycerol--serine O-phosphatidyltransferase (EC 2.7.8.8)	Glycerolipid and Glycerophospholipid Metabolism in Bacteria	D23_1c1285	Neut_1258
fig/66666666.60966.peg.1281	CDS	1186028	1185378	-2	-	651	Phosphatidylserine decarboxylase (EC 4.1.1.65)	Glycerolipid and Glycerophospholipid Metabolism in Bacteria	D23_1c1286	Neut_1259
fig/66666666.60966.peg.1282	CDS	1187049	1186033	-3	-	1017	Ketol-acid reductoisomerase (EC 1.1.1.86)	Branched-Chain Amino Acid Biosynthesis; Coenzyme A Biosynthesis	D23_1c1287	Neut_1260

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 GbkId	C91 Alias
fig/66666666.60966.peg.1283	CDS	1187629	1187138	-1	-	492	Acetolactate synthase small subunit (EC 2.2.1.6)	Acetolactate synthase subunits; Biosynthesis Chain Amino Acid	D23_1c1288	Neut_1261
fig/66666666.60966.peg.1284	CDS	1189337	1187634	-2	-	1704	Acetolactate synthase large subunit (EC 2.2.1.6)	Acetolactate synthase subunits; Biosynthesis Chain Amino Acid	D23_1c1289	Neut_1262
fig/66666666.60966.peg.1285	CDS	1189440	1189327	-3	-	114	hypothetical protein	-none- CBSS-316057.3.peg.563; CBSS-354.1.peg.2917; Putative TldE-TldD proteolytic complex	D23_1c1290	NA
fig/66666666.60966.peg.1287	CDS	1191111	1189663	-3	-	1449	TldD protein, part of TldE/TldD proteolytic complex		D23_1c1292	Neut_1263
fig/66666666.60966.peg.1288	CDS	1192101	1191238	-3	-	864	FIG003879: Predicted amidohydrolase/Aliphatic amidase AmiE (EC 3.5.1.4)	CBSS-354.1.peg.2917	D23_1c1293	Neut_1264
fig/66666666.60966.peg.1289	CDS	1196193	1192303	-3	-	3891	FIG005080: Possible exported protein	CBSS-354.1.peg.2917	D23_1c1294	Neut_1265
fig/66666666.60966.peg.1290	CDS	1196421	1199210	3	+	2790	Glutamate-ammonia-lyase	Ammonia assimilation; CBSS-316057.3.peg.3521	D23_1c1295	Neut_1266
fig/66666666.60966.peg.1291	CDS	1200720	1200058	-3	-	663	adenylyltransferase (EC 2.7.7.42)	-none-	D23_1c1300	Neut_1268
fig/66666666.60966.peg.1292	CDS	1200901	1200761	-1	-	141	hypothetical protein	-none- Arginine Biosynthesis--glo; Arginine Biosynthesis extended Protein deglycation	D23_1c1301	NA
fig/66666666.60966.peg.1293	CDS	1200921	1202306	3	+	1386	Argininosuccinate lyase (EC 4.3.2.1)		D23_1c1302	Neut_1269
fig/66666666.60966.peg.1294	CDS	1202336	1203241	2	+	906	Ribulosamine/erythrosamine 3-kinase	Protein deglycation	D23_1c1303	Neut_1270
fig/66666666.60966.peg.1295	CDS	1203386	1204876	2	+	1491	potentially involved in protein deglycation	-none-	D23_1c1304	Neut_1271
fig/66666666.60966.peg.1297	CDS	1205984	1205553	-2	-	432	hypothetical protein	-none- -none- -none- -none- -none- Flagellar motility	D23_1c1305	Neut_1272
fig/66666666.60966.peg.1298	CDS	1206313	1205987	-1	-	327	hypothetical protein		D23_1c1306	Neut_1273
fig/66666666.60966.peg.1299	CDS	1206643	1206329	-1	-	315	CcrB protein		D23_1c1307	Neut_1274
fig/66666666.60966.peg.1300	CDS	1206663	1206902	3	+	240	hypothetical protein		D23_1c1308	NA
fig/66666666.60966.peg.1301	CDS	1207586	1206945	-2	-	642	Chemotaxis response-phosphatase CheZ		D23_1c1309	Neut_1275
fig/66666666.60966.peg.1302	CDS	1208030	1207635	-2	-	396	Chemotaxis regulator-transmits chemoreceptor signals to flagellar motor components CheY	Flagellar motility	D23_1c1310	Neut_1276

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 GbkId	C91 Alias
fig/6666666.60966.peg.1303 fig/6666666.60966.peg.1305 fig/6666666.60966.peg.1306 fig/6666666.60966.peg.1307 fig/6666666.60966.peg.1308	CDS	1209139	1208375	-1	-	765	Mobile element protein	-none-	D23_1c1311	NA
	CDS	1210466	1210134	-2	-	333	hypothetical protein	-none-	D23_1c1313	NA
	CDS	1210482	1210661	3	+	180	hypothetical protein	-none-	D23_1c1314	Neut_1280
	CDS	1212693	1210768	-3	-	1926	Cytochrome c, class I	-none-	D23_1c1315	Neut_1281
	CDS	1213523	1213044	-2	-	480	FIG00858481: hypothetical protein	-none-	D23_1c1317	Neut_1282
fig/6666666.60966.peg.1309	CDS	1213544	1214608	2	+	1065	Folate-dependent protein for Fe/S cluster synthesis/repair in oxidative stress	-none-	D23_1c1318	Neut_1283
fig/6666666.60966.peg.1310 fig/6666666.60966.peg.1311	CDS	1215290	1214631	-2	-	660	FOG; Ankyrin repeat	-none-	D23_1c1319	Neut_1284
	CDS	1215984	1215292	-3	-	693	Putative deoxyribonuclease YcH	YcH	D23_1c1320	Neut_1285
fig/6666666.60966.peg.1313	CDS	1216517	1216125	-2	-	393	Queuosine biosynthesis QueD, PTPS-I	Queuosine-Archaeosine Biosynthesis; regulated enzymes; -tRNA modification	D23_1c1321	Neut_1286
fig/6666666.60966.peg.1315 fig/6666666.60966.peg.1316 fig/6666666.60966.peg.1318 fig/6666666.60966.peg.1319	CDS	1216624	1217484	1	+	861	Radical SAM domain protein	Bacteria	D23_1c1322	Neut_1287
	CDS	1217801	1217514	-2	-	288	FIG00858571: hypothetical protein	-none-	D23_1c1323	Neut_1288
	CDS	1219170	1218214	-3	-	957	Cobalt-zinc-cadmium resistance protein CzcD	Cobalt-zinc-cadmium resistance	D23_1c1325	Neut_1289
	CDS	1221684	1219195	-3	-	2490	Lead, cadmium, zinc and mercury transporting ATPase (EC 3.6.3.3) (EC 3.6.3.5); Copper-translocating P-type ATPase (EC 3.6.3.4)	Copper Transport System; Copper homeostasis	D23_1c1326	Neut_1290
fig/6666666.60966.peg.1320	CDS	1223842	1221797	-1	-	2046	1,4-alpha-glucan (glycogen) branching enzyme, GH-13-type (EC 2.4.1.18)	Glycogen metabolism	D23_1c1327	Neut_1291
fig/6666666.60966.peg.1321 fig/6666666.60966.peg.1322	CDS	1223815	1224054	1	+	240	hypothetical protein	-none-	D23_1c1328	NA
	CDS	1224141	1225418	3	+	1278	Glucose-1-phosphate adenylyltransferase (EC 2.7.7.27)	Glycogen metabolism	D23_1c1329	Neut_1292
fig/6666666.60966.peg.1323	CDS	1225484	1227202	2	+	1719	Amylopullulanase (EC 3.2.1.1)(EC 3.2.1.41)	-none-	D23_1c1330	Neut_1293
fig/6666666.60966.peg.1324	CDS	1227268	1229292	1	+	2025	Alpha-amylase (EC 3.2.1.1)	-none-	D23_1c1331	Neut_1294
fig/6666666.60966.peg.1325 fig/6666666.60966.peg.1326	CDS	1229324	1232014	2	+	2691	hypothetical protein	-none-	D23_1c1332	Neut_1295
	CDS	1232973	1232236	-3	-	738	Uracil-DNA glycosylase, family 4	Uracil-DNA glycosylase	D23_1c1334	Neut_1296
fig/6666666.60966.peg.1327	CDS	1233520	1233032	-1	-	489	Ribosomal-protein-S18p-alanine	Bacterial RNA-metabolizing Zn-	D23_1c1335	Neut_1297

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 GbkId	C91 Alias
fig/66666666.60966.peg.1328	CDS	1234194	1233523	-3	-	672	acetyltransferase (EC 2.3.1.—)	dependent hydrolases; Ribosome biogenesis bacterial -none-	D23_1c1336	Neut_1298
fig/66666666.60966.peg.1329	CDS	1234742	1234209	-2	-	534	Inactive homolog of metal-dependent proteases, putative molecular chaperone 2'-5' RNA ligase	RNA processing orphans	D23_1c1337	Neut_1299
fig/66666666.60966.peg.1330	CDS	1235245	1234742	-1	-	504	2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase (EC 4.6.1.12)	Isoprenoid Biosynthesis; Nonmevalonate Branch of Isoprenoid Biosynthesis; Possible RNA degradation cluster; Stationary phase repair cluster	D23_1c1338	Neut_1300
fig/66666666.60966.peg.1331	CDS	1235544	1235356	-3	-	189	hypothetical protein	-none-	D23_1c1339	Neut_0314
fig/66666666.60966.peg.1332	CDS	1235706	1235566	-3	-	141	hypothetical protein	-none-	D23_1c1340	Neut_0314
fig/66666666.60966.peg.1335	CDS	1236565	1236008	-1	-	558	Translation elongation factor P	Translation elongation factors bacterial	D23_1c1341	Neut_1302
fig/66666666.60966.peg.1336	CDS	1237794	1236640	-3	-	1155	hypothetical protein	-none-	D23_1c1342	Neut_1303
fig/66666666.60966.peg.1337	CDS	1237927	1238928	1	+	1002	D-lactate dehydrogenase (EC 1.1.1.28)	Fermentations: Lactate	D23_1c1343	Neut_1304
fig/66666666.60966.peg.1338	CDS	1240296	1239037	-3	-	1260	putative membrane protein	-none-	D23_1c1344	Neut_1315
fig/66666666.60966.peg.1339	CDS	1241400	1240354	-3	-	1047	Peptide chain release factor 2; programmed frameshift-containing	Programmed frameshift; Programmed frameshift; Translation termination factors bacterial	D23_1c1345	Neut_1316
fig/66666666.60966.peg.1340	CDS	1243378	1241576	-1	-	1803	patatin family protein	-none-	D23_1c1346	Neut_1317
fig/66666666.60966.peg.1341	CDS	1244736	1243531	-3	-	1206	Catalase (EC 1.11.1.6)	Oxidative stress; Photorespiration (oxidative C2 cycle); Protection from	D23_1c1348	NA
fig/66666666.60966.peg.1342	CDS	1246268	1244739	-2	-	1530	Peroxidase (EC 1.11.1.7)	Reactive Oxygen Species Oxidative stress; Protection from	D23_1c1349	NA
fig/66666666.60966.peg.1343	CDS	1247585	1246281	-2	-	1305	Peroxidase (EC 1.11.1.7)	Reactive Oxygen Species Oxidative stress; Protection from	D23_1c1350	NA
fig/66666666.60966.peg.1344	CDS	1248078	1247623	-3	-	456	hypothetical protein	Reactive Oxygen Species	D23_1c1351	NA
fig/66666666.60966.peg.1345	CDS	1248448	1251117	1	+	2670	FIG00860108; hypothetical protein	 Protection from -none-	D23_1c1352	Neut_0141

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.									
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	C91 Alias
fig 66666666.60966.peg.1346	CDS	1251218	1253155	2	+	1938	Choline dehydrogenase (EC 1.1.1.99.1)	-none-	D23_1c1353 NA
fig 66666666.60966.peg.1347	CDS	1256051	1253184	-2	-	2868	Peroxidase (EC 1.11.1.8)	-none-	D23_1c1354 NA
fig 66666666.60966.peg.1348	CDS	1257505	1256081	-1	-	1425	Hemagglutinin	-none-	D23_1c1355 NA
fig 66666666.60966.peg.1349	CDS	1258257	1257547	-3	-	711	hypothetical protein	-none-	D23_1c1356 NA
fig 66666666.60966.peg.1350	CDS	1259233	1258262	-1	-	972	hypothetical protein	-none-	D23_1c1357 NA
fig 66666666.60966.peg.1351	CDS	1261078	1259246	-1	-	1833	hypothetical protein	-none-	D23_1c1358 NA
fig 66666666.60966.peg.1352	CDS	1262776	1261109	-1	-	1668	Arachidonate 15-lipoxygenase (EC 1.13.11.33)	-none-	D23_1c1359 NA
fig 66666666.60966.peg.1353	CDS	1264449	1262845	-3	-	1605	putative cyclooxygenase-2	-none-	D23_1c1360 NA
fig 66666666.60966.peg.1354	CDS	1266134	1264638	-2	-	1497	hypothetical protein	-none-	D23_1c1361 NA
fig 66666666.60966.peg.1355	CDS	1266382	1266167	-1	-	216	hypothetical protein	-none-	D23_1c1362 NA
fig 66666666.60966.peg.1356	CDS	1266704	1266444	-2	-	261	hypothetical protein	-none-	D23_1c1363 NA
fig 66666666.60966.peg.1357	CDS	1267804	1266983	-1	-	822	Kazal-type serine protease inhibitor domain	-none-	D23_1c1364 NA
fig 66666666.60966.peg.1358	CDS	1268650	1267970	-1	-	681	Kazal-type serine protease inhibitor domain	-none-	D23_1c1365 NA
fig 66666666.60966.peg.1359	CDS	1268913	1268800	-3	-	114	hypothetical protein	-none-	D23_1c1366 NA
fig 66666666.60966.peg.1360	CDS	1269830	1269126	-2	-	705	Mobile element protein	-none-	D23_1c1367 Neut_1318
fig 66666666.60966.peg.1361	CDS	1270083	1269853	-3	-	231	Mobile element protein	-none-	D23_1c1368 Neut_1318
fig 66666666.60966.peg.1362	CDS	1270317	1270111	-3	-	207	Mobile element protein	-none-	D23_1c1369 Neut_2405
fig 66666666.60966.peg.1363	CDS	1271332	1270409	-1	-	924	alpha/beta hydrolase fold	-none-	D23_1c1370 NA
fig 66666666.60966.peg.1364	CDS	1271721	1271329	-3	-	393	hypothetical protein	-none-	D23_1c1371 NA
fig 66666666.60966.peg.1365	CDS	1272160	1273251	1	+	1092	Putrescine transport ATP-binding protein PotA (TC 3.A.1.11.1)	Polyamine Metabolism	D23_1c1372 Neut_1328
fig 66666666.60966.peg.1366	CDS	1273248	1274150	3	+	903	Spermidine Putrescine ABC transporter permease component	Polyamine Metabolism	D23_1c1373 Neut_1329
fig 66666666.60966.peg.1367	CDS	1274170	1274955	1	+	786	PotB (TC 3.A.1.11.1) Spermidine Putrescine ABC transporter permease component	Polyamine Metabolism	D23_1c1374 Neut_1330
fig 66666666.60966.peg.1368	CDS	1274952	1276061	3	+	1110	potC (TC_3.A.1.11.1) ABC transporter, periplasmic spermidine putrescine-binding protein PotD (TC 3.A.1.11.1)	Polyamine Metabolism	D23_1c1375 Neut_1331
fig 66666666.60966.peg.1369	CDS	1276066	1276374	1	+	309	Ferredoxin, 2Fe—2S	Alanine biosynthesis; Soluble cytochromes and functionally related	D23_1c1376 Neut_1332

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.									
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld C91 Alias
fig/66666666.60966.peg.1370	CDS	1278030	1276423	-3	-	1608	Type I restriction-modification system, DNA-methyltransferase subunit M (EC 2.1.1.72)	electron carriers; RNA modification Bacteria Restriction-Modification System; Type I Restriction-Modification	D23_1c1377 Neut_0541
fig/66666666.60966.peg.1371	CDS	1279031	1278027	-2	-	1005	Putative DNA-binding protein in cluster with Type I restriction-modification system	Restriction-Modification System	D23_1c1378 NA
fig/66666666.60966.peg.1372	CDS	1282963	1279028	-1	-	3936	Type I restriction-modification system, restriction subunit R (EC 3.1.21.3)	Restriction-Modification System; Type I Restriction-Modification	D23_1c1379 Neut_0537
fig/66666666.60966.peg.1373	CDS	1283272	1282982	-1	-	291	Type I restriction-modification system, restriction subunit R (EC 3.1.21.3)	Restriction-Modification System; Type I Restriction-Modification	D23_1c1380 NA
fig/66666666.60966.peg.1374	CDS	1287342	1283581	-3	-	3762	macromolecule metabolism; macromolecule synthesis, modification; dna-replication, repair, restr./modif.	-none-	D23_1c1381 Neut_1336
fig/66666666.60966.peg.1375	CDS	1287828	1287349	-3	-	480	FIG00858549; hypothetical protein	-none-	D23_1c1382 Neut_1337
fig/66666666.60966.peg.1376	CDS	1288968	1287859	-3	-	1110	2-keto-3-deoxy-D-arabino-heptulosonate-7-phosphate synthase I alpha (EC 2.5.1.54)	Chorismate Synthesis; Common Pathway For Synthesis of Aromatic Compounds (DAHP synthase to chorismate)	D23_1c1383 Neut_1338
fig/66666666.60966.peg.1377	CDS	1290768	1289113	-3	-	1656	MutS domain protein, family 6	DNA repair, bacterial MutL-MutS system	D23_1c1384 Neut_1339
fig/66666666.60966.peg.1378	CDS	1290915	1291493	3	+	579	FIG00858435; hypothetical protein	-none-	D23_1c1385 Neut_1340
fig/66666666.60966.peg.1379	CDS	1291506	1292660	3	+	1155	Probable Co/Zn/Cd efflux system membrane fusion protein	Cobalt-zinc-cadmium resistance	D23_1c1386 Neut_1341
fig/66666666.60966.peg.1380	CDS	1292693	1293328	2	+	636	ABC transporter ATP-binding protein YvcR	-none-	D23_1c1387 Neut_1342
fig/66666666.60966.peg.1381	CDS	1293325	1294527	1	+	1203	ABC transporter permease protein	-none-	D23_1c1388 Neut_1343
fig/66666666.60966.peg.1382	CDS	1294533	1295732	3	+	1200	putative ABC transporter protein	-none-	D23_1c1389 Neut_1344

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld	C91 Alias
fig/66666666.60966.peg.1383	CDS	1296266	1295742	-2	-	525	FIG00859169: hypothetical protein	-none-	D23_1c1390	Neut_1345
fig/66666666.60966.peg.1385	CDS	1296862	1297086	1	+	225	ADA regulatory protein/Methylated-DNA--protein-cysteine methyltransferase (EC 2.1.1.63)	DNA repair, bacterial; DNA repair, bacterial	D23_1c1391	Neut_1346
fig/66666666.60966.peg.1386	CDS	1297089	1297802	3	+	714	hypothetical protein	-none-	D23_1c1392	Neut_1347
fig/66666666.60966.peg.1387	CDS	1298101	1297967	-1	-	135	hypothetical protein	-none-	D23_1c1393	NA
fig/66666666.60966.peg.1389	CDS	1298212	1298838	1	+	627	Alkylated DNA repair protein	-none-	D23_1c1395	Neut_1349
fig/66666666.60966.peg.1390	CDS	1300170	1298923	-3	-	1248	Mobile element protein	-none-	D23_1c1396	Neut_0357
fig/66666666.60966.peg.1391	CDS	1300681	1300355	-1	-	327	hypothetical protein	-none-	D23_1c1397	Neut_1350
fig/66666666.60966.peg.1392	CDS	1301157	1300963	-3	-	195	Glutaredoxin	Glutaredoxins; Glutathione: Redox cycle; Phage DNA synthesis	D23_1c1398	Neut_1351
fig/66666666.60966.peg.1393	CDS	1302170	1301697	-2	-	474	Mobile element protein	-none-	D23_1c1399	Neut_1353
fig/66666666.60966.peg.1394	CDS	1303008	1302400	-3	-	609	Glutathione S-transferase (EC 2.5.1.18)	Glutathione: Non-redox reactions	D23_1c1400	Neut_1354
fig/66666666.60966.peg.1395	CDS	1303729	1303556	-1	-	174	hypothetical protein	-none-	D23_1c1402	NA
fig/66666666.60966.peg.1396	CDS	1303948	1307514	1	+	3567	Exodeoxyribonuclease V gamma chain (EC 3.1.11.5)	DNA repair, bacterial RecBCD pathway	D23_1c1403	Neut_1355
fig/66666666.60966.peg.1397	CDS	1307530	1311216	1	+	3687	Exodeoxyribonuclease V beta chain (EC 3.1.11.5)	DNA repair, bacterial RecBCD pathway	D23_1c1404	Neut_1356
fig/66666666.60966.peg.1398	CDS	1311213	1313258	3	+	2046	Exodeoxyribonuclease V alpha chain (EC 3.1.11.5)	DNA repair, bacterial RecBCD pathway	D23_1c1405	Neut_1357
fig/66666666.60966.peg.1399	CDS	1314581	1313277	-2	-	1305	Aspartyl aminopeptidase	-none-	D23_1c1406	Neut_1358
fig/66666666.60966.peg.1400	CDS	1315032	1314604	-3	-	429	PIN domain family protein	-none-	D23_1c1407	Neut_1359
fig/66666666.60966.peg.1401	CDS	1315334	1315032	-2	-	303	DNA-binding protein, CopG family	-none-	D23_1c1408	Neut_1360
fig/66666666.60966.peg.1402	CDS	1316699	1315362	-2	-	1338	Sensor protein PhoQ (EC 2.7.13.3)	-none-	D23_1c1409	Neut_1361
fig/66666666.60966.peg.1403	CDS	1317382	1316696	-1	-	687	DNA-binding response regulator	-none-	D23_1c1410	Neut_1362
fig/66666666.60966.peg.1404	CDS	1317741	1317451	-3	-	291	hypothetical protein	-none-	D23_1c1411	Neut_1363
fig/66666666.60966.peg.1405	CDS	1318192	1317827	-1	-	366	Putative metal chaperone, involved in Zn homeostasis, GTPase of COG0523 family	G3E family of P-loop GTPases (metalcenter biosynthesis); Zinc regulated enzymes	D23_1c1412	NA
fig/66666666.60966.peg.1406	CDS	1318485	1319036	3	+	552	Protein of unknown function DUF924	-none-	D23_1c1413	Neut_1365

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld	C91 Alias
fig/66666666.60966.peg.1407	CDS	1320160	1319171	-1	-	990	Integron integrase IntIPac	Integrans	D23_1c1414	Neut_1366
fig/66666666.60966.peg.1408	CDS	1320314	1321651	2	+	1338	DNA modification methyltransferase	-none-	D23_1c1415	NA
fig/66666666.60966.peg.1409	CDS	1321654	1322532	1	+	879	hypothetical protein	-none-	D23_1c1416	NA
fig/66666666.60966.peg.1410	CDS	1322665	1322880	1	+	216	hypothetical protein	-none-	D23_1c1417	NA
fig/66666666.60966.peg.1411	CDS	1323474	1324154	3	+	681	Thi/Pfp family protein	-none-	D23_1c1418	NA
fig/66666666.60966.peg.1412	CDS	1324808	1324990	2	+	183	hypothetical protein	-none-	D23_1c1419	NA
fig/66666666.60966.peg.1413	CDS	1325941	1324979	-1	-	963	Mobile element protein	-none-	D23_1c1420	Neut_1746
fig/66666666.60966.peg.1415	CDS	1326558	1328531	3	+	1974	Monoamine oxidase (1,4,3,4)	Auxin biosynthesis; Glycine and Serine Utilization; Threonine degradation	D23_1c1421	NA
fig/66666666.60966.peg.1417	CDS	1328848	1328711	-1	-	138	Mobile element protein	-none-	D23_1c1422	Neut_2501
fig/66666666.60966.peg.1418	CDS	1329144	1328821	-3	-	324	Mobile element protein	-none-	D23_1c1423	Neut_1624
fig/66666666.60966.peg.1419	CDS	1329566	1329129	-2	-	438	Mobile element protein	-none-	D23_1c1424	Neut_1888
fig/66666666.60966.peg.1420	CDS	1330200	1329892	-3	-	309	Death on curing protein, Doc toxin	Phd-Doc, YdcE-YdcD toxin-antitoxin (programmed cell death) systems	D23_1c1425	NA
fig/66666666.60966.peg.1421	CDS	1330677	1330231	-3	-	447	Prevent host death protein, Phd antitoxin	Phd-Doc, YdcE-YdcD toxin-antitoxin (programmed cell death) systems	D23_1c1426	NA
fig/66666666.60966.peg.1423	CDS	1331253	1331387	3	+	135	Mobile element protein	-none-	D23_1c1427	Neut_2500
fig/66666666.60966.peg.1424	CDS	1331444	1332292	2	+	849	Mobile element protein	-none-	D23_1c1428	Neut_1888
fig/66666666.60966.peg.1425	CDS	1332744	1333214	3	+	471	3-demethylubiquinone-9 3-methyltransferase	-none-	D23_1c1429	Neut_1376
fig/66666666.60966.peg.1426	CDS	1333410	1335026	3	+	1617	Dihydroxyacetone kinase, ATP-dependent (EC 2.7.1.29)	Dihydroxyacetone kinases	D23_1c1431	Neut_1377
fig/66666666.60966.peg.1427	CDS	1335644	1335210	-2	-	435	hypothetical protein	-none-	D23_1c1432	Neut_1378
fig/66666666.60966.peg.1428	CDS	1339065	1335670	-3	-	3396	PDZ domain	-none-	D23_1c1433	Neut_1379
fig/66666666.60966.peg.1429	CDS	1339241	1340350	2	+	1110	COGs COG0823	-none-	D23_1c1435	Neut_1380
fig/66666666.60966.peg.1430	CDS	1341715	1340438	-1	-	1278	Purative diHEME cytochrome c-553	Soluble cytochromes and functionally related electron carriers	D23_1c1436	Neut_1381
fig/66666666.60966.peg.1431	CDS	1342431	1341712	-3	-	720	Probable cytochrome c2	Soluble cytochromes and functionally related electron carriers	D23_1c1437	Neut_1382
fig/66666666.60966.peg.1432	CDS	1342445	1342963	2	+	519	FIG00859968: hypothetical protein	-none-	D23_1c1438	Neut_1383
fig/66666666.60966.peg.1433	CDS	1344805	1342952	-1	-	1854	Cell division protein FtsH (EC 3.4.24.—)	Bacterial Cell Division	D23_1c1439	Neut_1384
fig/66666666.60966.peg.1434	CDS	1345012	1346118	1	+	1107	S-(hydroxymethyl)glutathione	Glutathione-dependent pathway of	D23_1c1440	Neut_1385

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld	C91 Alias
fig/66666666.60966.peg.1435	CDS	1346131	1347000	1	+	870	dehydrogenase (EC 1.1.1.284)	formaldehyde detoxification	D23_1c1441	Neut_1386
							S-formylglutathione hydrolase (EC 3.1.2.12)	Glutathione-dependent pathway of formaldehyde detoxification		
fig/66666666.60966.peg.1436	CDS	1348436	1347264	-2	-	1173	Ornithine decarboxylase (EC 4.1.1.17)	Arginine and Ornithine Degradation;	D23_1c1442	Neut_1387
							Arginine decarboxylase (EC 4.1.1.19)	 Arginine and Ornithine Degradation;		
fig/66666666.60966.peg.1437	CDS	1348961	1349230	2	+	270	FIG00858492;	Metabolism	D23_1c1443	Neut_1388
							hypothetical protein	-none-		
fig/66666666.60966.peg.1438	CDS	1349336	1349214	-2	-	123	hypothetical protein	-none-	D23_1c1444	NA
fig/66666666.60966.peg.1439	CDS	1349660	1350622	2	+	963	Mobile element protein	-none-	D23_1c1446	Neut_1862
fig/66666666.60966.peg.1440	CDS	1350742	1352079	1	+	1338	Ribosomal protein S12p Asp88 (<i>E. coli</i>)	Methylthiotransferases;	D23_1c1447	Neut_1389
fig/66666666.60966.peg.1441	CDS	1353046	1352255	-1	-	792	methylthiotransferase	S12p Asp	D23_1c1448	Neut_1390
							hypothetical protein	methylthiotransferase		
fig/66666666.60966.peg.1442	CDS	1353301	1353618	1	+	318	Cell division protein	-none-	D23_1c1449	Neut_1391
fig/66666666.60966.peg.1443	CDS	1353569	1354282	2	+	714	LemA family protein	Bacterial Cell Division	D23_1c1450	Neut_1392
							Beta-propeller domains of methanol	-none-		
fig/66666666.60966.peg.1444	CDS	1354309	1355193	1	+	885		-none-	D23_1c1451	Neut_1393
fig/66666666.60966.peg.1445	CDS	1355207	1355710	2	+	504	dehydrogenase type	-none-	D23_1c1452	Neut_1394
							FIG004694;	-none-		
fig/66666666.60966.peg.1446	CDS	1356010	1356513	1	+	504	Hypothetical protein		D23_1c1454	Neut_1395
							DNA-directed RNA polymerase specialized sigma subunit, sigma24-like	-none-		
fig/66666666.60966.peg.1447	CDS	1356510	1357232	3	+	723	FIG00859011;	-none-	D23_1c1455	Neut_1396
fig/66666666.60966.peg.1448	CDS	1357326	1359644	3	+	2319	hypothetical protein	-none-	D23_1c1456	Neut_1397
							ABC transporter, transmembrane region: ABC transporter related	-none-		
fig/66666666.60966.peg.1449	CDS	1359641	1360108	2	+	468	DUF1854 domain-containing protein	-none-	D23_1c1457	Neut_1398
							hypothetical protein	-none-		
fig/66666666.60966.peg.1451	CDS	1360927	1362021	1	+	1095	hypothetical protein	-none-	D23_1c1458	Neut_1860
fig/66666666.60966.peg.1452	CDS	1362049	1362342	1	+	294	Mobile element protein	-none-	D23_1c1459	Neut_1719
fig/66666666.60966.peg.1453	CDS	1362441	1363319	3	+	879	Mobile element protein	-none-	D23_1c1460	Neut_1720
fig/66666666.60966.peg.1454	CDS	1363397	1365250	2	+	1854	hypothetical protein	-none-	D23_1c1461	NA
fig/66666666.60966.peg.1455	CDS	1365453	1365262	-3	-	192	Mobile element protein	-none-	D23_1c1462	Neut_2502

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.									
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	C91 Alias
fig/66666666.60966.peg.1456	CDS	1365648	1367945	3	+	2298	Cyanophycin synthase (EC 6.3.2.29)(EC 6.3.2.30)	Cyanophycin Metabolism	Neut_1401
fig/66666666.60966.peg.1457	CDS	1367966	1370572	2	+	2607	Cyanophycin synthase (EC 6.3.2.29)(EC 6.3.2.30)	Cyanophycin Metabolism	Neut_1402
fig/66666666.60966.peg.1458	CDS	1371664	1370720	-1	-	945	Copper-containing nitrite reductase (EC 1.7.2.1)	Denitrification; Dentrifying reductase gene clusters	Neut_1403
fig/66666666.60966.peg.1459	CDS	1372091	1371708	-2	-	384	cytochrome c, class IC	-none-	Neut_1404
fig/66666666.60966.peg.1460	CDS	1372789	1372091	-1	-	699	Cytochrome c, class I	-none-	Neut_1405
fig/66666666.60966.peg.1461	CDS	1373888	1372827	-2	-	1062	Multicopper oxidase	Copper homeostasis	Neut_1406
fig/66666666.60966.peg.1462	CDS	1374021	1374137	3	+	117	hypothetical protein	-none-	NA
fig/66666666.60966.peg.1463	CDS	1374189	1374653	3	+	465	Nitrite-sensitive transcriptional repressor NsrR	Nitrosative stress; Oxidative stress; Rrf2 family transcriptional regulators	Neut_1407
fig/66666666.60966.peg.1464	CDS	1378537	1374641	-1	-	3897	FIG00858660: hypothetical protein	-none-	Neut_1408
fig/66666666.60966.peg.1465	CDS	1380248	1378539	-2	-	1710	Outer membrane protein	-none-	Neut_1409
fig/66666666.60966.peg.1466	CDS	1380471	1380340	-3	-	132	hypothetical protein	-none-	NA
fig/66666666.60966.peg.1467	CDS	1380491	1381141	2	+	651	Uracil-DNA glycosylase, family 5	Uracil-DNA glycosylase	Neut_1410
fig/66666666.60966.peg.1468	CDS	1381162	1381350	1	+	189	putative isomerase	-none-	Neut_1411
fig/66666666.60966.peg.1469	CDS	1381364	1383178	2	+	1815	Exonuclease ABC subunit C	DNA repair, UvrABC system	Neut_1412
fig/66666666.60966.peg.1470	CDS	1383321	1384283	3	+	963	Mobile element protein	-none-	Neut_1746
fig/66666666.60966.peg.1471	CDS	1384401	1384814	3	+	414	ABC-type multidrug transport system, permease component	CBSS-196164.1.peg.1690	NA
fig/66666666.60966.peg.1472	CDS	1384811	1385149	2	+	339	ABC-type multidrug transport system, permease component	CBSS-196164.1.peg.1690	NA
fig/66666666.60966.peg.1474	CDS	1385894	1388275	2	+	2382	Penicillin acylase II	-none-	Neut_1415
fig/66666666.60966.peg.1475	CDS	1388559	1388395	-3	-	165	hypothetical protein	-none-	NA
fig/66666666.60966.peg.1476	CDS	1388615	1389973	2	+	1359	Response regulatory protein	-none-	Neut_1416
fig/66666666.60966.peg.1477	CDS	1390014	1391603	3	+	1590	Long-chain-fatty-acid-CoA ligase (EC 6.2.1.3)	Biotin biosynthesis; Biotin synthesis cluster; Fatty acid metabolism cluster	Neut_1417
fig/66666666.60966.peg.1478	CDS	1391630	1392835	2	+	1206	Diaminopimelate decarboxylase (EC 4.1.1.20)	Lysine Biosynthesis DAP Pathway, GJO scratch	Neut_1418

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld	C91 Alias
fig/66666666.60966.peg.1479	CDS	1392996	1394843	3	+	1848	Asparagine synthetase [glutamine-hydrolyzing] (EC 6.3.5.4)	Cyanophycin Metabolism; Glutamate and Aspartate uptake in Bacteria; Glutamine, Glutamate, Aspartate and Asparagine Biosynthesis	D23_1c1485	Neut_1419
fig/66666666.60966.peg.1480	CDS	1394795	1394911	2	+	117	hypothetical protein	-none-	D23_1c1486	NA
fig/66666666.60966.peg.1482	CDS	1395163	1395975	1	+	813	FIG00858746; hypothetical protein	-none-	D23_1c1488	Neut_1420
fig/66666666.60966.peg.1483	CDS	1397102	1396140	-2	-	963	Mobile element protein	-none-	D23_1c1490	Neut_1746
fig/66666666.60966.peg.1484	CDS	1397178	1397462	3	+	285	COGs COG0226	-none-	D23_1c1491	Neut_1423
fig/66666666.60966.peg.1485	CDS	1397455	1398681	1	+	1227	FIG00859800; hypothetical protein	-none-	D23_1c1492	Neut_1424
fig/66666666.60966.peg.1486	CDS	1398693	1401635	3	+	2943	hypothetical protein diguanylate cyclase/phosphodiesterase (GGDEF & EAL domains) with PAS/PAC sensor(s)	-none-	D23_1c1493	Neut_1425
fig/66666666.60966.peg.1488	CDS	1402082	1401924	-2	-	159	hypothetical protein	-none-	D23_1c1494	NA
fig/66666666.60966.peg.1489	CDS	1402531	1402115	-1	-	417	OsmC/Ohr family protein	-none-	D23_1c1495	Neut_1426
fig/66666666.60966.peg.1490	CDS	1403584	1402532	-1	-	1053	DNA polymerase III delta subunit (EC 2.7.7.7)	CBSS-208964.1.peg.3988	D23_1c1496	Neut_1427
fig/66666666.60966.peg.1491	CDS	1404106	1403612	-1	-	495	LPS-assembly lipoprotein RlpB precursor (Rare lipoprotein B)	CBSS-208964.1.peg.3988; Lipopolysaccharide assembly	D23_1c1497	Neut_1428
fig/66666666.60966.peg.1492	CDS	1406732	1404126	-2	-	2607	Leucyl-tRNA synthetase (EC 6.1.1.4)	CBSS-208964.1.peg.3988; tRNA aminoacylation, Leu	D23_1c1498	Neut_1429
fig/66666666.60966.peg.1493	CDS	1406756	1407925	2	+	1170	S-adenosylmethionine:RNA ribosyltransferase-isomerase (EC 5.---.---.---)	CBSS-211586.1.peg.2832; Queuosine-Archaeosine Biosynthesis; Scaffold proteins for [4Fe-4S] cluster assembly (MRP family); tRNA modification	D23_1c1499	Neut_1430
fig/66666666.60966.peg.1494	CDS	1407922	1409007	1	+	1086	tRNA-guanine transglycosylase (EC 2.4.2.29)	Bacteria CBSS-211586.1.peg.2832; Queuosine-Archaeosine	D23_1c1500	Neut_1431

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.									
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	C91 Alias
fig/66666666.60966.peg.1495	CDS	1409072	1409536	2	+	465	Preprotein translocase subunit YajC (TC 3.A.5.1.1)	Biosynthesis; Scaffold proteins for [4Fe-4S] cluster assembly (MRP family); RNA modification	Neut_1432
fig/66666666.60966.peg.1496	CDS	1409575	1411470	1	+	1896	Protein-export membrane protein SecD (TC 3.A.5.1.1)	CBSS-211586.1.peg.2832	Neut_1433
fig/66666666.60966.peg.1497	CDS	1411495	1412427	1	+	933	Protein-export membrane protein SecF (TC 3.A.5.1.1)	CBSS-211586.1.peg.2832	Neut_1434
fig/66666666.60966.peg.1498	CDS	1412467	1412844	1	+	378	FIG028220; hypothetical protein co-occurring with HEAT repeat protein	-none-	Neut_1435
fig/66666666.60966.peg.1499	CDS	1412897	1413628	2	+	732	Ubiquinone/menaquinone biosynthesis methyltransferase UbiE (EC 2.1.1.—)	Menaquinone and Phyloquinone Biosynthesis; Ubiquinone Biosynthesis; Ubiquinone	Neut_1436
fig/66666666.60966.peg.1500	CDS	1413802	1413680	-1	-	123	Agmatinase (EC 3.5.3.11)	Biosynthesis-gio Arginine and Ornithine Degradation; Polyamine Metabolism	Neut_1437
fig/66666666.60966.peg.1501	CDS	1414650	1413808	-3	-	843	Agmatinase (EC 3.5.3.11)	Arginine and Ornithine Degradation; Polyamine Metabolism	Neut_1437
fig/66666666.60966.peg.1502	CDS	1415255	1414815	-2	-	441	Lipoprotein signal peptidase (EC 3.4.23.36)	Lipoprotein Biosynthesis; Signal peptidase	Neut_1438
fig/66666666.60966.peg.1503	CDS	1415346	1415224	-3	-	123	hypothetical protein	-none-	NA
fig/66666666.60966.peg.1504	CDS	1418173	1415339	-1	-	2835	Isoleucyl-tRNA synthetase (EC 6.1.1.5)	tRNA aminoacylation, Ile	Neut_1439
fig/66666666.60966.peg.1505	CDS	1418945	1418148	-2	-	798	Riboflavin kinase (EC 2.7.1.26)/FMN adenylyltransferase (EC 2.7.7.2)	Riboflavin, FMN and FAD metabolism; Riboflavin, FMN and FAD metabolism; Riboflavin, FMN and FAD metabolism in plants; Riboflavin, FMN and FAD metabolism in plants;	Neut_1440

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 GbkId	C91 Alias
fig/66666666.60966.peg.1506	CDS	1419539	1419237	-2	-	303	possible sec-independent protein translocase protein	 riboflavin to FAD; riboflavin to FAD -none-	D23_1c1513	Neut_1441
fig/66666666.60966.peg.1507	CDS	1420011	1419886	-3	-	126	TatC hypothetical protein	-none-	D23_1c1514	Neut_1442
fig/66666666.60966.peg.1508	CDS	1420112	1422079	2	+	1968	Dipeptide-binding ABC transporter, periplasmic substrate-binding component (TC 3.A.1.5.2)	ABC transporter dipeptide (TC 3.A.1.5.2)	D23_1c1514	Neut_1442
fig/66666666.60966.peg.1509	CDS	1422089	1423066	2	+	978	3.A.1.5.2 Oligopeptide transport system permease protein OppB (TC 3.A.1.5.1)	ABC transporter oligopeptide (TC 3.A.1.5.1)	D23_1c1515	Neut_1443
fig/66666666.60966.peg.1511	CDS	1423196	1424851	2	+	1656	DnaJ-class molecular chaperone CbpA	Protein chaperones	D23_1c1516	Neut_1444
fig/66666666.60966.peg.1512	CDS	1424900	1425844	2	+	945	DnaJ-class molecular chaperone CbpA	Protein chaperones	D23_1c1517	Neut_1445
fig/66666666.60966.peg.1513	CDS	1425856	1426152	1	+	297	InterPro IPR000551	-none-	D23_1c1518	Neut_1446
fig/66666666.60966.peg.1514	CDS	1426240	1427103	1	+	864	FIG00858431: hypothetical protein	-none-	D23_1c1519	Neut_1447
fig/66666666.60966.peg.1515	CDS	1427120	1427659	2	+	540	FAD pyrophosphatase (EC 3.6.1.18), projected from PMID:18815383	-none-	D23_1c1520	Neut_1448
fig/66666666.60966.peg.1516	CDS	1428946	1428491	-1	-	456	Mobile element protein	-none-	D23_1c1522	Neut_2502
fig/66666666.60966.peg.1517	CDS	1429295	1428909	-2	-	387	Mobile element protein	-none-	D23_1c1523	Neut_0884
fig/66666666.60966.peg.1518	CDS	1430487	1429300	-3	-	1188	Mobile element protein	-none-	D23_1c1524	Neut_2405
fig/66666666.60966.peg.1519	CDS	1430505	1430621	3	+	117	paratin family protein	-none-	D23_1c1525	Neut_1317
fig/66666666.60966.peg.1520	CDS	1430834	1430670	-2	-	165	hypothetical protein	-none-	D23_1c1526	NA
fig/66666666.60966.peg.1521	CDS	1431909	1431424	-3	-	486	hypothetical protein	-none-	D23_1c1527	Neut_1449
fig/66666666.60966.peg.1522	CDS	1432156	1431899	-1	-	258	hypothetical protein	-none-	D23_1c1528	Neut_1450
fig/66666666.60966.peg.1523	CDS	1432650	1432159	-3	-	492	phage-related hypothetical protein	-none-	D23_1c1529	Neut_1451
fig/66666666.60966.peg.1524	CDS	1432970	1432650	-2	-	321	hypothetical protein	-none-	D23_1c1530	Neut_1452
fig/66666666.60966.peg.1525	CDS	1433143	1432967	-1	-	177	hypothetical protein	-none-	D23_1c1531	Neut_1453
fig/66666666.60966.peg.1526	CDS	1434487	1433552	-1	-	936	hypothetical protein	-none-	D23_1c1533	Neut_1454
fig/66666666.60966.peg.1527	CDS	1437226	1434497	-1	-	2730	hypothetical protein	-none-	D23_1c1534	Neut_1455
fig/66666666.60966.peg.1528	CDS	1437449	1437267	-2	-	183	Phage protein	-none-	D23_1c1535	Neut_1456
fig/66666666.60966.peg.1529	CDS	1437694	1437467	-1	-	228	hypothetical protein	-none-	D23_1c1536	Neut_1457
fig/66666666.60966.peg.1530	CDS	1438472	1437702	-2	-	771	hypothetical protein	-none-	D23_1c1537	Neut_1458
fig/66666666.60966.peg.1531	CDS	1439638	1438469	-1	-	1170	hypothetical protein	-none-	D23_1c1538	Neut_1459
fig/66666666.60966.peg.1532	CDS	1441511	1439631	-2	-	1881	Phage tail length tape- measure protein	Phage tail proteins; Phage tail proteins 2	D23_1c1539	Neut_1460
fig/66666666.60966.peg.1533	CDS	1442755	1441508	-1	-	1248	Mobile element protein	-none-	D23_1c1540	Neut_0357
fig/66666666.60966.peg.1534	CDS	1443390	1442854	-3	-	537	Phage protein	-none-	D23_1c1541	Neut_1460
fig/66666666.60966.peg.1535	CDS	1443638	1443390	-2	-	249	hypothetical protein	-none-	D23_1c1542	Neut_1461

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.									
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	C91 Alias
fig/66666666.60966.peg.1536	CDS	1444027	1443677	-1	-	351	Phage protein	-non-	Neut_1462
fig/66666666.60966.peg.1537	CDS	1444752	1444036	-3	-	717	major tail protein, putative	-non-	Neut_1463
fig/66666666.60966.peg.1538	CDS	1445120	1444758	-2	-	363	hypothetical protein	-non-	Neut_1464
fig/66666666.60966.peg.1539	CDS	1445638	1445117	-1	-	522	FIG00959132: hypothetical protein	-non-	Neut_1465
fig/66666666.60966.peg.1540	CDS	1445985	1445650	-3	-	336	Phage protein	-non-	Neut_1466
fig/66666666.60966.peg.1541	CDS	1446544	1445987	-1	-	558	Similar to Gene Transfer Agent (GTA) ORFG06	-non-	Neut_1467
fig/66666666.60966.peg.1542	CDS	1446900	1446601	-3	-	300	hypothetical protein	-non-	Neut_1468
fig/66666666.60966.peg.1543	CDS	1448136	1446910	-3	-	1227	Phage major capsid protein	Phage capsid proteins	Neut_1469
fig/66666666.60966.peg.1544	CDS	1448949	1448218	-3	-	732	Prophage Ctp protease-like protein	cAMP signaling in bacteria	Neut_1470
fig/66666666.60966.peg.1545	CDS	1450207	1448915	-1	-	1293	Phage portal protein	Phage packaging machinery	Neut_1471
fig/66666666.60966.peg.1546	CDS	1451877	1450204	-3	-	1674	Phage terminase large subunit	Phage packaging machinery	Neut_1472
fig/66666666.60966.peg.1547	CDS	1452346	1451882	-1	-	465	Phage terminase, small subunit	Phage packaging machinery	Neut_1473
fig/66666666.60966.peg.1548	CDS	1452801	1452478	-3	-	324	Phage holin	-non-	Neut_1474
fig/66666666.60966.peg.1549	CDS	1453267	1452887	-1	-	381	hypothetical protein	-non-	Neut_1475
fig/66666666.60966.peg.1550	CDS	1453430	1453269	-2	-	162	hypothetical protein	-non-	NA
fig/66666666.60966.peg.1551	CDS	1453630	1453451	-1	-	180	hypothetical protein	-non-	NA
fig/66666666.60966.peg.1552	CDS	1453862	1453623	-2	-	240	hypothetical protein	-non-	Neut_1477
fig/66666666.60966.peg.1553	CDS	1456504	1454171	-1	-	2334	DNA primase, phage associated # P4-type	-non-	Neut_1478
fig/66666666.60966.peg.1554	CDS	1456740	1456501	-3	-	240	hypothetical protein	-non-	NA
fig/66666666.60966.peg.1555	CDS	1456744	1456887	1	+	144	Phage-related protein	-non-	Neut_1480
fig/66666666.60966.peg.1556	CDS	1456894	1457172	1	+	279	Helix-turn-helix motif	-non-	Neut_1481
fig/66666666.60966.peg.1557	CDS	1457678	1457277	-2	-	402	hypothetical protein	-non-	NA
fig/66666666.60966.peg.1558	CDS	1458137	1458478	2	+	342	hypothetical protein	-non-	NA
fig/66666666.60966.peg.1559	CDS	1458868	1459158	1	+	291	hypothetical protein	-non-	NA
fig/66666666.60966.peg.1561	CDS	1459812	1459672	-3	-	141	hypothetical protein	-non-	NA
fig/66666666.60966.peg.1563	CDS	1460270	1460563	2	+	294	Mobile element protein	-non-	Neut_1719
fig/66666666.60966.peg.1564	CDS	1460662	1461540	1	+	879	Mobile element protein	-non-	Neut_1720
fig/66666666.60966.peg.1565	CDS	1462142	1461954	-2	-	189	hypothetical protein	-non-	Neut_1489
fig/66666666.60966.peg.1567	CDS	1462722	1462844	3	+	123	hypothetical protein	-non-	NA
fig/66666666.60966.peg.1568	CDS	1463137	1463829	1	+	693	putative nuclease	-non-	Neut_1491
fig/66666666.60966.peg.1569	CDS	1464806	1463826	-2	-	981	Abortive infection bacteriophage resistance protein	-non-	NA
fig/66666666.60966.peg.1570	CDS	1466353	1465106	-1	-	1248	Mobile element protein	-non-	Neut_0357
fig/66666666.60966.peg.1573	CDS	1466953	1467108	1	+	156	hypothetical protein	-non-	NA
fig/66666666.60966.peg.1574	CDS	1467105	1467389	3	+	285	hypothetical protein	-non-	Neut_1493
fig/66666666.60966.peg.1575	CDS	1467386	1467868	2	+	483	hypothetical protein	-non-	Neut_1494
fig/66666666.60966.peg.1576	CDS	1467883	1468245	1	+	363	hypothetical protein	-non-	Neut_1495
fig/66666666.60966.peg.1577	CDS	1468255	1468638	1	+	384	hypothetical protein	-non-	Neut_1496

Annotation of genes in SEQ ID NO: 1.

Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 GbkId	C91 Alias
fig/66666666.60966.peg.1580	CDS	1469337	1470362	3	+	1026	Integrase	-none-	D23_1c1580	Neut_1498
fig/66666666.60966.peg.1581	CDS	1470687	1470989	3	+	303	Exodeoxyribonuclease VII small subunit (EC 3.1.11.6)	DNA repair, bacterial; Purine salvage cluster	D23_1c1582	Neut_1499
fig/66666666.60966.peg.1582	CDS	1470979	1471872	1	+	894	Octaprenyl diphosphate synthase (EC 2.5.1.90) Dimethylallyltransferase (EC 2.5.1.1)/(2E,6E)-farnesyl diphosphate synthase (EC 2.5.1.10) Geranylgeranyl diphosphate synthase (EC 2.5.1.29)	Isoprenoid Biosynthesis; Isoprenoid Biosynthesis; Isoprenoid Biosynthesis; Interconversions; Isoprenoids for Quinones; Isoprenoids for Quinones; Isoprenoids for Quinones; Isoprenoids for Quinones; Isoprenoids for Quinones; Polyprenyl Diphosphate Biosynthesis; Polyprenyl Diphosphate Biosynthesis; Polyprenyl Diphosphate Biosynthesis	D23_1c1583	Neut_1500
fig/66666666.60966.peg.1583	CDS	1471935	1473779	3	+	1845	1-deoxy-D-xylose 5-phosphate synthase (EC 2.2.1.7)	Biosynthesis Isoprenoid Biosynthesis; Nonmevalonate Branch of Isoprenoid Biosynthesis; Pyridoxin (Vitamin B6) Biosynthesis; Thiamin biosynthesis	D23_1c1584	Neut_1501
fig/66666666.60966.peg.1584	CDS	1473895	1474698	1	+	804	GTP cyclohydrolase I (EC 3.5.4.16) type 2	Folate Biosynthesis; Queuosine-Archaeosine Biosynthesis; Zinc regulated enzymes -none-	D23_1c1585	Neut_1502
fig/66666666.60966.peg.1585	CDS	1474865	1475485	2	+	621	InterPro IPR005134 COGs COG2862		D23_1c1586	Neut_1503
fig/66666666.60966.peg.1586	CDS	1476254	1475514	-2	-	741	FolM Alternative dihydrofolate reductase 1	Folate Biosynthesis	D23_1c1587	Neut_1504
fig/66666666.60966.peg.1587	CDS	1476327	1477502	3	+	1176	COG1565: Uncharacterized conserved protein	-none-	D23_1c1588	Neut_1505

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld	C91 Alias
fig/66666666.60966.peg.1588	CDS	1478242	1477499	-1	-	744	5'-methylthioadenosine nucleosidase (EC 3.2.2.16)/S-adenosylhomocysteine nucleosidase (EC 3.2.2.9)	Adenosyl nucleosidases; Adenosyl nucleosidases; CBSS-320388.3.peg.3759; CBSS-320388.3.peg.3759; Methionine Biosynthesis; Methionine Degradation; Polyamine Metabolism	D23_1c1589	Neut_1506
fig/66666666.60966.peg.1589	CDS	1480269	1478239	-3	-	2031	Squalene-hopene cyclase (EC 5.4.99.17)	CBSS-320388.3.peg.3759; Hopanes	D23_1c1590	Neut_1507
fig/66666666.60966.peg.1590	CDS	1481103	1480384	-3	-	720	Surface lipoprotein	-none-	D23_1c1591	Neut_1508
fig/66666666.60966.peg.1591	CDS	1481997	1481389	-3	-	609	Ferric siderophore transport system, biopolymer transport protein ExbB	Ton and Tol transport systems	D23_1c1592	Neut_1509
fig/66666666.60966.peg.1592	CDS	1482298	1483644	1	+	1347	Exodeoxyribonuclease VII large subunit (EC 3.1.11.6)	DNA repair, bacterial; Purine salvage cluster	D23_1c1594	Neut_1510
fig/66666666.60966.peg.1593	CDS	1483709	1484089	2	+	381	COG2363	-none-	D23_1c1595	Neut_1511
fig/66666666.60966.peg.1594	CDS	1484150	1485328	2	+	1179	23S rRNA (guanine-N2-)-methyltransferase rhlM EC 2.1.1.-)	RNA methylation	D23_1c1596	Neut_1512
fig/66666666.60966.peg.1595	CDS	1485345	1485833	3	+	489	Periplasmic thiol:disulfide oxidoreductase DsbB, required for DsbA reoxidation	Biogenesis of c-type cytochromes; Periplasmic disulfide interchange	D23_1c1597	Neut_1513
fig/66666666.60966.peg.1596	CDS	1486114	1486503	1	+	390	Endoribonuclease L-PSP	CBSS-176299.4.peg.1996A	D23_1c1598	Neut_1514
fig/66666666.60966.peg.1598	CDS	1486998	1486651	-3	-	348	FIG016027; protein of unknown function YeaO	-none-	D23_1c1599	Neut_1515
fig/66666666.60966.peg.1599	CDS	1487053	1487727	1	+	675	AtfE component of AtfFGH ABC transport system	AtfFGH ABC Transport System	D23_1c1600	Neut_1516
fig/66666666.60966.peg.1600	CDS	1487724	1490273	3	+	2550	AtfF component of AtfFGH ABC transport system/AtfG component of AtfFGH ABC transport system	AtfFGH ABC Transport System; AtfFGH ABC Transport System	D23_1c1601	Neut_1517
fig/66666666.60966.peg.1601	CDS	1490273	1491343	2	+	1071	AtfH component of AtfFGH ABC transport system	AtfFGH ABC Transport System	D23_1c1602	Neut_1518

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 GbkId	C91 Alias
fig/66666666.60966.peg.1602	CDS	1491424	1491666	1	+	243	Molybdopterin biosynthesis protein B	-none-	D23_1c1603	Neut_1519
fig/66666666.60966.peg.1603	CDS	1491738	1491619	-3	-	120	hypothetical protein	-none-	D23_1c1604	NA
fig/66666666.60966.peg.1605	CDS	1492158	1492982	3	+	825	Particulate methane monooxygenase C-subunit (EC 1.14.13.25)	Particulate methane monooxygenase (pMMO)	D23_1c1605	Neut_1520
fig/66666666.60966.peg.1607	CDS	1494238	1493309	-1	-	930	Cytochrome c551 peroxidase (EC 1.11.1.5)	Protection from Reactive Oxygen Species	D23_1c1607	Neut_1521
fig/66666666.60966.peg.1609	CDS	1494804	1495091	3	+	288	Mobile element protein	-none-	D23_1c1608	Neut_2502
fig/66666666.60966.peg.1611	CDS	1495989	1495288	-3	-	702	2-C-methyl-D-erythritol 4-phosphate cytidyltransferase (EC 2.7.7.60)	Isoprenoid Biosynthesis; Nonmevalonate Branch of Isoprenoid Biosynthesis; Possible RNA degradation cluster; Stationary phase repair cluster	D23_1c1609	Neut_1525
fig/66666666.60966.peg.1612	CDS	1496057	1496497	2	+	441	Inosine-5'-monophosphate dehydrogenase (EC 1.1.1.205)	Purine conversions; Purine salvage cluster	D23_1c1610	Neut_1526
fig/66666666.60966.peg.1613	CDS	1497472	1496552	-1	-	921	Cell division protein FisX	Bacterial Cell Division; Heat shock Cell division Proteases and a Methyltransferase	D23_1c1611	Neut_1527
fig/66666666.60966.peg.1614	CDS	1498131	1497469	-3	-	663	Cell division transporter, ATP-binding protein FisE (TC 3.A.5.1.1)	Bacterial Cell Division; Heat shock Cell division Proteases and a Methyltransferase	D23_1c1612	Neut_1528
fig/66666666.60966.peg.1615	CDS	1499195	1498158	-2	-	1038	Signal recognition particle receptor protein FisY (=alpha subunit) (TC 3.A.5.1.1)	Bacterial Cell Division; Bacterial signal recognition particle (SRP); Heat shock Cell division Proteases and a Methyltransferase; Universal GTPases	D23_1c1613	Neut_1529
fig/66666666.60966.peg.1616	CDS	1499262	1500653	3	+	1392	FIG015547; peptidase, M16 family	-none-	D23_1c1614	Neut_1530
fig/66666666.60966.peg.1617	CDS	1500780	1501865	3	+	1086	Alanine racemase (EC 5.1.1.1)	Alanine biosynthesis; Pyruvate Alanine Serine Interconversions	D23_1c1615	Neut_1531
fig/66666666.60966.peg.1618	CDS	1502688	1501894	-3	-	795	Peptidyl-prolyl cis-trans isomerase (EC 5.2.1.8)	Queuosine-Archaeosine Biosynthesis	D23_1c1616	Neut_1532
fig/66666666.60966.peg.1619	CDS	1502854	1502738	-1	-	117	hypothetical protein	-none-	D23_1c1617	NA
fig/66666666.60966.peg.1620	CDS	1503164	1502862	-2	-	303	YcL protein	Broadly distributed proteins not in subsystems; CBSS-211586.9.peg.2729	D23_1c1618	Neut_1533

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld	C91 Alias
fig/66666666.60966.peg.1621	CDS	1504431	1503277	-3	-	1155	Rubredoxin-NAD(+) reductase (EC 1.18.1.1)	Ruberythrin	D23_1c1619	Neut_1534
fig/66666666.60966.peg.1622	CDS	1504748	1505626	2	+	879	Probable protease htpX homolog (EC 3.4.24.-)	-none-	D23_1c1620	Neut_1535
fig/66666666.60966.peg.1623	CDS	1505626	1505739	1	+	114	hypothetical protein	-none-	D23_1c1621	NA
fig/66666666.60966.peg.1624	CDS	1506477	1505698	-3	-	780	Surface lipoprotein	-none-	D23_1c1622	Neut_1536
fig/66666666.60966.peg.1625	CDS	1507909	1506626	-1	-	1284	Glutamate-1-semialdehyde aminotransferase (EC 5.4.3.8)	CBS5-196164.1,peg.461; Heme and Siroheme Biosynthesis	D23_1c1623	Neut_1537
fig/66666666.60966.peg.1626	CDS	1508584	1507946	-1	-	639	Thiamin-phosphate pyrophosphorylase (EC 2.5.1.3)	5-FCL-like protein; Thiamin biosynthesis	D23_1c1624	Neut_1538
fig/66666666.60966.peg.1627	CDS	1509419	1508577	-2	-	843	Phosphomethylpyrimidine kinase (EC 2.7.4.7)	5-FCL-like protein; Thiamin biosynthesis	D23_1c1625	Neut_1539
fig/66666666.60966.peg.1628	CDS	1509508	1509660	1	+	153	Rubredoxin	Ruberythrin	D23_1c1626	Neut_1540
fig/66666666.60966.peg.1629	CDS	1509660	1510049	3	+	390	Lactoylglutathione lyase (EC 4.4.1.5)	Glutathione: Non-redox reactions; Methylglyoxal Metabolism	D23_1c1627	Neut_1541
fig/66666666.60966.peg.1630	CDS	1510238	1510603	2	+	366	probable iron binding protein from the HesB_IscA_SufA family	 Metabolism	D23_1c1628	Neut_1542
fig/66666666.60966.peg.1631	CDS	1511703	1510612	-3	-	1092	Anhydro-N-acetylmuramic acid kinase (EC 2.7.1.-)	Recycling of Peptidoglycan Amino Sugars	D23_1c1629	Neut_1543
fig/66666666.60966.peg.1632	CDS	1513070	1511739	-2	-	1332	Peptidase, M23/M37 family	-none-	D23_1c1630	Neut_1544
fig/66666666.60966.peg.1633	CDS	1513163	1514383	2	+	1221	Tyrosyl-tRNA synthetase (EC 6.1.1.1)	tRNA aminoacylation, Tyr	D23_1c1631	Neut_1545
fig/66666666.60966.peg.1634	CDS	1514577	1515674	3	+	1098	Myo-inositol 2-dehydrogenase (EC 1.1.1.18)	-none-	D23_1c1632	Neut_1547
fig/66666666.60966.peg.1635	CDS	1515687	1516748	3	+	1062	N-Acetylneuraminatase cytidyltransferase (EC 2.7.7.43)	CMP-N-acetylneuraminatase Biosynthesis; Sialic Acid Metabolism	D23_1c1633	Neut_1548
fig/66666666.60966.peg.1636	CDS	1516926	1518425	3	+	1500	N-acetylneuraminatase synthase (EC 2.5.1.56)	CMP-N-acetylneuraminatase Biosynthesis; Sialic Acid Metabolism	D23_1c1634	Neut_1549
fig/66666666.60966.peg.1637	CDS	1518919	1518437	-1	-	483	Ribonucleotide reductase transcriptional regulator NrdR	Ribonucleotide reduction	D23_1c1635	Neut_1551

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld	C91 Alias
fig/66666666.60966.peg.1638	CDS	1520246	1518996	-2	-	1251	Serine hydroxymethyltransferase (EC 2.1.2.1)	5-FLC-like protein; Glycine Biosynthesis; Glycine and Serine Utilization; Photorespiration (oxidative C2 cycle); Serine Biosynthesis	D23_1c1636	Neut_1552
fig/66666666.60966.peg.1639	CDS	1520440	1521180	1	+	741	PqqC-like protein	Folate Biosynthesis	D23_1c1638	Neut_1553
fig/66666666.60966.peg.1640	CDS	1522218	1521283	-3	-	936	Transcriptional activator MetR	LysR-family proteins in <i>Escherichia coli</i> ; LysR-family proteins in <i>Salmonella enterica Typhimurium</i> ; Methionine Biosynthesis	D23_1c1639	Neut_1554
fig/66666666.60966.peg.1641	CDS	1522318	1524594	1	+	2277	5-methyltetrahydropteroyltriglutamate--homocysteine methyltransferase (EC 2.1.1.14)	Methionine Biosynthesis	D23_1c1640	Neut_1555
fig/66666666.60966.peg.1643	CDS	1526020	1524767	-1	-	1254	UDP-N-acetylglucosamine 1-carboxyvinyltransferase (EC 2.5.1.7)	Peptidoglycan Biosynthesis; UDP-N-acetylmuramate from Fructose-6-phosphate Biosynthesis	D23_1c1641	Neut_1556
fig/66666666.60966.peg.1644	CDS	1526108	1526497	2	+	390	Putative translation initiation inhibitor, yjgF family	-none-	D23_1c1642	Neut_1557
fig/66666666.60966.peg.1645	CDS	1526528	1528585	2	+	2058	ATP-dependent DNA helicase RecG (EC 3.6.1.—)	-none-	D23_1c1643	Neut_1558
fig/66666666.60966.peg.1646	CDS	1529436	1528588	-3	-	849	Hypothetical ATP-binding protein UPF0042, contains P-loop	-none-	D23_1c1644	Neut_1559
fig/66666666.60966.peg.1647	CDS	1529521	1530072	1	+	552	Transcription elongation factor GreB	CBSS-243265.1.peg.198; Transcription factors bacterial	D23_1c1645	Neut_1560
fig/66666666.60966.peg.1648	CDS	1532787	1530136	-3	-	2652	Malto-oligosyltrehalose synthase (EC 5.4.99.15)	-none-	D23_1c1646	NA
fig/66666666.60966.peg.1649	CDS	1533038	1532850	-2	-	189	hypothetical protein	-none-	D23_1c1647	NA
fig/66666666.60966.peg.1650	CDS	1534716	1533031	-3	-	1686	Malto-oligosyltrehalose trehalohydrolase (EC 3.2.1.141)	-none-	D23_1c1648	Neut_1291
fig/66666666.60966.peg.1651	CDS	1535327	1534896	-2	-	432	Trehalose synthase, nucleoside diphosphate glucose dependent	-none-	D23_1c1649	NA

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld	C91 Alias
fig/66666666.60966.peg.1652	CDS	1535502	1535386	-3	-	117	hypothetical protein	-none-	D23_1c1650	NA
	CDS	1535501	1536028	2	+	528	Sensory histidine kinase QseC		D23_1c1651	Neut_1565
fig/66666666.60966.peg.1654	CDS	1535992	1536651	1	+	660	Sensory histidine kinase QseC	-none-	D23_1c1652	Neut_1565
fig/66666666.60966.peg.1655	CDS	1537148	1537849	2	+	702	Protein of unknown function DUF484	-none-	D23_1c1653	Neut_1566
fig/66666666.60966.peg.1656	CDS	1537833	1538798	3	+	966	Tyrosine recombinase XerC	-none-	D23_1c1654	Neut_1567
fig/66666666.60966.peg.1657	CDS	1539747	1538845	-3	-	903	Arogenate dehydrogenase (EC 1.3.1.43)	Chorismate Synthesis; Phenylalanine and Tyrosine Branches from Chorismate	D23_1c1655	Neut_1568
fig/66666666.60966.peg.1658	CDS	1540893	1539775	-3	-	1119	Biosynthetic Aromatic amino acid aminotransferase beta (EC 2.6.1.57)		D23_1c1656	Neut_1569
fig/66666666.60966.peg.1659	CDS	1541973	1540915	-3	-	1059	Chorismate mutase I (EC 5.4.99.5)	Chorismate Synthesis; Synthesis; Phenylalanine and Tyrosine Branches from Chorismate; Phenylalanine and Tyrosine Branches from Chorismate	D23_1c1657	Neut_1570
							Prephenate dehydratase (EC 4.2.1.51)			
fig/66666666.60966.peg.1660	CDS	1543230	1542013	-3	-	1218	D-3-phosphoglycerate dehydrogenase (EC 1.1.1.95)	Glycine and Serine Utilization; Pyridoxin (Vitamin B6) Biosynthesis; Serine Biosynthesis	D23_1c1658	Neut_1571
fig/66666666.60966.peg.1661	CDS	1544329	1543223	-1	-	1107	Phosphoserine aminotransferase (EC 2.6.1.52)	Glycine and Serine Utilization; Pyridoxin (Vitamin B6) Biosynthesis; Serine Biosynthesis	D23_1c1659	Neut_1572
fig/66666666.60966.peg.1662	CDS	1546893	1544347	-3	-	2547	DNA gyrase subunit A (EC 5.99.1.3)	 Serine Biosynthesis Cell Division Subsystem including YidCD; DNA gyrase subunits; DNA replication cluster 1; DNA topoisomerases, Type II, ATP-dependent; Resistance to fluoroquinolones	D23_1c1660	Neut_1573

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld	C91 Alias
fig/6666666.60966.peg.1663	CDS	1547556	1546963	-3	-	594	COGs COG2854	-none-	D23_1c1661	Neat_1574
	CDS	1548691	1547573	-1	-	1119	Glycosyl transferase, family 2	-none-	D23_1c1662	Neat_1575
fig/6666666.60966.peg.1665	CDS	1549918	1548755	-1	-	1164	Possible Fe—S oxidoreductase	-none-	D23_1c1663	Neat_1576
fig/6666666.60966.peg.1666	CDS	1550134	1550247	1	+	114	hypothetical protein	-none-	D23_1c1664	NA
	CDS	1550439	1552481	3	+	2043	Transketolase (EC 2.2.1.1)	 Calvin-Benson cycle; Pentose phosphate pathway	D23_1c1666	Neat_1577
fig/6666666.60966.peg.1669	CDS	1552544	1553551	2	+	1008	NADPH-dependent glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.13)/NAD-dependent glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12)	Calvin-Benson cycle; Calvin-Benson cycle; Glycolysis and Gluconeogenesis; Glycolysis and Gluconeogenesis; Pyridoxin (Vitamin B6) Biosynthesis	D23_1c1667	Neat_1578
fig/6666666.60966.peg.1670	CDS	1553775	1553659	-3	-	117	hypothetical protein	-none-	D23_1c1668	NA
	CDS	1553870	1555048	2	+	1179	Phosphoglycerate kinase (EC 2.7.2.3)	Calvin-Benson cycle; Glycolysis and Gluconeogenesis	D23_1c1669	Neat_1579
fig/6666666.60966.peg.1672	CDS	1555083	1556573	3	+	1491	Pyruvate kinase (EC 2.7.1.40)	Glycerate metabolism; Glycolysis and Gluconeogenesis; Pyruvate metabolism I; anaplerotic reactions, PEP	D23_1c1670	Neat_1580
fig/6666666.60966.peg.1673	CDS	1556682	1557746	3	+	1065	Fructose-bisphosphate aldolase class II (EC 4.1.2.13)	Calvin-Benson cycle; Glycolysis and Gluconeogenesis	D23_1c1671	Neat_1581
fig/6666666.60966.peg.1674	CDS	1558432	1560090	1	+	1659	Cytochrome c oxidases subunit CcoN (EC 1.9.3.1)	Terminal cytochrome C oxidases	D23_1c1672	Neat_1582
fig/6666666.60966.peg.1675	CDS	1560080	1560688	2	+	609	Cytochrome c oxidase subunit CcoO (EC 1.9.3.1)	Terminal cytochrome C oxidases	D23_1c1673	Neat_1583
fig/6666666.60966.peg.1676	CDS	1560753	1561364	3	+	612	Copper-containing nitrite reductase (EC 1.7.2.1)	Denitrification; Denitrifying reductase gene clusters	D23_1c1674	Neat_1584
fig/6666666.60966.peg.1677	CDS	1561404	1562054	3	+	651	Cytochrome oxidase biogenesis protein ScoI/ScoC/PrnC, putative copper metallochaperone	Biogenesis of cytochrome c oxidases	D23_1c1675	Neat_1585

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.									
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld C91 Alias
fig/66666666.60966.peg.1678	CDS	1562150	1562635	2	+	486	hypothetical cytochrome oxidase associated membrane protein	-none-	D23_1c1676 Neut_1586
fig/66666666.60966.peg.1679	CDS	1563699	1562746	-3	-	954	Rare lipoprotein A precursor	Peptidoglycan Biosynthesis	D23_1c1677 Neut_1587
fig/66666666.60966.peg.1680	CDS	1564813	1563704	-1	-	1110	Rod shape-determining protein RodA	Bacterial Cytoskeleton; Bacterial cell division cluster; Peptidoglycan Biosynthesis	D23_1c1678 Neut_1588
fig/66666666.60966.peg.1681	CDS	1566713	1564836	-2	-	1878	Penicillin-binding protein 2 (PBP-2)	16S rRNA modification within P site of ribosome; Bacterial cell division cluster; CBSS-83331.1.peg.3039; Peptidoglycan Biosynthesis	D23_1c1679 Neut_1589
fig/66666666.60966.peg.1682	CDS	1567261	1566710	-1	-	552	Rod shape-determining protein MreD	Bacterial Cell Division; Bacterial Cytoskeleton; Bacterial cell division cluster; CBSS-354.1.peg.2917	D23_1c1680 Neut_1590
fig/66666666.60966.peg.1683	CDS	1568123	1567233	-2	-	891	Rod shape-determining protein MreC	Bacterial Cell Division; Bacterial Cytoskeleton; Bacterial cell division cluster; CBSS-	D23_1c1681 Neut_1591
fig/66666666.60966.peg.1684	CDS	1569547	1568486	-1	-	1062	Rod shape-determining protein MreB	354.1.peg.2917 Bacterial Cell Division; Bacterial Cytoskeleton; Bacterial cell division cluster; CBSS-	D23_1c1682 Neut_1592
fig/66666666.60966.peg.1685	CDS	1569665	1569997	2	+	333	Aspartyl-tRNA(Asn) amidotransferase subunit C (EC 6.3.5.6) @ Glutamyl-tRNA(Gln) amidotransferase subunit C (EC 6.3.5.7)	Bacterial Cell Division; Bacterial Cytoskeleton; Bacterial cell division cluster	D23_1c1683 Neut_1593
fig/66666666.60966.peg.1686	CDS	1570062	1571522	3	+	1461	Aspartyl-tRNA(Asn) amidotransferase subunit A (EC 6.3.5.6) @ Glutamyl-tRNA(Gln)	tRNA aminoacylation, Asp and Asn; tRNA aminoacylation, Glu and Gln	D23_1c1684 Neut_1594

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.									
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkid C91 Alias
fig/66666666.60966.peg.1687	CDS	1571587	1573023	1	+	1437	amidotransferase subunit A (EC 6.3.5.7) Aspartyl-tRNA(Asn) amidotransferase subunit B (EC 6.3.5.6) @ Glutamyl-tRNA(Gln) amidotransferase subunit B (EC 6.3.5.7) hypothetical protein FIG00859257; hypothetical protein hear shock protein, Hsp20 family	tRNA aminoacylation, Asp and Asn; tRNA aminoacylation, Glu and Gln	D23_1c1685 Neut_1595
fig/66666666.60966.peg.1688	CDS	1573150	1573001	-1	-	150	hypothetical protein	-none-	D23_1c1686 NA
fig/66666666.60966.peg.1689	CDS	1573320	1573129	-3	-	192	hypothetical protein	-none-	D23_1c1687 NA
fig/66666666.60966.peg.1691	CDS	1574140	1573703	-1	-	438	hypothetical protein Hsp20 family	-none-	D23_1c1688 Neut_1596
fig/66666666.60966.peg.1692	CDS	1574662	1574802	1	+	141	Integrase	-none-	D23_1c1690 Neut_1498
fig/66666666.60966.peg.1693	CDS	1575055	1574942	-1	-	114	hypothetical protein	-none-	D23_1c1692 NA
fig/66666666.60966.peg.1694	CDS	1575572	1575369	-2	-	204	hypothetical protein	-none-	D23_1c1693 NA
fig/66666666.60966.peg.1696	CDS	1575911	1575753	-2	-	159	Mobile element protein	-none-	D23_1c1694 Neut_1094
fig/66666666.60966.peg.1697	CDS	1576265	1578433	2	+	2169	GTP pyrophosphokinase (EC 2.7.6.5), (p)ppGpp synthetase II/ Guanosine-3',5'-bis(diphosphate) 3'-	Stringent Response, (p)ppGpp metabolism; Stringent Response, (p)ppGpp metabolism	D23_1c1695 Neut_1601
fig/66666666.60966.peg.1698	CDS	1578452	1579135	2	+	684	pyrophosphohydrolase (EC 3.1.7.2)	-none-	D23_1c1696 Neut_1602
fig/66666666.60966.peg.1699	CDS	1579798	1579145	-1	-	654	hypothetical protein Periplasmic thiol:disulfide interchange protein DsbA	Biogenesis of c-type cytochromes; Periplasmic disulfide interchange	D23_1c1697 Neut_1603
fig/66666666.60966.peg.1700	CDS	1580580	1579915	-3	-	666	Cell division protein	-none-	D23_1c1698 Neut_1604
fig/66666666.60966.peg.1701	CDS	1582304	1580604	-2	-	1701	Arginyl-tRNA synthetase (EC 6.1.1.19)	tRNA aminoacylation, Arg	D23_1c1699 Neut_1605
fig/66666666.60966.peg.1702	CDS	1582613	1583272	2	+	660	FIG00859197; hypothetical protein	-none-	D23_1c1700 Neut_1606
fig/66666666.60966.peg.1703	CDS	1583410	1584297	1	+	888	Methylenetetrahydrofolate dehydrogenase (NADP+) (EC 1.5.1.5)/ Methylenetetrahydrofolate cyclohydrolase (EC 3.5.4.9)	5-FCL-like protein; -br>One-carbon metabolism by tetrahydropterines; -br>One-carbon metabolism by tetrahydropterines	D23_1c1701 Neut_1607
fig/66666666.60966.peg.1704	CDS	1584339	1586999	3	+	2661	Pyruvate dehydrogenase E1 component (EC 1.2.4.1)	5-FCL-like protein; -br>Dehydrogenase complexes; -br>Methionine	D23_1c1702 Neut_1608

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.									
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkid C91 Alias
fig/6666666.60966.peg.1705	CDS	1587072	1588421	3	+	1350	Dihydroliipoamide acetyltransferase component of pyruvate dehydrogenase complex (EC 2.3.1.12)	Degradation; Pyruvate metabolism II: acetyl-CoA, acetogenesis from pyruvate	D23_1c1703 Neut_1609
fig/6666666.60966.peg.1706	CDS	1588457	1589170	2	+	714	Nicotinate-nucleotide adenyltransferase (EC 2.7.7.18)	5-FCL-like protein; Dehydrogenase complexes; Pyruvate metabolism II: acetyl-CoA, acetogenesis from pyruvate	D23_1c1704 Neut_1610
fig/6666666.60966.peg.1707	CDS	1589167	1589532	1	+	366	lojap protein	-none-	D23_1c1705 Neut_1611
fig/6666666.60966.peg.1708	CDS	1589624	1590091	2	+	468	LSU m3Psi1915 methyltransferase RlmH	RNA methylation; Ribosome biogenesis bacterial; rRNA modification	D23_1c1706 Neut_1612
fig/6666666.60966.peg.1709	CDS	1590169	1590792	1	+	624	Septum formation protein Maf	Bacteria	D23_1c1707 Neut_1613
fig/6666666.60966.peg.1710	CDS	1590825	1592276	3	+	1452	Cytoplasmic axial filament protein CafA and Ribonuclease G (EC 3.1.4.—)	Bacterial Cell Division; Bacterial	D23_1c1708 Neut_1614
fig/6666666.60966.peg.1711	CDS	1592500	1593222	1	+	723	Ferric siderophore transport system.	Cytoskeleton; Bacterial cell division cluster; CBSS-	D23_1c1709 Neut_1615
fig/6666666.60966.peg.1712	CDS	1593226	1594023	1	+	798	periplasmic binding protein TonB MotA/TolQ/ExbB proton channel family protein	354.1.peg.2917 Bacterial Cell Division; CBSS-	D23_1c1710 Neut_1616
fig/6666666.60966.peg.1713	CDS	1594023	1594448	3	+	426	Biopolymer transport protein ExbD/TolR	354.1.peg.2917; RNA processing and degradation, bacterial	D23_1c1711 Neut_1617
fig/6666666.60966.peg.1714	CDS	1595300	1594518	-2	-	783	23S rRNA (guanosine-2'-O-)-methyltransferase rlmB (EC 2.1.1.—)	Ton and Tol transport systems	D23_1c1712 Neut_1618
fig/6666666.60966.peg.1715	CDS	1597532	1595328	-2	-	2205	3'-to-5' exonuclease RNase R	Ton and Tol transport systems	D23_1c1713 Neut_1619

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 GbkId	C91 Alias
fig/66666666.60966.peg.1717	CDS	1598152	1599303	1	+	1152	DNA polymerase IV (EC 2.7.7.7)	DNA repair, bacterial	D23_1c1715	Neut_1620
fig/66666666.60966.peg.1718	CDS	1599507	1599391	-3	-	117	hypothetical protein	-none-	D23_1c1716	NA
fig/66666666.60966.peg.1719	CDS	1599639	1601015	3	+	1377	Flagellar regulatory protein FlcQ	Flagellum	D23_1c1717	Neut_1621
fig/66666666.60966.peg.1720	CDS	1601809	1601045	-1	-	765	hypothetical protein	-none-	D23_1c1718	Neut_1622
fig/66666666.60966.peg.1721	CDS	1604085	1601893	-3	-	2193	hypothetical protein	-none-	D23_1c1719	Neut_1623
fig/66666666.60966.peg.1722	CDS	1605035	1604157	-2	-	879	Mobile element protein	-none-	D23_1c1720	Neut_1720
fig/66666666.60966.peg.1723	CDS	1605427	1605134	-1	-	294	Mobile element protein	-none-	D23_1c1721	Neut_1719
fig/66666666.60966.peg.1724	CDS	1606591	1605455	-1	-	1137	8-amino-7-oxononanoate synthase (EC 2.3.1.47)	Biotin biosynthesis; Biotin biosynthesis Experimental; Biotin synthesis cluster	D23_1c1722	Neut_2137
fig/66666666.60966.peg.1725	CDS	1608741	1606597	-3	-	2145	hypothetical protein	-none-	D23_1c1723	NA
fig/66666666.60966.peg.1726	CDS	1610072	1608738	-2	-	1335	hypothetical protein	-none-	D23_1c1724	NA
fig/66666666.60966.peg.1727	CDS	1610798	1610106	-2	-	693	hypothetical protein	-none-	D23_1c1725	Neut_1626
fig/66666666.60966.peg.1728	CDS	1611531	1610833	-3	-	699	hypothetical protein	-none-	D23_1c1726	Neut_1626
fig/66666666.60966.peg.1729	CDS	1611922	1612056	1	+	135	hypothetical protein	-none-	D23_1c1727	NA
fig/66666666.60966.peg.1731	CDS	1612368	1612991	3	+	624	InterPro IPR000379 COGs COG2945	-none-	D23_1c1729	Neut_1627
fig/66666666.60966.peg.1732	CDS	1614001	1613204	-1	-	798	Carbonic anhydrase (EC 4.2.1.1)	Zinc regulated enzymes	D23_1c1730	Neut_1628
fig/66666666.60966.peg.1734	CDS	1614522	1614355	-3	-	168	hypothetical protein	-none-	D23_1c1731	NA
fig/66666666.60966.peg.1735	CDS	1614827	1614699	-2	-	129	hypothetical protein	-none-	D23_1c1732	NA
fig/66666666.60966.peg.1736	CDS	1614794	1616134	2	+	1341	Probable transmembrane protein	-none-	D23_1c1733	Neut_1630
fig/66666666.60966.peg.1737	CDS	1617043	1616225	-1	-	819	rRNA methylases	-none-	D23_1c1734	Neut_1631
fig/66666666.60966.peg.1738	CDS	1617516	1617061	-3	-	456	Mobile element protein	-none-	D23_1c1735	Neut_2502
fig/66666666.60966.peg.1739	CDS	1617913	1617479	-1	-	435	Mobile element protein	-none-	D23_1c1736	Neut_0884
fig/66666666.60966.peg.1740	CDS	1618427	1618293	-2	-	135	hypothetical protein	-none-	D23_1c1738	NA
fig/66666666.60966.peg.1741	CDS	1619838	1618420	-3	-	1419	Dihydroliopamide dehydrogenase (EC 1.8.1.4)	5-FLC-like protein; Glycine cleavage system; Photorepiration (oxidative C2 cycle); TCA Cycle	D23_1c1739	Neut_1632
fig/66666666.60966.peg.1742	CDS	1620067	1621053	1	+	987	Malate dehydrogenase (EC 1.1.1.37)	TCA Cycle	D23_1c1740	Neut_1633
fig/66666666.60966.peg.1743	CDS	1621562	1621107	-2	-	456	Thiol peroxidase, Bcp-type (EC 1.11.1.15)	CBSS-316057.3.peg.3521; Thioredoxin-disulfide reductase	D23_1c1741	Neut_1634
fig/66666666.60966.peg.1744	CDS	1622938	1621595	-1	-	1344	Cytochrome c heme lyase subunit CcmH	Biogenesis of c-type cytochromes; Copper homeostasis	D23_1c1742	Neut_1635
fig/66666666.60966.peg.1745	CDS	1623399	1622935	-3	-	465	Cytochrome c heme lyase subunit CcmL	Biogenesis of c-type cytochromes	D23_1c1743	Neut_1636

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld	C91 Alias
fig/66666666.60966.peg.1746	CDS	1623982	1623458	-1	-	525	Cytochrome c-type biogenesis protein CcmG/DsbE, thiol:disulfide oxidoreductase	Biogenesis of c-type cytochromes; Periplasmic disulfide interchange	D23_1c1744	Neut_1637
fig/66666666.60966.peg.1747	CDS	1626024	1623979	-3	-	2046	Cytochrome c heme lyase subunit CcmF	Biogenesis of c-type cytochromes; Copper homeostasis	D23_1c1745	Neut_1638
fig/66666666.60966.peg.1748	CDS	1626514	1626065	-1	-	450	Cytochrome c-type biogenesis protein CcmE, heme chaperone	Biogenesis of c-type cytochromes	D23_1c1746	Neut_1639
fig/66666666.60966.peg.1749	CDS	1627373	1626690	-2	-	684	Cytochrome c-type biogenesis protein CcmC, putative heme lyase for CcmE	Biogenesis of c-type cytochromes	D23_1c1747	Neut_1641
fig/66666666.60966.peg.1750	CDS	1628195	1627536	-2	-	660	ABC transporter involved in cytochrome c biogenesis, CcmB subunit	Biogenesis of c-type cytochromes	D23_1c1748	Neut_1642
fig/66666666.60966.peg.1751	CDS	1628729	1628202	-2	-	528	ABC transporter involved in cytochrome c biogenesis, ATPase component CcmA	Biogenesis of c-type cytochromes	D23_1c1749	Neut_1643
fig/66666666.60966.peg.1752	CDS	1629793	1628927	-1	-	867	tRNA pseudouridine synthase B (EC 4.2.1.70)	CBSS-138119.3.peg.2719; RNA pseudouridine syntheses; Riboflavin, FMN and FAD metabolism in plants; tRNA modification Bacteria; tRNA processing	D23_1c1750	Neut_1644
fig/66666666.60966.peg.1753	CDS	1630354	1630001	-1	-	354	Ribosome-binding factor A	CBSS-138119.3.peg.2719; NusA-TFH Cluster; Translation initiation factors bacterial	D23_1c1751	Neut_1645
fig/66666666.60966.peg.1754	CDS	1633073	1630407	-2	-	2667	Translation initiation factor 2	CBSS-138119.3.peg.2719; NusA-TFH Cluster; Translation initiation factors bacterial; Universal GTPases	D23_1c1752	Neut_1646
fig/66666666.60966.peg.1756	CDS	1634664	1633192	-3	-	1473	Transcription termination protein NusA	NusA-TFH Cluster; Transcription factors bacterial	D23_1c1753	Neut_1647

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld	C91 Alias
fig/66666666.60966.peg.1757	CDS	1635199	1634717	-1	-	483	COG0779: clustered with transcription termination protein NusA	-none-	D23_1c1754	Neut_1648
fig/66666666.60966.peg.1760	CDS	1636174	1636611	1	+	438	FIG00859331: hypothetical protein	-none-	D23_1c1755	Neut_1650
fig/66666666.60966.peg.1761	CDS	1637012	1636686	-2	-	327	Cytochrome c4	Soluble cytochromes and functionally related electron carriers	D23_1c1756	Neut_1651
fig/66666666.60966.peg.1762	CDS	1637371	1637045	-1	-	327	Putative periplasmic cytochrome type-C oxidoreductase signal peptide protein (EC 1.---,---,---)	-none-	D23_1c1757	Neut_1652
fig/66666666.60966.peg.1763	CDS	1639430	1637484	-2	-	1947	COG0488: ATPase components of ABC transporters with duplicated ATPase domains	-none-	D23_1c1758	Neut_1653
fig/66666666.60966.peg.1765	CDS	1639660	1640709	1	+	1050	Selenide, water dikinase (EC 2.7.9.3)	Selenocysteine metabolism; rRNA modification Bacteria	D23_1c1759	Neut_1654
fig/66666666.60966.peg.1766	CDS	1640702	1641868	2	+	1167	Selenophosphate-dependent tRNA 2-selenouridine synthase	Selenocysteine metabolism; rRNA modification Bacteria	D23_1c1760	Neut_1655
fig/66666666.60966.peg.1767	CDS	1642217	1642050	-2	-	168	hypothetical protein	-none-	D23_1c1761	NA
fig/66666666.60966.peg.1768	CDS	1643785	1642217	-1	-	1569	4-cresol dehydrogenase [hydroxylating] flavoprotein subunit (EC 1.17.99.1)	Cresol degradation	D23_1c1762	Neut_1656
fig/66666666.60966.peg.1769	CDS	1644507	1643782	-3	-	726	PchX protein	-none-	D23_1c1763	Neut_1657
fig/66666666.60966.peg.1770	CDS	1644874	1644518	-1	-	357	4-cresol dehydrogenase [hydroxylating] cytochrome c subunit precursor	Cresol degradation	D23_1c1764	Neut_1658
fig/66666666.60966.peg.1771	CDS	1646175	1645138	-3	-	1038	Dihydroorotase (EC 3.5.2.3)	De Novo Pyrimidine Synthesis; Zinc regulated enzymes	D23_1c1766	Neut_1659
fig/66666666.60966.peg.1772	CDS	1647130	1646273	-1	-	858	rRNA small subunit methyltransferase I	16S rRNA modification within P site of ribosome; CBSS-160492.1.peg.550; Heat shock dnaK gene cluster extended	D23_1c1767	Neut_1660
fig/66666666.60966.peg.1773	CDS	1647311	1649101	2	+	1791	ABC transporter, multidrug efflux family	-none-	D23_1c1769	Neut_1661
fig/66666666.60966.peg.1774	CDS	1649206	1649556	1	+	351	Predicted endonuclease distantly related to	CBSS-160492.1.peg.550	D23_1c1770	Neut_1662

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.									
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	C91 Alias
fig/66666666.60966.peg.1776	CDS	1651143	1649929	-3	-	1215	archaeal Holliday junction resolvase	-none-	Neut_1663
fig/66666666.60966.peg.1777	CDS	1651568	1651446	-2	-	123	hypothetical protein	-none-	NA
fig/66666666.60966.peg.1778	CDS	1651714	1652223	1	+	510	Putative lipoprotein	-none-	Neut_1664
fig/66666666.60966.peg.1779	CDS	1652740	1661706	1	+	8967	Cyclic beta-1,2-glucan synthase (EC 2.4.1.—)	Synthesis of osmoregulated periplasmic glucans	Neut_1665
fig/66666666.60966.peg.1780	CDS	1662138	1661821	-3	-	318	Mobile element protein	-none-	Neut_1666
fig/66666666.60966.peg.1781	CDS	1662168	1662344	3	+	177	hypothetical protein	-none-	NA
fig/66666666.60966.peg.1782	CDS	1662546	1662406	-3	-	141	Mobile element protein	-none-	Neut_2190
fig/66666666.60966.peg.1783	CDS	1663286	1663158	-2	-	129	hypothetical protein	-none-	NA
fig/66666666.60966.peg.1784	CDS	1663556	1664518	2	+	963	Mobile element protein	-none-	Neut_1746
fig/66666666.60966.peg.1785	CDS	1666051	1664585	-1	-	1467	ATP-dependent RNA helicase	ATP-dependent RNA helicases, bacterial	Neut_1668
fig/66666666.60966.peg.1786	CDS	1666377	1666147	-3	-	231	Bcep18194_A5658	-none-	Neut_2088
fig/66666666.60966.peg.1787	CDS	1666620	1666441	-3	-	180	Mobile element protein	-none-	Neut_0332
fig/66666666.60966.peg.1788	CDS	1667117	1666644	-2	-	474	Mobile element protein	-none-	NA
fig/66666666.60966.peg.1789	CDS	1668091	1667294	-1	-	798	FIG00861229;	-none-	Neut_1669
fig/66666666.60966.peg.1790	CDS	1668254	1668093	-2	-	162	hypothetical protein	-none-	NA
fig/66666666.60966.peg.1791	CDS	1669037	1668330	-2	-	708	Cytochrome c family protein	-none-	Neut_2333
fig/66666666.60966.peg.1792	CDS	1670220	1669099	-3	-	1122	FIG00859557;	-none-	Neut_1792
fig/66666666.60966.peg.1793	CDS	1671929	1670217	-2	-	1713	Hydroxylamine oxidoreductase	-none-	Neut_2335
fig/66666666.60966.peg.1794	CDS	1672280	1672017	-2	-	264	precursor (EC 1.7.3.4) SSU ribosomal protein S20p	-none-	NA
fig/66666666.60966.peg.1795	CDS	1672473	1672634	3	+	162	hypothetical protein	-none-	NA
fig/66666666.60966.peg.1796	CDS	1672870	1672601	-1	-	270	Mobile element protein	-none-	Neut_2450
fig/66666666.60966.peg.1797	CDS	1673187	1672894	-3	-	294	hypothetical protein	-none-	Neut_2449
fig/66666666.60966.peg.1798	CDS	1673875	1673648	-1	-	228	hypothetical protein	-none-	Neut_1676
fig/66666666.60966.peg.1799	CDS	1674438	1674052	-3	-	387	FIG002082; Protein SirB2	A Gammaproteobacteria Cluster Relating to Translation	Neut_1677
fig/66666666.60966.peg.1800	CDS	1674844	1674500	-1	-	345	FIG00858740;	-none-	Neut_1678
fig/66666666.60966.peg.1801	CDS	1675595	1674939	-2	-	657	hypothetical protein	-none-	Neut_1679
fig/66666666.60966.peg.1802	CDS	1675804	1675601	-1	-	204	Probable membrane protein	-none-	Neut_1679
fig/66666666.60966.peg.1803	CDS	1676931	1675825	-3	-	1107	GbcA Glycine betaine demethylase subunit A	-none-	Neut_1680
fig/66666666.60966.peg.1804	CDS	1677301	1678686	1	+	1386	FIG00858667; hypothetical protein	-none-	Neut_1681

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld	C91 Alias
fig/66666666.60966.peg.1805	CDS	1678742	1680964	2	+	2223	Helicase PriA essential for oriC/DnaA-independent DNA replication	-none-	D23_1c1803	Neut_1682
fig/66666666.60966.peg.1806	CDS	1681478	1681032	-2	-	447	Universal stress protein	-none-	D23_1c1804	Neut_1683
fig/66666666.60966.peg.1807	CDS	1683246	1681597	-3	-	1650	Folate transporter 3	-none-	D23_1c1805	Neut_1684
fig/66666666.60966.peg.1808	CDS	1684397	1683258	-2	-	1140	Outer membrane stress sensor protease DegS	Periplasmic Stress Response; Proteolysis in bacteria, ATP-dependent	D23_1c1806	Neut_1685
fig/66666666.60966.peg.1809	CDS	1684396	1685181	1	+	786	FIG137478: Hypothetical protein YbgI	-none-	D23_1c1807	Neut_1686
fig/66666666.60966.peg.1810	CDS	1685374	1685249	-1	-	126	hypothetical protein	-none-	D23_1c1808	NA
fig/66666666.60966.peg.1811	CDS	1685328	1686644	3	+	1317	Membrane-bound lytic murein transglycosylase A precursor (EC 3.2.1.—)	Murein Hydrolases	D23_1c1809	Neut_1687
fig/66666666.60966.peg.1812	CDS	1687397	1686684	-2	-	714	Thiol:disulfide interchange protein DsbC	Periplasmic disulfide interchange	D23_1c1810	Neut_1688
fig/66666666.60966.peg.1813	CDS	1688609	1687443	-2	-	1167	2-octaprenyl-3-methyl-6-methoxy-1,4-benzoquinol hydroxylase (EC 1.14.13.—)	CBSS-87626.3.peg.3639; Ubiquinone Biosynthesis;	D23_1c1811	Neut_1689
fig/66666666.60966.peg.1814	CDS	1688840	1690807	2	+	1968	Acetyl-coenzyme A synthetase (EC 6.2.1.1)	 Ubiquinone Biosynthesis-glo Pynvate metabolism II: acetyl-CoA, acetogenesis from pynvate	D23_1c1812	Neut_1690
fig/66666666.60966.peg.1815	CDS	1690825	1691799	1	+	975	Beta-lactamase related protein	-none-	D23_1c1813	Neut_1691
fig/66666666.60966.peg.1816	CDS	1693183	1691903	-1	-	1281	amidohydrolase	-none-	D23_1c1814	Neut_1692
fig/66666666.60966.peg.1817	CDS	1693463	1693633	2	+	171	Mobile element protein	-none-	D23_1c1816	Neut_1693
fig/66666666.60966.peg.1818	CDS	1694024	1695034	2	+	1011	Fatty acid desaturase	-none-	D23_1c1817	Neut_1694
fig/66666666.60966.peg.1819	CDS	1696649	1695402	-2	-	1248	Mobile element protein	-none-	D23_1c1818	Neut_0357
fig/66666666.60966.peg.1820	CDS	1697303	1696830	-2	-	474	Mobile element protein	-none-	D23_1c1820	Neut_1256
fig/66666666.60966.peg.1821	CDS	1697769	1697377	-3	-	393	hypothetical protein	-none-	D23_1c1821	Neut_2449
fig/66666666.60966.peg.1822	CDS	1698235	1698032	-1	-	204	hypothetical protein	-none-	D23_1c1822	Neut_0363
fig/66666666.60966.peg.1823	CDS	1698639	1698322	-3	-	318	hypothetical protein	-none-	D23_1c1823	Neut_1695
fig/66666666.60966.peg.1824	CDS	1698977	1698639	-2	-	339	Mobile element protein	-none-	D23_1c1824	Neut_1696
fig/66666666.60966.peg.1825	CDS	1699393	1702311	1	+	2919	Na(+)/H(+) antiporter subunit A/Na(+)/H(+) antiporter subunit B	Multi-subunit cation antiporter; Multi-subunit cation antiporter	D23_1c1825	Neut_1697
fig/66666666.60966.peg.1826	CDS	1702311	1702655	3	+	345	Na(+)/H(+) antiporter subunit C	Multi-subunit cation antiporter	D23_1c1826	Neut_1698
fig/66666666.60966.peg.1827	CDS	1702652	1704298	2	+	1647	Na(+)/H(+) antiporter subunit D	Multi-subunit cation antiporter	D23_1c1827	Neut_1699
fig/66666666.60966.peg.1828	CDS	1704295	1704780	1	+	486	Na(+)/H(+) antiporter subunit E	Multi-subunit cation antiporter	D23_1c1828	Neut_1700

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.									
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld C91 Alias
fig/66666666.60966.peg.1829	CDS	1704777	1705058	3	+	282	Na(+) H(+) antiporter subunit F	Multi-subunit cation antiporter	D23_1c1829 Neut_1701
fig/66666666.60966.peg.1830	CDS	1705055	1705480	2	+	426	Na(+) H(+) antiporter subunit G	Multi-subunit cation antiporter	D23_1c1830 Neut_1702
fig/66666666.60966.peg.1831	CDS	1708136	1705638	-2	-	2499	FIG00809136: hypothetical protein	-none-	D23_1c1831 Neut_1703
fig/66666666.60966.peg.1832	CDS	1709066	1708266	-2	-	801	ABC transporter ATP-binding protein YycR	-none-	D23_1c1832 Neut_1704
fig/66666666.60966.peg.1833	CDS	1709065	1709676	1	+	612	Arylesterase precursor (EC 3.1.1.2)	-none-	D23_1c1833 Neut_1705
fig/66666666.60966.peg.1834	CDS	1710780	1709716	-3	-	1065	Ribosomal large subunit pseudouridine synthase C (EC 4.2.1.70)	RNA pseudouridine syntheses; Ribosome biogenesis bacterial	D23_1c1834 Neut_1706
fig/66666666.60966.peg.1835	CDS	1711189	1713738	1	+	2550	Ribonuclease E (EC 3.1.26.12)	RNA processing and degradation, bacterial; Ribosome biogenesis bacterial	D23_1c1835 Neut_1707
fig/66666666.60966.peg.1836	CDS	1714787	1713870	-2	-	918	Tyrosine recombinase XerD	-none-	D23_1c1836 Neut_1708
fig/66666666.60966.peg.1837	CDS	1715754	1714909	-3	-	846	CsaA-related protein	-none-	D23_1c1837 Neut_1709
fig/66666666.60966.peg.1838	CDS	1715897	1717246	2	+	1350	Signal recognition particle, subunit Ffh	Bacterial signal recognition particle (SRP); Universal GTPases	D23_1c1838 Neut_1710
fig/66666666.60966.peg.1839	CDS	1717562	1718059	2	+	498	Cytosine/adenosine deaminases	-none-	D23_1c1840 Neut_1711
fig/66666666.60966.peg.1840	CDS	1719089	1718079	-2	-	1011	collagen triple helix repeat domain protein	-none-	D23_1c1841 Neut_1712
fig/66666666.60966.peg.1841	CDS	1719941	1719435	-2	-	507	Mobile element protein	-none-	D23_1c1843 Neut_1353
fig/66666666.60966.peg.1843	CDS	1720927	1720256	-1	-	672	Putative TEGT family carrier/transport protein	CBSS-326442.4.peg.1852	D23_1c1844 Neut_1715
fig/66666666.60966.peg.1844	CDS	1721245	1721081	-1	-	165	hypothetical protein	-none-	D23_1c1845 Neut_1716
fig/66666666.60966.peg.1845	CDS	1721933	1721268	-2	-	666	Mobile element protein	-none-	D23_1c1846 Neut_1717
fig/66666666.60966.peg.1846	CDS	1722745	1722137	-1	-	609	Aldehyde dehydrogenase (EC 1.2.1.3)	Glycerolipid and Glycerophospholipid Metabolism in Bacteria; Methylglyoxal Metabolism; Methylglyoxal Metabolism; Pyruvate metabolism II: acetyl-CoA, acetogenesis from pyruvate	D23_1c1847 Neut_0700
fig/66666666.60966.peg.1847	CDS	1722896	1723189	2	+	294	Mobile element protein	-none-	D23_1c1848 Neut_1719
fig/66666666.60966.peg.1848	CDS	1723288	1724166	1	+	879	Mobile element protein	-none-	D23_1c1849 Neut_1720

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld	C91 Alias
fig/66666666.60966.peg.1849	CDS	1724724	1724179	-3	-	546	Aldehyde dehydrogenase (EC 1.2.1.3)	Glycerolipid and Glycerophospholipid Metabolism in Bacteria; Methylglyoxal Metabolism; Methylglyoxal Metabolism; Pyruvate metabolism II: acetyl-CoA, acetogenesis from pyruvate	D23_1c1850	Neut_0700
fig/66666666.60966.peg.1850	CDS	1726242	1724995	-3	-	1248	Mobile element protein	-none-	D23_1c1851	Neut_0357
fig/66666666.60966.peg.1851	CDS	1727368	1727126	-1	-	243	Mobile element protein	-none-	D23_1c1854	Neut_2190
fig/66666666.60966.peg.1852	CDS	1727480	1727641	2	+	162	hypothetical protein	-none-	D23_1c1855	NA
fig/66666666.60966.peg.1853	CDS	1727757	1727873	3	+	117	hypothetical protein	-none-	D23_1c1856	NA
fig/66666666.60966.peg.1854	CDS	1727880	1728818	3	+	939	Major facilitator family transporter	-none-	D23_1c1857	NA
fig/66666666.60966.peg.1855	CDS	1728936	1730123	3	+	1188	Cytosine deaminase (EC 3.5.4.1)	CBSS-326442.4.peg.1852; Creatine and Creatinine Degradation; pyrimidine conversions	D23_1c1858	Neut_1722
fig/66666666.60966.peg.1857	CDS	1730427	1730624	3	+	198	Mobile element protein	-none-	D23_1c1859	Neut_1748
fig/66666666.60966.peg.1858	CDS	1730716	1730937	1	+	222	Mobile element protein	-none-	D23_1c1860	Neut_1747
fig/66666666.60966.peg.1859	CDS	1731165	1733414	3	+	2250	Lead, cadmium, zinc and mercury transporting ATPase (EC 3.6.3.3) (EC 3.6.3.5); Copper-translocating P-type ATPase (EC 3.6.3.4)	Copper Transport System; Copper homeostasis	D23_1c1862	Neut_1724
fig/66666666.60966.peg.1860	CDS	1733755	1733946	1	+	192	hypothetical protein	-none-	D23_1c1863	Neut_1734
fig/66666666.60966.peg.1861	CDS	1734020	1735762	2	+	1743	Asparagine synthetase [glutamine-hydrolyzing] (EC 6.3.5.4)	Cyanophycin Metabolism; Glutamate and Aspartate uptake in Bacteria; Glutamine, Glutamate, Aspartate and Asparagine Biosynthesis	D23_1c1864	Neut_1735
fig/66666666.60966.peg.1862	CDS	1735929	1737425	3	+	1497	major facilitator superfamily MFS_1	-none-	D23_1c1865	Neut_1736
fig/66666666.60966.peg.1863	CDS	1737575	1737841	2	+	267	Mobile element protein	-none-	D23_1c1866	NA
fig/66666666.60966.peg.1864	CDS	1737835	1737975	1	+	141	Mobile element protein	-none-	D23_1c1867	NA
fig/66666666.60966.peg.1865	CDS	1738803	1738006	-3	-	798	Mobile element protein	-none-	D23_1c1868	Neut_1888
fig/66666666.60966.peg.1866	CDS	1739186	1738917	-2	-	270	Mobile element protein	-none-	D23_1c1869	Neut_2500
fig/66666666.60966.peg.1868	CDS	1740582	1739374	-3	-	1209	hypothetical protein	-none-	D23_1c1870	Neut_1740

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld	C91 Alias
fig/66666666.60966.peg.1869	CDS	1742717	1740588	-2	-	2130	Ferrichrome-iron receptor	-none-	D23_1c1871	Neut_1741
fig/66666666.60966.peg.1870	CDS	1744008	1742806	-3	-	1203	Vibrioferin decarboxylase protein PvsE	-none-	D23_1c1872	Neut_1742
fig/66666666.60966.peg.1871	CDS	1745819	1744005	-2	-	1815	Vibrioferin amide bond forming protein PvsD @ Siderophore synthetase superfamily, group A	-none-	D23_1c1873	Neut_1743
fig/66666666.60966.peg.1872	CDS	1747001	1745829	-2	-	1173	Vibrioferin membrane-spanning transport protein PvsC	-none-	D23_1c1874	Neut_1744
fig/66666666.60966.peg.1873	CDS	1748855	1747044	-2	-	1812	Anthrachelin biosynthesis protein AsbB @ Siderophore synthetase superfamily, group C @ Siderophore synthetase component, ligase	-none-	D23_1c1875	Neut_1745
fig/66666666.60966.peg.1874	CDS	1749327	1749151	-3	-	177	Mobile element protein	-none-	D23_1c1876	Neut_1747
fig/66666666.60966.peg.1876	CDS	1750647	1749685	-3	-	963	Mobile element protein	-none-	D23_1c1878	Neut_1278
fig/66666666.60966.peg.1877	CDS	1751012	1752274	2	+	1263	hypothetical protein	-none-	D23_1c1879	Neut_1749
fig/66666666.60966.peg.1878	CDS	1752303	1753133	3	+	831	Potassium efflux system KefA protein/Small-conductance mechanosensitive channel	Potassium homeostasis	D23_1c1880	NA
fig/66666666.60966.peg.1879	CDS	1753722	1753153	-3	-	570	hypothetical protein	-none-	D23_1c1881	NA
fig/66666666.60966.peg.1880	CDS	1753823	1754629	2	+	807	Glutamate racemase (EC 5.1.1.3)	Glutamine, Glutamate, Aspartate and Asparagine Biosynthesis; Peptidoglycan Biosynthesis	D23_1c1882	Neut_1752
fig/66666666.60966.peg.1881	CDS	1754733	1755689	3	+	957	Universal stress protein	-none-	D23_1c1883	Neut_1753
fig/66666666.60966.peg.1882	CDS	1756978	1755725	-1	-	1254	DNA repair protein RadA	DNA repair, bacterial; Proteolysis in bacteria, ATP-dependent	D23_1c1884	Neut_1754
fig/66666666.60966.peg.1883	CDS	1757073	1757201	3	+	129	hypothetical protein	-none-	D23_1c1885	NA
fig/66666666.60966.peg.1884	CDS	1758757	1757372	-1	-	1386	L-serine dehydratase (EC 4.3.1.17)	Glycine and Serine Utilization; Pyruvate Alanine Serine Interconversions	D23_1c1887	Neut_1760
fig/66666666.60966.peg.1886	CDS	1759312	1760013	1	+	702	Serine protease precursor MucD/AlgY associated with sigma factor RpoE	Transcription initiation, bacterial sigma factors	D23_1c1890	Neut_1761

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld	C91 Alias
fig/66666666.60966.peg.1887	CDS	1760622	1761584	3	+	963	Mobile element protein	-none- Cell wall related cluster	D23_1c1892	Neut_1278
fig/66666666.60966.peg.1888	CDS	1761643	1762524	1	+	882	FIG071646: Sugar transferase		D23_1c1893	Neut_1762
fig/66666666.60966.peg.1889	CDS	1762574	1764655	2	+	2082	Sensory transduction histidine kinases	-none-	D23_1c1894	Neut_1763
fig/66666666.60966.peg.1890	CDS	1764739	1766601	1	+	1863	Lipid A export ATP-binding/permease protein MsbA	-none-	D23_1c1895	Neut_1764
fig/66666666.60966.peg.1891	CDS	1769731	1766636	-1	-	3096	Cobalt-zinc-cadmium resistance protein CzeA; Cation efflux system	Cobalt-zinc-cadmium resistance; Cobalt-zinc-cadmium resistance	D23_1c1896	Neut_1765
fig/66666666.60966.peg.1892	CDS	1770897	1769734	-3	-	1164	Cobalt/zinc/cadmium efflux RND transporter, protein CusA	Cobalt-zinc-cadmium resistance	D23_1c1897	Neut_1766
fig/66666666.60966.peg.1893	CDS	1772112	1770907	-3	-	1206	Heavy metal RND efflux protein, CzeB family	Cobalt-zinc-cadmium resistance	D23_1c1898	Neut_1767
fig/66666666.60966.peg.1894	CDS	1773163	1772504	-1	-	660	FIG00859115; hypothetical protein	-none-	D23_1c1899	Neut_1768
fig/66666666.60966.peg.1895	CDS	1773250	1773996	1	+	747	Glycerophosphoryl diester phosphodiesterase (EC 3.1.4.46)	CBSS-1762994.peg.1996A; Glycerol and Glycerol-3-phosphate Uptake and Utilization	D23_1c1900	Neut_1769
fig/66666666.60966.peg.1896	CDS	1774006	1775181	1	+	1176	Aerobic glycerol-3-phosphate dehydrogenase (EC 1.1.5.3)	Glycerol and Glycerol-3-phosphate Uptake and Utilization; Glycerolipid and Glycerophospholipid Metabolism in Bacteria; Respiratory dehydrogenases 1	D23_1c1901	Neut_1770
fig/66666666.60966.peg.1897	CDS	1775543	1775211	-2	-	333	FIG00859262; hypothetical protein	-none-	D23_1c1902	Neut_1771
fig/66666666.60966.peg.1898	CDS	1776076	1775636	-1	-	441	FIG00859309; hypothetical protein	-none-	D23_1c1903	Neut_1772
fig/66666666.60966.peg.1899	CDS	1777538	1776078	-2	-	1461	Dihydrolipouamide dehydrogenase of 2-oxoglutarate dehydrogenase (EC 1.8.1.4)	Dehydrogenase complexes; TCA Cycle	D23_1c1904	Neut_1773
fig/66666666.60966.peg.1900	CDS	1778658	1777612	-3	-	1047	Beta N-acetylglucosaminidase (EC 3.2.1.52)	Murein Hydrolases; Recycling of Peptidoglycan Amino Sugars	D23_1c1905	Neut_1774

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld	C91 Alias
fig/66666666.60966.peg.1901	CDS	1779038	1778661	-2	-	378	Holo-[acyl-carrier protein] synthase (EC 2.7.8.7)	Fatty Acid Biosynthesis FASII	D23_1c1906	Neut_1775
fig/66666666.60966.peg.1902	CDS	1779760	1779035	-1	-	726	Pyridoxine 5'-phosphate synthase (EC 2.6.99.2)	Pyridoxin (Vitamin B6) Biosynthesis	D23_1c1907	Neut_1776
fig/66666666.60966.peg.1903	CDS	1780768	1779878	-1	-	891	GTP-binding protein Era	Bacterial Cell Division; Glycyl-tRNA synthetase containing cluster; Universal GTPases	D23_1c1908	Neut_1777
fig/66666666.60966.peg.1904	CDS	1781583	1780846	-3	-	738	Ribonuclease III (EC 3.1.26.3)	RNA processing and degradation, bacterial	D23_1c1909	Neut_1778
fig/66666666.60966.peg.1905	CDS	1781963	1781580	-2	-	384	possible transmembrane protein	-none-	D23_1c1910	Neut_1779
fig/66666666.60966.peg.1906	CDS	1782802	1781999	-1	-	804	Signal peptidase I (EC 3.4.21.89)	Signal peptidase	D23_1c1911	Neut_1780
fig/66666666.60966.peg.1907	CDS	1784670	1782874	-3	-	1797	Translation elongation factor LepA	Heat shock dnaK gene cluster extended; Translation elongation factors bacterial; Universal GTPases	D23_1c1912	Neut_1781
fig/66666666.60966.peg.1908	CDS	1784806	1784651	-1	-	156	hypothetical protein	-none-	D23_1c1913	NA
fig/66666666.60966.peg.1909	CDS	1785160	1784972	-1	-	189	COGs COG0526	-none-	D23_1c1914	Neut_1782
fig/66666666.60966.peg.1910	CDS	1786580	1785222	-2	-	1359	Serine protease precursor MucD/AlgY associated with sigma factor RpoE	Transcription initiation, bacterial sigma factors	D23_1c1915	Neut_1783
fig/66666666.60966.peg.1911	CDS	1787446	1786871	-1	-	576	InterPro IPR001687 COGs COG3073	-none-	D23_1c1916	Neut_1784
fig/66666666.60966.peg.1912	CDS	1788062	1787460	-2	-	603	RNA polymerase sigma factor RpoE	Transcription initiation, bacterial sigma factors	D23_1c1917	Neut_1785
fig/66666666.60966.peg.1913	CDS	1789111	1788260	-1	-	852	Magnesium and cobalt efflux protein CorC	CBSS- 56780.10.peg.1536; Copper homeostasis; copper tolerance; Glycyl-tRNA synthetase containing cluster; Magnesium transport; tRNA-methylthiotransferase containing cluster	D23_1c1919	Neut_1786
fig/66666666.60966.peg.1914	CDS	1789590	1789165	-3	-	426	Metal-dependent hydrolase YbeY, involved in rRNA and/or	CBSS- 56780.10.peg.1536; Glycyl-tRNA	D23_1c1920	Neut_1787

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 GbkId	C91 Alias
fig/66666666.60966.peg.1915	CDS	1790629	1789619	-1	-	1011	ribosome maturation and assembly	synthetase containing cluster; RNA-methylthiotransferase containing cluster	D23_1c1921	Neut_1788
fig/66666666.60966.peg.1916	CDS	1792022	1790691	-2	-	1332	Phosphate starvation-inducible ATPase PhoH with RNA binding motif tRNA-(6)A37 methylthiotransferase	Methylthiotransferases; RNA-methylthiotransferase containing cluster; RNA modification Bacteria; RNA processing -none-	D23_1c1922	Neut_1789
fig/66666666.60966.peg.1917	CDS	1793040	1792321	-3	-	720	Cytochrome c-type protein TorY		D23_1c1923	Neut_1790
fig/66666666.60966.peg.1918	CDS	1793750	1793043	-2	-	708	Cytochrome c family protein	-none-	D23_1c1924	Neut_2333
fig/66666666.60966.peg.1919	CDS	1794933	1793812	-3	-	1122	FIG00859557; hypothetical protein	-none-	D23_1c1925	Neut_1792
fig/66666666.60966.peg.1920	CDS	1796642	1794930	-2	-	1713	Hydroxylamine oxidoreductase	-none-	D23_1c1926	Neut_2335
fig/66666666.60966.peg.1921	CDS	1801076	1796862	-2	-	4215	precursor (EC 1.7.3.4) DNA-directed RNA polymerase beta' subunit (EC 2.7.7.6)	Mycobacterium virulence operon involved in DNA transcription; RNA polymerase bacterial	D23_1c1927	Neut_1794
fig/66666666.60966.peg.1922	CDS	1805314	1801235	-1	-	4080	DNA-directed RNA polymerase beta subunit (EC 2.7.7.6)	Mycobacterium virulence operon involved in DNA transcription; RNA polymerase bacterial	D23_1c1928	Neut_1795
fig/66666666.60966.peg.1923	CDS	1806051	1805680	-3	-	372	LSU ribosomal protein L7/L12 (P1/P2)		D23_1c1929	Neut_1796
fig/66666666.60966.peg.1924	CDS	1806648	1806133	-3	-	516	LSU ribosomal protein L10p (P0)	LSU ribosomal proteins cluster	D23_1c1930	Neut_1797
fig/66666666.60966.peg.1925	CDS	1807793	1807098	-2	-	696	LSU ribosomal protein Lip (L10Ae)	LSU ribosomal proteins cluster	D23_1c1931	Neut_1798
fig/66666666.60966.peg.1926	CDS	1808121	1807795	-3	-	327	LSU ribosomal protein L11p (L12e)	LSU ribosomal proteins cluster	D23_1c1932	Neut_1799
fig/66666666.60966.peg.1927	CDS	1808877	1808344	-3	-	534	Transcription antitermination protein NusG	LSU ribosomal proteins cluster; Transcription factors bacterial	D23_1c1934	Neut_1800
fig/66666666.60966.peg.1928	CDS	1809240	1808896	-3	-	345	Preprotein translocase subunit SecE (TC 3.A.5.1.1)	LSU ribosomal proteins cluster	D23_1c1935	Neut_1801

Annotation of genes in SEQ ID NO: 1.									
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	C91 Alias
fig 66666666.609666.peg.1929	CDS	1810674	1809484	-3	-	1191	Translation elongation factor Tu	<i>Mycobacterium</i> virulence operon involved in protein synthesis (SSU ribosomal proteins); Translation elongation factors bacterial; Universal GTPases	D23_1c1937 Neut_1802
fig 66666666.609666.peg.1930	CDS	1813382	1811220	-2	-	2163	Type IV pilus biogenesis	-none- -none- -none- -none- -none-	Neut_1803
fig 66666666.609666.peg.1931	CDS	1813906	1813379	-1	-	528	Type IV pilus biogenesis protein PilP		Neut_1804
fig 66666666.609666.peg.1932	CDS	1814553	1813903	-3	-	651	Type IV pilus biogenesis protein PilO		Neut_1805
fig 66666666.609666.peg.1933	CDS	1815161	1814550	-2	-	612	Type IV pilus biogenesis protein PilN		Neut_1806
fig 66666666.609666.peg.1934	CDS	1816207	1815158	-1	-	1050	Type IV pilus biogenesis protein PilM		Neut_1807
fig 66666666.609666.peg.1936	CDS	1816441	1818753	1	+	2313	Multimodular transpeptidase-transglycosylase (EC 2.4.1.129) (EC 3.4.--.--)	Peptidoglycan Biosynthesis	Neut_1808
fig 66666666.609666.peg.1937	CDS	1819961	1819041	-2	-	921	Deacetylases, including yeast histone deacetylase and acetoin utilization protein	-none-	Neut_1809
fig 66666666.609666.peg.1939	CDS	1820319	1820170	-3	-	150	hypothetical protein	-none-	Neut_1810
fig 66666666.609666.peg.1941	CDS	1820799	1820638	-3	-	162	Addition module antidiote protein	-none-	Neut_1811
fig 66666666.609666.peg.1942	CDS	1821096	1820914	-3	-	183	hypothetical protein	-none-	Neut_1812
fig 66666666.609666.peg.1943	CDS	1821848	1821985	2	+	138	Prevent host death protein, Phd antitoxin	Phd-Doc, YdeE-YdeD toxin-antitoxin (programmed cell death) systems	NA
fig 66666666.609666.peg.1944	CDS	1821982	1822278	1	+	297	Death on curing protein, Doc toxin	Phd-Doc, YdeE-YdeD toxin-antitoxin (programmed cell death) systems	NA
fig 66666666.609666.peg.1945	CDS	1822608	1822423	-3	-	186	hypothetical protein	-none-	Neut_1816
fig 66666666.609666.peg.1946	CDS	1822933	1822688	-1	-	246	hypothetical protein	-none-	Neut_1817
fig 66666666.609666.peg.1947	CDS	1823491	1823150	-1	-	342	Predicted transcriptional regulator	-none-	NA
fig 66666666.609666.peg.1948	CDS	1823708	1823484	-2	-	225	Phage-related protein	-none-	NA
fig 66666666.609666.peg.1952	CDS	1824892	1824704	-1	-	189	hypothetical protein	-none-	NA
fig 66666666.609666.peg.1953	CDS	1825371	1825036	-3	-	336	Mobile element protein	-none-	Neut_1624
fig 66666666.609666.peg.1954	CDS	1825828	1825352	-1	-	477	Mobile element protein	-none-	Neut_1888
fig 66666666.609666.peg.1955	CDS	1826127	1825942	-3	-	186	Mobile element protein	-none-	Neut_2500

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.									
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	C91 Alias
fig/66666666.60966.peg.1956	CDS	1826660	1826472	-2	-	189	Mobile element protein	-none-	Neut_1821
fig/66666666.60966.peg.1958	CDS	1827279	1826920	-3	-	360	Flagellin protein FlaG	Flagellum	Neut_1822
fig/66666666.60966.peg.1959	CDS	1827931	1827644	-1	-	288	Exonuclease ABC, C subunit-like	-none-	Neut_1823
fig/66666666.60966.peg.1960	CDS	1829557	1828115	-1	-	1443	Flagellin protein FlaB	Flagellum; Flagellum in <i>Campylobacter</i>	Neut_1824
fig/66666666.60966.peg.1961	CDS	1830108	1830230	3	+	123	hypothetical protein	-none-	NA
fig/66666666.60966.peg.1962	CDS	1831067	1830276	-2	-	792	Mobile element protein	-none-	Neut_1888
fig/66666666.60966.peg.1963	CDS	1831366	1831181	-1	-	186	Mobile element protein	-none-	Neut_2500
fig/66666666.60966.peg.1964	CDS	1831759	1831616	-1	-	144	hypothetical protein	-none-	NA
fig/66666666.60966.peg.1965	CDS	1832465	1831749	-2	-	717	hypothetical protein	-none-	Neut_1827
fig/66666666.60966.peg.1967	CDS	1835117	1833291	-2	-	1827	DNA mismatch repair protein MutL	DNA repair, bacterial MutL-MutS system	Neut_1828
fig/66666666.60966.peg.1968	CDS	1835287	1835946	1	+	660	Ribose 5-phosphate isomerase A (EC 5.3.1.6)	Calvin-Benson cycle; D-ribose utilization; Pentose phosphate pathway	Neut_1829
fig/66666666.60966.peg.1969	CDS	1836022	1836150	1	+	129	hypothetical protein	-none-	NA
fig/66666666.60966.peg.1970	CDS	1836140	1836859	2	+	720	Phosphate transport system regulatory protein PhoU	High affinity phosphate transporter and control of PHO regulon; Phosphate metabolism	Neut_1830
fig/66666666.60966.peg.1971	CDS	1836819	1838420	3	+	1602	Exopolysphatase (EC 3.6.1.11)	Phosphate metabolism; Polyphosphate	Neut_1831
fig/66666666.60966.peg.1972	CDS	1838878	1838429	-1	-	450	Type IV pilus biogenesis protein PilE	-none-	Neut_1832
fig/66666666.60966.peg.1973	CDS	1842290	1838922	-2	-	3369	Type IV fimbrial biogenesis protein PilY1	-none-	Neut_1833
fig/66666666.60966.peg.1974	CDS	1843213	1842362	-1	-	852	Type IV fimbrial biogenesis protein PilX	-none-	Neut_1834
fig/66666666.60966.peg.1975	CDS	1844303	1843239	-2	-	1065	Type IV fimbrial biogenesis protein PilW	-none-	Neut_1835
fig/66666666.60966.peg.1976	CDS	1844806	1844321	-1	-	486	Type IV fimbrial biogenesis protein PilV	-none-	Neut_1836
fig/66666666.60966.peg.1977	CDS	1845339	1844830	-3	-	510	Type IV fimbrial biogenesis protein FimT	-none-	Neut_1837
fig/66666666.60966.peg.1978	CDS	1845616	1845783	1	+	168	hypothetical protein	-none-	NA
fig/66666666.60966.peg.1979	CDS	1846400	1845768	-2	-	633	DNA-binding response regulator, LuxR family	-none-	Neut_1839
fig/66666666.60966.peg.1980	CDS	1847948	1846452	-2	-	1497	Sensory box histidine kinase/response regulator	-none-	Neut_1840
fig/66666666.60966.peg.1981	CDS	1848084	1848206	3	+	123	hypothetical protein	-none-	NA
fig/66666666.60966.peg.1982	CDS	1850161	1848614	-1	-	1548	pilin glycosylation enzyme, putative	-none-	Neut_1841

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld	C91 Alias
fig/66666666.60966.peg.1985	CDS	1852587	1850815	-3	-	1773	Gamma-glutamyltranspeptidase (EC 2.3.2.2)	Glutathione; Biosynthesis and gamma-glutamyl cycle	D23_1c1989	Neut_1843
fig/66666666.60966.peg.1986	CDS	1853865	1852903	-3	-	963	Mobile element protein	-non-	D23_1c1990	Neut_1746
fig/66666666.60966.peg.1987	CDS	1854432	1853923	-3	-	510	putative transmembrane protein	-non-	D23_1c1991	Neut_1845
fig/66666666.60966.peg.1988	CDS	1855175	1854606	-2	-	570	hypothetical protein	-non-	D23_1c1992	Neut_1849
fig/66666666.60966.peg.1989	CDS	1855822	1855421	-1	-	402	Glyoxalase family protein	-non-	D23_1c1993	Neut_1850
fig/66666666.60966.peg.1990	CDS	1855948	1855835	-1	-	114	hypothetical protein	-non-	D23_1c1994	NA
fig/66666666.60966.peg.1991	CDS	1856595	1856026	-3	-	570	hypothetical protein	-non-	D23_1c1995	Neut_1851
fig/66666666.60966.peg.1992	CDS	1856572	1856703	1	+	132	hypothetical protein	-non-	D23_1c1996	NA
fig/66666666.60966.peg.1993	CDS	1858627	1856756	-1	-	1872	hypothetical protein	-non-	D23_1c1997	Neut_1853
fig/66666666.60966.peg.1994	CDS	1860549	1858642	-3	-	1908	hypothetical protein	-non-	D23_1c1998	Neut_1854
fig/66666666.60966.peg.1995	CDS	1860536	1860667	2	+	132	hypothetical protein	-non-	D23_1c1999	NA
fig/66666666.60966.peg.1996	CDS	1861687	1860761	-1	-	927	Expressed protein precursor	-non-	D23_1c2000	Neut_1857
fig/66666666.60966.peg.1997	CDS	1862145	1861684	-3	-	462	hypothetical protein	-non-	D23_1c2001	Neut_1858
fig/66666666.60966.peg.1998	CDS	1865471	1862334	-2	-	3138	Proline dehydrogenase (EC 1.5.99.8) (Proline oxidase)/Delta-1-pyrroline-5-carboxylate dehydrogenase (EC 1.5.1.12)	Proline, 4-hydroxyproline uptake and utilization; Respiratory dehydrogenases 1	D23_1c2002	Neut_1859
fig/66666666.60966.peg.2000	CDS	1865859	1865701	-3	-	159	hypothetical protein	-non-	D23_1c2003	Neut_1860
fig/66666666.60966.peg.2001	CDS	1866618	1866328	-3	-	291	hypothetical protein	-non-	D23_1c2004	NA
fig/66666666.60966.peg.2002	CDS	1866574	1866861	1	+	288	Probable transmembrane protein	-non-	D23_1c2005	Neut_1861
fig/66666666.60966.peg.2004	CDS	1867176	1867955	3	+	780	hypothetical protein	-non-	D23_1c2006	Neut_1863
fig/66666666.60966.peg.2005	CDS	1870128	1868077	-3	-	2052	Serine peptidase	-non-	D23_1c2007	Neut_1864
fig/66666666.60966.peg.2006	CDS	1870373	1870546	2	+	174	hypothetical protein	-non-	D23_1c2008	NA
fig/66666666.60966.peg.2007	CDS	1871555	1870827	-2	-	729	1-acyl-sn-glycerol-3-phosphate acyltransferase (EC 2.3.1.51)	Glycerolipid and Glycerophospholipid Metabolism in Bacteria	D23_1c2009	Neut_1866
fig/66666666.60966.peg.2008	CDS	1872097	1871555	-1	-	543	Histidinol-phosphatase (EC 3.1.3.15)	Histidine Biosynthesis	D23_1c2010	Neut_1867
fig/66666666.60966.peg.2009	CDS	1874271	1872124	-3	-	2148	Glycyl-tRNA synthetase beta chain (EC 6.1.1.14)	Glycyl-tRNA synthetase; Glycyl-tRNA synthetase containing cluster; tRNA aminoacylation, Gly	D23_1c2011	Neut_1868
fig/66666666.60966.peg.2010	CDS	1875191	1874268	-2	-	924	Glycyl-tRNA synthetase alpha chain (EC 6.1.1.14)	Glycyl-tRNA synthetase; Glycyl-tRNA synthetase containing cluster; tRNA aminoacylation, Gly	D23_1c2012	Neut_1869

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 GbkId	C91 Alias
fig/66666666.60966.peg.2011	CDS	1876717	1875224	-1	-	1494	Apolipoprotein N-acyltransferase (EC 2.3.1.—)Copper homeostasis protein CutE	Copper homeostasis: copper tolerance; Lipoprotein Biosynthesis; RNA-methylthiotransferase containing cluster; RNA-methylthiotransferase containing cluster -none-	D23_1c2013	Neut_1870
fig/66666666.60966.peg.2012	CDS	1877111	1876773	-2	-	339	FIG00859587: hypothetical protein		D23_1c2014	Neut_1871
fig/66666666.60966.peg.2013	CDS	1877265	1877122	-3	-	144	hypothetical protein		D23_1c2015	NA
fig/66666666.60966.peg.2014	CDS	1877596	1879305	1	+	1710	Multi-copper oxidase	Copper homeostasis	D23_1c2016	Neut_1872
fig/66666666.60966.peg.2015	CDS	1879305	1880399	3	+	1095	Zinc ABC transporter, periplasmic-binding protein ZnuA	-none-	D23_1c2017	Neut_1873
fig/66666666.60966.peg.2016	CDS	1881044	1880847	-2	-	198	hypothetical protein		D23_1c2018	NA
fig/66666666.60966.peg.2017	CDS	1881486	1881962	3	+	477	Cytochrome c oxidase (B(O/a)3-type) chain II (EC 1.9.3.1)	Terminal cytochrome C oxidases	D23_1c2020	Neut_1874
fig/66666666.60966.peg.2018	CDS	1882000	1883478	1	+	1479	Cytochrome c oxidase (B(O/a)3-type) chain I (EC 1.9.3.1)	Terminal cytochrome C oxidases	D23_1c2021	Neut_1875
fig/66666666.60966.peg.2019	CDS	1883574	1884203	3	+	630	Cytochrome oxidase biogenesis protein ScoI/ScoC/PrrC, putative copper metallochaperone	Biogenesis of cytochrome c oxidases	D23_1c2022	Neut_1876
fig/66666666.60966.peg.2020	CDS	1884187	1884756	1	+	570	hypothetical cytochrome oxidase associated membrane protein	-none-	D23_1c2023	Neut_1877
fig/66666666.60966.peg.2021	CDS	1885726	1884809	-1	-	918	Nitrite transporter from formate/nitrite family	-none-	D23_1c2024	Neut_1878
fig/66666666.60966.peg.2022	CDS	1887084	1886077	-3	-	1008	UDP-glucose 4-epimerase (EC 5.1.3.2)	CBSS-296591.1.peg.2330; N-linked Glycosylation in Bacteria; Rhamnose containing glycans; dTDP-rhamnose synthesis -none-	D23_1c2026	Neut_1879
fig/66666666.60966.peg.2023	CDS	1888047	1887160	-3	-	888	Glucose-1-phosphate thymidyltransferase (EC 2.7.7.24)	Glycerate metabolism; Photorespiration (oxidative C2 cycle)	D23_1c2027	Neut_1880
fig/66666666.60966.peg.2024	CDS	1888279	1888148	-1	-	132	hypothetical protein		D23_1c2028	NA
fig/66666666.60966.peg.2025	CDS	1888278	1889198	3	+	921	2-hydroxy-3-oxopropionate reductase (EC 1.1.1.60)		D23_1c2029	Neut_1881

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 GbkId	C91 Alias
fig/66666666.60966.peg.2026	CDS	1889188	1890642	1	+	1455	Glycolate dehydrogenase (EC 1.1.99.14), subunit GlcD	Glycolate, glyoxylate interconversions; Photorespiration (oxidative C2 cycle)	D23_1c2030	Neut_1882
fig/66666666.60966.peg.2027	CDS	1890667	1891767	1	+	1101	Glycolate dehydrogenase (EC 1.1.99.14), FAD-binding subunit GlcE	Glycolate, glyoxylate interconversions; Photorespiration (oxidative C2 cycle)	D23_1c2031	Neut_1883
fig/66666666.60966.peg.2028	CDS	1891771	1893039	1	+	1269	Glycolate dehydrogenase (EC 1.1.99.14), iron-sulfur subunit GlcF	Glycolate, glyoxylate interconversions; Photorespiration (oxidative C2 cycle)	D23_1c2032	Neut_1884
fig/66666666.60966.peg.2029	CDS	1893781	1893065	-1	-	717	Putative predicted metal-dependent hydrolase	Restriction-Modification System	D23_1c2033	Neut_1885
fig/66666666.60966.peg.2030	CDS	1894582	1893806	-1	-	777	5'-methylthioadenosine nucleosidase (EC 3.2.2.16)/S-adenosylhomocysteine nucleosidase (EC 3.2.2.9)	Adenosyl nucleosidases; Adenosyl nucleosidases; CBSS-320388.3.peg.3759; CBSS-320388.3.peg.3759; Methionine Biosynthesis; Methionine Degradation; Polyamine Metabolism	D23_1c2034	Neut_1886
fig/66666666.60966.peg.2031	CDS	1896116	1894656	-2	-	1461	Exodeoxyribonuclease I (EC 3.1.11.1)	DNA Repair Base Excision	D23_1c2035	Neut_1887
fig/66666666.60966.peg.2032	CDS	1896438	1897133	3	+	696	FIG00657740: hypothetical protein	-none-	D23_1c2036	Neut_1890
fig/66666666.60966.peg.2033	CDS	1898282	1897428	-2	-	855	5,10-methylenetetrahydrofolate reductase (EC 1.5.1.20)	5-FCL-like protein; Methionine Biosynthesis; One-carbon metabolism by tetrahydropterines	D23_1c2038	Neut_1891
fig/66666666.60966.peg.2034	CDS	1898327	1898494	2	+	168	hypothetical protein	-none-	D23_1c2039	NA
fig/66666666.60966.peg.2035	CDS	1899999	1898563	-3	-	1437	Adenosylhomocysteine synthetase (EC 3.3.1.1)	Methionine Biosynthesis; Methionine Degradation	D23_1c2041	Neut_1892
fig/66666666.60966.peg.2036	CDS	1901244	1900135	-3	-	1110	S-adenosylmethionine synthetase (EC 2.5.1.6)	Methionine Biosynthesis; Methionine Degradation	D23_1c2042	Neut_1893
fig/66666666.60966.peg.2037	CDS	1901543	1902289	2	+	747	Short chain dehydrogenase	-none-	D23_1c2043	Neut_1894

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 GbkId	C91 Alias
fig/66666666.60966.peg.2038	CDS	1902336	1902812	3	+	477	ATPase YjeF, predicted to have essential role in cell wall biosynthesis	-none-	D23_1c2044	Neut_1895
fig/66666666.60966.peg.2039	CDS	1903046	1904116	2	+	1071	N-acetylmuramoyl-L-alanine amidase (EC 3.5.1.28)	Murein Hydrolases; Recycling of Peptidoglycan Amino Acids; Zinc regulated enzymes	D23_1c2045	Neut_1896
fig/66666666.60966.peg.2040	CDS	1904922	1904170	-3	-	753	FIG00859340: hypothetical protein	-none-	D23_1c2046	Neut_1897
fig/66666666.60966.peg.2041	CDS	1905860	1904919	-2	-	942	Ribosomal protein L11 methyltransferase (EC 2.1.1.—)	Heat shock dnaK gene cluster extended; Ribosome biogenesis bacterial	D23_1c2047	Neut_1898
fig/66666666.60966.peg.2042	CDS	1907254	1905896	-1	-	1359	Biotin carboxylase of acetyl-CoA carboxylase (EC 6.3.4.14)	Fatty Acid Biosynthesis FASII	D23_1c2048	Neut_1899
fig/66666666.60966.peg.2043	CDS	1907784	1907326	-3	-	459	Biotin carboxyl carrier protein of acetyl-CoA carboxylase	Fatty Acid Biosynthesis FASII	D23_1c2049	Neut_1900
fig/66666666.60966.peg.2045	CDS	1908294	1907989	-3	-	306	3-dehydroquininate dehydratase II (EC 4.2.1.10)	Chorismate Synthesis; Common Pathway For Synthesis of Aromatic Compounds (DAHPSynthase to chorismate); Quinate degradation	D23_1c2050	Neut_1901
fig/66666666.60966.peg.2047	CDS	1908606	1909715	3	+	1110	Glycine oxidase ThiO (EC 1.4.3.19)	Thiamin biosynthesis	D23_1c2051	Neut_1902
fig/66666666.60966.peg.2048	CDS	1910693	1909722	-2	-	972	4-hydroxy-3-methylbut-2-enyl diphosphate reductase (EC 1.17.1.2)	Isoprenoid Biosynthesis; Nonmevalonate Branch of Isoprenoid Biosynthesis	D23_1c2052	Neut_1903
fig/66666666.60966.peg.2049	CDS	1910983	1911267	1	+	285	FIG00858797: hypothetical protein	-none-	D23_1c2053	Neut_1904
fig/66666666.60966.peg.2050	CDS	1911307	1912386	1	+	1080	Histidinol-phosphate aminotransferase (EC 2.6.1.9)	Histidine Biosynthesis	D23_1c2054	Neut_1905
fig/66666666.60966.peg.2051	CDS	1912429	1913016	1	+	588	Imidazoleglycerol-phosphate dehydratase (EC 4.2.1.19)	Histidine Biosynthesis	D23_1c2055	Neut_1906
fig/66666666.60966.peg.2052	CDS	1913079	1913687	3	+	609	Imidazole glycerol phosphate synthase amidotransferase subunit (EC 2.4.2.—)	Histidine Biosynthesis	D23_1c2056	Neut_1907
fig/66666666.60966.peg.2053	CDS	1913772	1914518	3	+	747	Phosphoribosylformimino-5-aminimidazole	Chorismate: Intermediate for	D23_1c2057	Neut_1908

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld	C91 Alias
fig/66666666.60966.peg.2054	CDS	1914604	1915380	1	+	777	carboxamide ribotide isomerase (EC 5.3.1.16)	synthesis of Tryptophan, PAPA antibiotics, PABA, 3-hydroxyanthranilate and more; Histidine Biosynthesis	D23_1c2058	Neut_1909
								Histidine Biosynthesis		
fig/66666666.60966.peg.2055	CDS	1915424	1915909	2	+	486	Imidazole glycerol phosphate synthase cyclase subunit (EC 4.1.3.—)	Histidine Biosynthesis; Zinc regulated enzymes	D23_1c2059	Neut_1910
								 Zinc regulated enzymes		
fig/66666666.60966.peg.2056	CDS	1915906	1916259	1	+	354	Phosphoribosyl-ATP pyrophosphatase (EC 3.6.1.31)	Histidine Biosynthesis; Riboflavin synthesis cluster	D23_1c2060	Neut_1911
								 Riboflavin synthesis cluster		
fig/66666666.60966.peg.2057	CDS	1916261	1916611	2	+	351	FIG146285; Diadenosine tetraphosphate (Ap4A) hydrolase and other HIT family hydrolases	-none-	D23_1c2061	Neut_1912
								-none-		
fig/66666666.60966.peg.2058	CDS	1916631	1916864	3	+	234	Twin-arginine translocation protein TatA	Cluster-based Subsystem Grouping Hypotheticals- perhaps Proteosome Related; Twin-arginine translocation system	D23_1c2062	Neut_1913
fig/66666666.60966.peg.2060	CDS	1916950	1917342	1	+	393	Twin-arginine translocation protein TatB	Twin-arginine translocation system	D23_1c2063	Neut_1914
fig/66666666.60966.peg.2061	CDS	1917433	1918209	1	+	777	Twin-arginine translocation protein TatC	Cluster-based Subsystem Grouping Hypotheticals- perhaps Proteosome Related; Twin-arginine translocation system	D23_1c2064	Neut_1915
fig/66666666.60966.peg.2062	CDS	1918395	1920518	3	+	2124	Outer membrane vitamin B12 receptor BtuB	-none-	D23_1c2065	Neut_1916
fig/66666666.60966.peg.2063	CDS	1920528	1921118	3	+	591	Optional hypothetical component of the B12 transporter BtuM	-none-	D23_1c2066	Neut_1917
fig/66666666.60966.peg.2064	CDS	1921118	1921726	2	+	609	CobT(yalam) adenosyltransferase (EC 2.5.1.17)	-none-	D23_1c2067	Neut_1918
fig/66666666.60966.peg.2065	CDS	1922713	1922826	1	+	114	hypothetical protein	-none-	D23_1c2069	NA
fig/66666666.60966.peg.2067	CDS	1923447	1924253	3	+	807	Cytochrome bd-type quinol oxidase, subunit 1	-none-	D23_1c2070	Neut_1920

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld	C91 Alias
fig/66666666.60966.peg.2068	CDS	1926393	1924288	-3	-	2106	Methionyl-tRNA synthetase (EC 6.1.1.10)	Scaffold proteins for [4Fe—4S] cluster assembly (MRP family); RNA	D23_1c2071	Neut_1921
fig/66666666.60966.peg.2069	CDS	1926506	1926390	-2	-	117	hypothetical protein	-non-	D23_1c2072	NA
fig/66666666.60966.peg.2070	CDS	1926499	1927584	1	+	1086	Scaffold protein for [4Fe—4S] cluster assembly ApbC, MRP-like	aminoacylation, Met	D23_1c2073	Neut_1922
fig/66666666.60966.peg.2071	CDS	1927630	1928163	1	+	534	Deoxycytidine triphosphate deaminase (EC 3.5.4.13)	[4Fe—4S] cluster assembly (MRP family)	D23_1c2074	Neut_1923
fig/66666666.60966.peg.2072	CDS	1928619	1928251	-3	-	369	COGs COG1917	pyrimidine conversions	D23_1c2075	Neut_1924
fig/66666666.60966.peg.2073	CDS	1929401	1928631	-2	-	771	Inner membrane protein	-non-	D23_1c2076	Neut_1925
fig/66666666.60966.peg.2074	CDS	1929596	1929889	2	+	294	Mobile element protein	-non-	D23_1c2077	Neut_1719
fig/66666666.60966.peg.2075	CDS	1929988	1930866	1	+	879	Mobile element protein	-non-	D23_1c2078	Neut_1720
fig/66666666.60966.peg.2076	CDS	1930961	1933699	2	+	2739	Ca ion P-type ATPase	-non-	D23_1c2079	Neut_1926
fig/66666666.60966.peg.2078	CDS	1934003	1934323	2	+	321	hypothetical protein	-non-	D23_1c2080	Neut_1927
fig/66666666.60966.peg.2079	CDS	1935608	1934517	-2	-	1092	Prophage Lp2 protein 6	-non-	D23_1c2081	Neut_1928
fig/66666666.60966.peg.2081	CDS	1935893	1936396	2	+	504	ABC transporter ATP-binding protein YycR	-non-	D23_1c2083	Neut_1936
fig/66666666.60966.peg.2083	CDS	1936587	1936880	3	+	294	Mobile element protein	-non-	D23_1c2085	Neut_1719
fig/66666666.60966.peg.2084	CDS	1936979	1937857	2	+	879	Mobile element protein	-non-	D23_1c2086	Neut_1720
fig/66666666.60966.peg.2085	CDS	1938082	1937936	-1	-	147	hypothetical protein	-non-	D23_1c2087	Neut_1937
fig/66666666.60966.peg.2086	CDS	1938637	1938188	-1	-	450	Ferric uptake regulation protein FUR	Bacterial RNA-metabolizing Zn-dependent hydrolases; Oxidative stress	D23_1c2088	Neut_1938
fig/66666666.60966.peg.2087	CDS	1938848	1939330	2	+	483	Outer membrane lipoprotein SmpA, a component of the essential YaeT outer-membrane protein assembly complex	 Lipopolysaccharide assembly	D23_1c2089	Neut_1939
fig/66666666.60966.peg.2088	CDS	1939327	1940133	1	+	807	Dihydrodipicolinate reductase (EC 1.3.1.26)	-non-	D23_1c2090	Neut_1940
fig/66666666.60966.peg.2089	CDS	1941855	1940347	-3	-	1509	ATPase	-non-	D23_1c2091	Neut_1941
fig/66666666.60966.peg.2090	CDS	1942209	1942087	-3	-	123	hypothetical protein	-non-	D23_1c2092	NA
fig/66666666.60966.peg.2091	CDS	1942220	1942459	2	+	240	Type I restriction-modification system, DNA-methyltransferase subunit M (EC 2.1.1.72)	Restriction-Modification System; Type I Restriction-Modification	D23_1c2093	Neut_1942
fig/66666666.60966.peg.2092	CDS	1942456	1942860	1	+	405	Cell filamentation protein fic	-non-	D23_1c2094	Neut_1943
fig/66666666.60966.peg.2093	CDS	1942823	1943278	2	+	456	Mobile element protein	-non-	D23_1c2095	Neut_2502
fig/66666666.60966.peg.2094	CDS	1943305	1944786	1	+	1482	Outer membrane	-non-	D23_1c2096	Neut_1945

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.									
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld C91 Alias
fig/66666666.60966.peg.2095	CDS	1944830	1946014	2	+	1185	component of tripartite multidrug resistance system	Multidrug Resistance Efflux Pumps	D23_1c2097 Neut_1946
fig/66666666.60966.peg.2096	CDS	1946018	1949131	2	+	3114	Membrane fusion protein of RND family multidrug efflux pump RND efflux system, inner membrane transporter CmeB	Multidrug Resistance Efflux Pumps	D23_1c2098 Neut_1947
fig/66666666.60966.peg.2097	CDS	1949252	1949386	2	+	135	Type I restriction-modification system, restriction subunit R (EC 3.1.21.3)	Restriction-Modification System; Type I Restriction-Modification	D23_1c2099 NA
fig/66666666.60966.peg.2098	CDS	1949446	1950246	1	+	801	Bis(5'-nucleosyl)-tetraphosphatase, symmetrical (EC 3.6.1.41)	EC49-61	D23_1c2100 Neut_1949
fig/66666666.60966.peg.2099	CDS	1951009	1950200	-1	-	810	1-acyl-sn-glycerol-3-phosphate acyltransferase (EC 2.3.1.51)	Glycerolipid and Glycerophospholipid Metabolism in Bacteria	D23_1c2101 Neut_1950
fig/66666666.60966.peg.2100	CDS	1952008	1951073	-1	-	936	InterPro IPR002173 COGs COG0524	-none-	D23_1c2103 Neut_1951
fig/66666666.60966.peg.2101	CDS	1953491	1952040	-2	-	1452	Glycine dehydrogenase [decarboxylating] (glycine cleavage system P2 protein) (EC 1.4.4.2)	Glycine and Serine Utilization; Glycine cleavage system; Photorespiration (oxidative C2 cycle)	D23_1c2104 Neut_1952
fig/66666666.60966.peg.2102	CDS	1954921	1953566	-1	-	1356	Glycine dehydrogenase [decarboxylating] (glycine cleavage system P1 protein) (EC 1.4.4.2)	Utilization; Glycine cleavage system; Photorespiration (oxidative C2 cycle)	D23_1c2105 Neut_1953
fig/66666666.60966.peg.2103	CDS	1955520	1955131	-3	-	390	Glycine cleavage system H protein	Glycine and Serine Utilization; Glycine cleavage system; Photorespiration (oxidative C2 cycle)	D23_1c2106 Neut_1954
fig/66666666.60966.peg.2104	CDS	1956682	1955591	-1	-	1092	Aminomethyltransferase (glycine cleavage system T protein) (EC 2.1.2.10)	CBSS-87626.3.peg.3639; Glycine and Serine Utilization; Glycine cleavage system; Photorespiration (oxidative C2 cycle)	D23_1c2107 Neut_1955
fig/66666666.60966.peg.2106	CDS	1957270	1957133	-1	-	138	hypothetical protein	-none-	D23_1c2108 NA
fig/66666666.60966.peg.2107	CDS	1958337	1957420	-3	-	918	Coproporphyrinogen III oxidase, aerobic (EC 1.3.3.3)	Heme and Siroheme Biosynthesis	D23_1c2109 Neut_1956

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld	C91 Alias
fig/66666666.60966.peg.2108	CDS	1958499	1959671	3	+	1173	Chorismate synthase (EC 4.2.3.5)	Chorismate Synthesis; Common Pathway For Synthesis of Aromatic Compounds (DAHHP synthase to chorismate)	D23_1c2110	Neut_1957
fig/66666666.60966.peg.2110	CDS	1960360	1961133	1	+	774	IncF plasmid conjugative transfer surface exclusion protein TraI	-none-	D23_1c2111	Neut_1959
fig/66666666.60966.peg.2111	CDS	1961191	1961517	1	+	327	hypothetical protein	-none-	D23_1c2112	NA
fig/66666666.60966.peg.2112	CDS	1961582	1962544	2	+	963	Mobile element protein	-none-	D23_1c2114	Neut_1862
fig/66666666.60966.peg.2113	CDS	1962848	1962729	-2	-	120	hypothetical protein	-none-	D23_1c2115	NA
fig/66666666.60966.peg.2115	CDS	1965546	1963441	-3	-	2106	Ferrichrome-iron receptor	-none-	D23_1c2116	Neut_1962
fig/66666666.60966.peg.2118	CDS	1967221	1966565	-1	-	657	Protein of unknown function DUF208	-none-	D23_1c2119	Neut_1964
fig/66666666.60966.peg.2119	CDS	1967389	1968672	1	+	1284	FIG00858634; hypothetical protein	-none-	D23_1c2120	Neut_1965
fig/66666666.60966.peg.2120	CDS	1968691	1969206	1	+	516	FIG00859317; hypothetical protein	-none-	D23_1c2121	Neut_1966
fig/66666666.60966.peg.2121	CDS	1969209	1969910	3	+	702	InterPro IPR000179	-none-	D23_1c2122	Neut_1967
fig/66666666.60966.peg.2122	CDS	1969960	1970334	1	+	375	InterPro IPR003807	-none-	D23_1c2123	Neut_1968
fig/66666666.60966.peg.2123	CDS	1971838	1970381	-1	-	1458	COGs COG2149 Catalase (EC 1.11.1.6)	Oxidative stress; Phototrespiration (oxidative C2 cycle); Protection from Reactive Oxygen Species	D23_1c2124	Neut_1969
fig/66666666.60966.peg.2124	CDS	1973211	1971865	-3	-	1347	FIG00858984; hypothetical protein	-none-	D23_1c2125	Neut_1970
fig/66666666.60966.peg.2125	CDS	1974046	1973246	-1	-	801	Protein of unknown function DUF81	-none-	D23_1c2126	Neut_1971
fig/66666666.60966.peg.2126	CDS	1974579	1974124	-3	-	456	Protein of unknown function DUF55	-none-	D23_1c2127	Neut_1972
fig/66666666.60966.peg.2127	CDS	1976514	1974985	-3	-	1530	Capsular polysaccharide biosynthesis protein	Capsular Polysaccharides Biosynthesis and Assembly	D23_1c2128	NA
fig/66666666.60966.peg.2128	CDS	1977322	1976504	-1	-	819	WcbQ Oxidoreductase, short-chain dehydrogenase/reductase family (EC 1.1.1.-)	Capsular Polysaccharides Biosynthesis and Assembly	D23_1c2129	Neut_1974
fig/66666666.60966.peg.2129	CDS	1978801	1977323	-1	-	1479	Glycosyltransferase	-none-	D23_1c2130	Neut_1975
fig/66666666.60966.peg.2130	CDS	1979888	1978803	-2	-	1086	possible spore protein [UI:20467420]	-none-	D23_1c2131	Neut_1976
fig/66666666.60966.peg.2131	CDS	1981090	1979921	-1	-	1170	FIG00858788; hypothetical protein	-none-	D23_1c2132	Neut_1977

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 GbkId	C91 Alias
fig/66666666.60966.psg.2132	CDS	1982214	1981231	-3	-	984	hypothetical protein	-none-	D23_1c2133	NA
fig/66666666.60966.psg.2133	CDS	1982778	1982218	-3	-	561	hypothetical protein	-none-	D23_1c2134	NA
fig/66666666.60966.psg.2134	CDS	1983862	1982783	-1	-	1080	Glycosyltransferase (EC 2.4.1.—)	-none-	D23_1c2135	Neut_1982
fig/66666666.60966.psg.2135	CDS	1984198	1983896	-1	-	303	hypothetical protein	-none-	D23_1c2136	Neut_1983
fig/66666666.60966.psg.2136	CDS	1984883	1984704	-2	-	180	Mobile element protein	-none-	D23_1c2137	Neut_1984
fig/66666666.60966.psg.2137	CDS	1985076	1984852	-3	-	225	hypothetical protein	-none-	D23_1c2138	NA
fig/66666666.60966.psg.2139	CDS	1985629	1986384	1	+	756	hypothetical protein	-none-	D23_1c2139	Neut_1985
fig/66666666.60966.psg.2140	CDS	1987287	1988645	3	+	1359	FIG00860556; hypothetical protein	-none-	D23_1c2140	Neut_1986
fig/66666666.60966.psg.2141	CDS	1988761	1989375	1	+	615	Cytochrome oxidase biogenesis protein Scc1/SenC/PrrC, putative copper metallochaperone	Biogenesis of cytochrome c oxidases	D23_1c2141	Neut_1987
fig/66666666.60966.psg.2142	CDS	1989381	1990007	3	+	627	FIG00859788; hypothetical protein	-none-	D23_1c2142	Neut_1988
fig/66666666.60966.psg.2145	CDS	1990866	1990588	-3	-	279	hypothetical membrane protein	-none-	D23_1c2144	Neut_1990
fig/66666666.60966.psg.2146	CDS	1991501	1991046	-2	-	456	InterPro IPR000485	-none-	D23_1c2145	Neut_1991
fig/66666666.60966.psg.2147	CDS	1993914	1991785	-3	-	2130	Phosphate acetyltransferase (EC 2.3.1.8)	Fermentations: Lactate; Pyruvate metabolism II: acetyl-CoA, acetogenesis from pyruvate	D23_1c2146	Neut_1995
fig/66666666.60966.psg.2148	CDS	1995278	1994253	-2	-	1026	Fructose-bisphosphate aldolase class I (EC 4.1.2.13)	Calvin-Benson cycle; Glycolysis and Gluconeogenesis	D23_1c2147	Neut_1996
fig/66666666.60966.psg.2150	CDS	1995705	1995851	3	+	147	hypothetical protein	-none-	D23_1c2148	NA
fig/66666666.60966.psg.2152	CDS	1995966	1996298	3	+	333	DNA-binding protein	-none-	D23_1c2150	Neut_1998
fig/66666666.60966.psg.2154	CDS	1996639	1996932	1	+	294	protein of unknown function DUF497	-none-	D23_1c2151	Neut_1999
fig/66666666.60966.psg.2155	CDS	1996922	1997203	2	+	282	hypothetical protein	-none-	D23_1c2152	Neut_2000
fig/66666666.60966.psg.2156	CDS	1997890	1998093	1	+	204	Mobile element protein	-none-	D23_1c2154	NA
fig/66666666.60966.psg.2157	CDS	1998565	1998266	-1	-	300	VapC toxin protein	Toxin-antitoxin replicon stabilization systems	D23_1c2155	NA
fig/66666666.60966.psg.2158	CDS	1998899	1998666	-2	-	234	VapB protein (antitoxin to VapC)	Toxin-antitoxin replicon stabilization systems	D23_1c2156	NA
fig/66666666.60966.psg.2159	CDS	1999186	1999344	1	+	159	hypothetical protein	-none-	D23_1c2157	NA
fig/66666666.60966.psg.2160	CDS	1999629	1999339	-3	-	291	transcriptional regulator, XRE family	-none-	D23_1c2158	NA
fig/66666666.60966.psg.2161	CDS	1999997	1999692	-2	-	306	Phage-related protein	-none-	D23_1c2159	NA
fig/66666666.60966.psg.2162	CDS	2000147	2000022	-2	-	126	hypothetical protein	-none-	D23_1c2160	NA
fig/66666666.60966.psg.2163	CDS	2000220	2000486	3	+	267	Mobile element protein	-none-	D23_1c2161	NA
fig/66666666.60966.psg.2164	CDS	2000507	2000995	2	+	489	Mobile element protein	-none-	D23_1c2162	NA
fig/66666666.60966.psg.2165	CDS	2001470	2000997	-2	-	474	Mobile element protein	-none-	D23_1c2163	Neut_1256
fig/66666666.60966.psg.2166	CDS	2001936	2001544	-3	-	393	hypothetical protein	-none-	D23_1c2164	Neut_2449
fig/66666666.60966.psg.2167	CDS	2002011	2002328	3	+	318	Mobile element protein	-none-	D23_1c2165	NA

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld	C91 Alias
fig/66666666.60966.peg.2168	CDS	2003529	2002369	-3	-	1161	CDP-4-dehydro-6-deoxy-D-glucose 3-dehydratase (EC 4.2.1.—)	-none-	D23_1c2166	NA
fig/66666666.60966.peg.2169	CDS	2004426	2003539	-3	-	888	NAD-dependent epimerase/dehydratase family protein	CBSS-296591.1.peg.2330	D23_1c2167	NA
fig/66666666.60966.peg.2170	CDS	2005586	2004468	-2	-	1119	GDP-mannose 4,6-dehydratase (EC 4.2.1.47)	-none-	D23_1c2168	Neut_0156
fig/66666666.60966.peg.2171	CDS	2007467	2005632	-2	-	1836	hypothetical protein	-none-	D23_1c2169	NA
fig/66666666.60966.peg.2172	CDS	2010568	2007473	-1	-	3096	Minor teichoic acid biosynthesis protein	-none-	D23_1c2170	NA
fig/66666666.60966.peg.2173	CDS	2012163	2010694	-3	-	1470	GgaB InterPro IPR001173	-none-	D23_1c2171	NA
fig/66666666.60966.peg.2174	CDS	2015828	2012172	-2	-	3657	Beta-1,3-glucosyltransferase	-none-	D23_1c2172	NA
fig/66666666.60966.peg.2175	CDS	2017307	2016147	-2	-	1161	Capsular polysaccharide export system inner membrane protein KpsE	Capsular Polysaccharides Biosynthesis and Assembly	D23_1c2173	NA
fig/66666666.60966.peg.2176	CDS	2017966	2017304	-1	-	663	Capsular polysaccharide ABC transporter, ATP-binding protein KpsT	Capsular Polysaccharides Biosynthesis and Assembly	D23_1c2174	Neut_0152
fig/66666666.60966.peg.2177	CDS	2018757	2017963	-3	-	795	Capsular polysaccharide ABC transporter, permease protein KpsM	Capsular Polysaccharides Biosynthesis and Assembly; Rhamnose containing glycans	D23_1c2175	NA
fig/66666666.60966.peg.2178	CDS	2019947	2018757	-2	-	1191	Capsular polysaccharide biosynthesis/export periplasmic protein WcbC	Capsular Polysaccharides Biosynthesis and Assembly	D23_1c2176	Neut_2119
fig/66666666.60966.peg.2179	CDS	2021279	2019957	-2	-	1323	8-amino-7-oxononanoate synthase (EC 2.3.1.47)	Biotin biosynthesis; Biotin biosynthesis Experimental; Biotin synthesis cluster	D23_1c2177	Neut_0461
fig/66666666.60966.peg.2180	CDS	2025121	2021276	-1	-	3846	Capsular polysaccharide biosynthesis fatty acid synthase WcbR	-none-	D23_1c2178	NA
fig/66666666.60966.peg.2181	CDS	2028811	2025266	-1	-	3546	Capsular polysaccharide biosynthesis fatty acid synthase WcbR	-none-	D23_1c2179	Neut_2003
fig/66666666.60966.peg.2182	CDS	2029053	2029166	3	+	114	hypothetical protein	-none-	D23_1c2180	NA
fig/66666666.60966.peg.2183	CDS	2029212	2029352	3	+	141	hypothetical protein	-none-	D23_1c2181	NA
fig/66666666.60966.peg.2184	CDS	2031017	2029434	-2	-	1584	L-aspartate oxidase (EC 1.4.3.16)	Mycobacterium virulence operon possibly involved in quinolinate biosynthesis;	D23_1c2182	Neut_2004

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.									
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld C91 Alias
fig/66666666.60966.peg.2185	CDS	2031267	2032187	3	+	921	Branched-chain amino acid aminotransferase (EC 2.6.1.42)	 NAD and NADP cofactor biosynthesis global Alanine biosynthesis; Branched-Chain Amino Acid Biosynthesis; Leucine Biosynthesis; Pyruvate Alanine Serine Interconversions -none-	D23_1c2184 Neut_2005
fig/66666666.60966.peg.2186	CDS	2032212	2032421	3	+	210	FIG00858455: hypothetical protein		
fig/66666666.60966.peg.2187	CDS	2032485	2033864	3	+	1380	Phosphomannomutase (EC 5.4.2.8) Phosphoglucosyltransferase (EC 5.4.2.2)		
fig/66666666.60966.peg.2188	CDS	2033901	2035508	3	+	1608	NAD synthetase (EC 6.3.1.5)/Glutamine amidotransferase chain of NAD synthetase	NAD and NADP cofactor biosynthesis global; NAD and NADP cofactor biosynthesis global -none-	D23_1c2185 Neut_2006 D23_1c2186 Neut_2007
fig/66666666.60966.peg.2189	CDS	2035472	2037178	2	+	1707	Exported zinc metalloprotease YfgC precursor		D23_1c2188 Neut_2009
fig/66666666.60966.peg.2190	CDS	2037257	2038402	2	+	1146	Macrolide-specific efflux protein MacA		D23_1c2189 Neut_2010
fig/66666666.60966.peg.2191	CDS	2038413	2039180	3	+	768	Macrolide export ATP-binding/permease protein MacB (EC 3.6.3.—)	Multidrug Resistance Efflux Pumps Multidrug Resistance Efflux Pumps	D23_1c2190 Neut_2011
fig/66666666.60966.peg.2192	CDS	2039177	2040400	2	+	1224	Macrolide export ATP-binding/permease protein MacB (EC 3.6.3.—)	Multidrug Resistance Efflux Pumps	D23_1c2191 Neut_2012
fig/66666666.60966.peg.2193	CDS	2040442	2040612	1	+	171	hypothetical protein	-none-	D23_1c2192 NA
fig/66666666.60966.peg.2195	CDS	2040690	2040875	3	+	186	hypothetical protein	-none-	D23_1c2193 NA
fig/66666666.60966.peg.2196	CDS	2040926	2042374	2	+	1449	ATP synthase beta chain (EC 3.6.3.14)	-none-	D23_1c2194 Neut_2013
fig/66666666.60966.peg.2197	CDS	2042371	2042781	1	+	411	ATP synthase epsilon chain (EC 3.6.3.14)	-none-	D23_1c2195 Neut_2014
fig/66666666.60966.peg.2198	CDS	2042778	2043059	3	+	282	ATP synthase protein I	-none-	D23_1c2196 Neut_2015
fig/66666666.60966.peg.2199	CDS	2043105	2043398	3	+	294	FIG048548: ATP synthase protein I2	-none-	D23_1c2197 Neut_2016
fig/66666666.60966.peg.2200	CDS	2043420	2044112	3	+	693	ATP synthase A chain (EC 3.6.3.14)	-none-	D23_1c2198 Neut_2017
fig/66666666.60966.peg.2201	CDS	2044115	2044393	2	+	279	ATP synthase C chain (EC 3.6.3.14)	-none-	D23_1c2199 Neut_2018
fig/66666666.60966.peg.2202	CDS	2044400	2045170	2	+	771	ATP synthase B chain (EC 3.6.3.14)	-none-	D23_1c2200 Neut_2019

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.									
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkid C91 Alias
fig 66666666.60966.peg.2203	CDS	2045183	2046733	2	+	1551	ATP synthase alpha chain (EC 3.6.3.14)	-none-	D23_1c2201 Neut_2020
fig 66666666.60966.peg.2204	CDS	2046810	2047607	3	+	798	ATP synthase gamma chain (EC 3.6.3.14)	-none-	D23_1c2202 Neut_2021
fig 66666666.60966.peg.2206	CDS	2048344	2050986	1	+	2643	DNA mismatch repair protein MutS	DNA repair, bacterial MutL-MutS system; DNA repair system including RecA, MutS and a hypothetical protein	D23_1c2204 Neut_2022
fig 66666666.60966.peg.2207	CDS	2051500	2051021	-1	-	480	FKBP-type peptidyl-prolyl cis-trans isomerase SlyD (EC 5.2.1.8)	G3E family of P-loop GTPases (metalcenter biosynthesis); Peptidyl-prolyl cis-trans isomerase; Potassium homeostasis	D23_1c2205 Neut_2023
fig 66666666.60966.peg.2208	CDS	2052207	2051656	-3	-	552	Ribonuclease HII (EC 3.1.26.4)	Ribonuclease H; Ribonucleases in Bacillus	D23_1c2206 Neut_2024
fig 66666666.60966.peg.2209	CDS	2052737	2052294	-2	-	444	(3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase (EC 4.2.1.-)	-none-	D23_1c2207 Neut_2025
fig 66666666.60966.peg.2210	CDS	2053333	2052770	-1	-	564	Outer membrane protein H precursor	Lipopolysaccharide assembly; Periplasmic Stress Response	D23_1c2208 Neut_2026
fig 66666666.60966.peg.2211	CDS	2055635	2053359	-2	-	2277	Outer membrane protein assembly factor YaeI precursor	Lipopolysaccharide assembly	D23_1c2209 Neut_2027
fig 66666666.60966.peg.2212	CDS	2057020	2055638	-1	-	1383	Membrane-associated zinc metalloprotease	-none-	D23_1c2210 Neut_2028
fig 66666666.60966.peg.2213	CDS	2058265	2057024	-1	-	1242	1-deoxy-D-xylulose 5-phosphate reductoisomerase (EC 1.1.1.267)	CBSS-83331.1.peg.3039; Isoprenoid Biosynthesis; Nonmevalonate Branch of Isoprenoid Biosynthesis	D23_1c2211 Neut_2029
fig 66666666.60966.peg.2214	CDS	2059089	2058262	-3	-	828	Phosphatidate cytidyltransferase (EC 2.7.7.41)	Glycerolipid and Glycerophospholipid Metabolism in Bacteria	D23_1c2212 Neut_2030
fig 66666666.60966.peg.2215	CDS	2059870	2059124	-1	-	747	Undecaprenyl diposphate synthase (EC 2.5.1.31)	CBSS-83331.1.peg.3039; Isoprenoid Biosynthesis; Isoprenoids for Quinones; Polyprenyl	D23_1c2213 Neut_2031

SUPPLEMENTARY TABLE 1-continued

Feature ID	Annotation of genes in SEQ ID NO: 1.					Subsystem	D23 Gbkld	C91 Alias
	Type	Start	Stop	Frame	Strand	Length (bp)	Function	
fig/66666666.60966.peg.2216	CDS	2060476	2059919	-1	-	558	Ribosome recycling factor	D23_1c2214 Neut_2032
fig/66666666.60966.peg.2217	CDS	2061239	2060577	-2	-	663	Uridylate kinase (EC 2.7.4.—)	D23_1c2215 Neut_2033
fig/66666666.60966.peg.2218	CDS	2061489	2061361	-3	-	129	hypothetical protein	D23_1c2216 NA
fig/66666666.60966.peg.2219	CDS	2062729	2061845	-1	-	885	Translation elongation factor 1s	D23_1c2217 Neut_2034
fig/66666666.60966.peg.2220	CDS	2063557	2062802	-1	-	756	SSU ribosomal protein S2p (SAe)	D23_1c2218 Neut_2035
fig/66666666.60966.peg.2221	CDS	2064515	2063772	-2	-	744	transcriptional regulator, Crp/Fnr family	D23_1c2219 Neut_2036
fig/66666666.60966.peg.2222	CDS	2064873	2064562	-3	-	312	Cytochrome c, class I	D23_1c2220 Neut_2037
fig/66666666.60966.peg.2223	CDS	2067162	2065111	-3	-	2052	Zinc-regulated outer membrane receptor	D23_1c2221 Neut_2038
fig/66666666.60966.peg.2225	CDS	2067907	2068395	1	+	489	Zinc uptake regulation protein ZUR	D23_1c2222 NA
fig/66666666.60966.peg.2226	CDS	2068438	2068560	1	+	123	Mobile element protein	D23_1c2223 Neut_0884
fig/66666666.60966.peg.2228	CDS	2068811	2069020	2	+	210	hypothetical protein	D23_1c2224 NA
fig/66666666.60966.peg.2231	CDS	2069466	2069332	-3	-	135	hypothetical protein	D23_1c2225 NA
fig/66666666.60966.peg.2232	CDS	2069455	2070327	1	+	873	COG0613, Predicted metal-dependent phosphoesterases (PHP family)	D23_1c2226 Neut_2039
fig/66666666.60966.peg.2233	CDS	2070450	2071082	3	+	633	YciO family	D23_1c2227 Neut_2040
fig/66666666.60966.peg.2234	CDS	2071075	2071740	1	+	666	FIG004556: membrane metalloprotease	D23_1c2228 Neut_2041
fig/66666666.60966.peg.2235	CDS	2071817	2073019	2	+	1203	Tryptophanyl-HRNA synthetase (EC 6.1.1.2)	D23_1c2229 Neut_2042

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld	C91 Alias
fig/66666666.60966.psg.2236	CDS	2073064	2073864	1	+	801	Segregation and condensation protein A	-none-	D23_1c2230	Neut_2043
fig/66666666.60966.psg.2237	CDS	2073830	2074480	2	+	651	Segregation and condensation protein B	-none-	D23_1c2231	Neut_2044
fig/66666666.60966.psg.2238	CDS	2074504	2075751	1	+	1248	Isocitrate dehydrogenase [NADP] (EC 1.1.1.42)	5-FCL-like protein; TCA Cycle	D23_1c2232	Neut_2045
fig/66666666.60966.psg.2239	CDS	2076095	2075892	-2	-	204	Cold shock protein CspD	Cold shock, CspA family of proteins	D23_1c2233	Neut_2046
fig/66666666.60966.psg.2240	CDS	2076577	2076407	-1	-	171	hypothetical protein	-none-	D23_1c2234	NA
fig/66666666.60966.psg.2241	CDS	2076576	2076779	3	+	204	ATP-dependent Clp protease adaptor protein ClpS	ClpAS cluster; Proteolysis in bacteria, ATP-dependent	D23_1c2235	Neut_2047
fig/66666666.60966.psg.2242	CDS	2076781	2079051	1	+	2271	ATP-dependent Clp protease ATP-binding subunit ClpA	ClpAS cluster; Proteolysis in bacteria, ATP-dependent; Ribosome recycling related cluster	D23_1c2236	Neut_2048
fig/66666666.60966.psg.2243	CDS	2079170	2080129	2	+	960	TRAP transporter solute receptor, unknown substrate 6	TRAP Transporter unknown substrate 6	D23_1c2237	Neut_2049
fig/66666666.60966.psg.2244	CDS	2080157	2080885	2	+	729	Orotate phosphoribosyltransferase (EC 2.4.2.10)	De Novo Pyrimidine Synthesis	D23_1c2238	Neut_2050
fig/66666666.60966.psg.2245	CDS	2081203	2082156	1	+	954	COGs COG0715	-none-	D23_1c2239	Neut_2051
fig/66666666.60966.psg.2246	CDS	2082161	2082997	2	+	837	Alkanesulfonates transport system	Alkanesulfonate assimilation	D23_1c2240	Neut_2052
fig/66666666.60966.psg.2247	CDS	2083232	2083038	-2	-	195	permease protein	-none-	D23_1c2241	NA
fig/66666666.60966.psg.2248	CDS	2083287	2083823	3	+	537	ABC-type nitrate/sulfonate/bicarbonate transport system, ATPase component	Alkanesulfonate assimilation	D23_1c2242	Neut_2053
fig/66666666.60966.psg.2249	CDS	2083938	2086493	3	+	2556	Glycogen phosphorylase (EC 2.4.1.1)	Glycogen metabolism	D23_1c2243	Neut_2054
fig/66666666.60966.psg.2251	CDS	2086807	2086661	-1	-	147	hypothetical protein	-none-	D23_1c2244	NA
fig/66666666.60966.psg.2252	CDS	2086806	2087288	3	+	483	Flagellar biosynthesis protein Flil	Flagellum	D23_1c2245	Neut_2055
fig/66666666.60966.psg.2253	CDS	2087301	2088293	3	+	993	Flagellar motor switch protein FlmM	Flagellar motility; Flagellum	D23_1c2246	Neut_2056
fig/66666666.60966.psg.2254	CDS	2088322	2088795	1	+	474	Flagellar motor switch protein Flin	 Flagellum	D23_1c2247	Neut_2057
fig/66666666.60966.psg.2255	CDS	2088822	2089265	3	+	444	Flagellar biosynthesis protein FlhQ	 Flagellum	D23_1c2248	Neut_2058
fig/66666666.60966.psg.2256	CDS	2089255	2090040	1	+	786	Flagellar biosynthesis protein Flp	Flagellum	D23_1c2249	Neut_2059

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld	C91 Alias
fig/66666666.60966.peg.2257	CDS	2090056	2090331	1	+	276	Flagellar biosynthesis protein FlhQ	Flagellum	D23_1c2250	Neut_2060
fig/66666666.60966.peg.2258	CDS	2090429	2091229	2	+	801	Flagellar biosynthesis protein FlhR	Flagellar motility; Flagellum	D23_1c2251	Neut_2061
fig/66666666.60966.peg.2259	CDS	2091389	2093485	2	+	2097	FIG00858519; hypothetical protein	-none-	D23_1c2252	Neut_2062
fig/66666666.60966.peg.2260	CDS	2093591	2095027	2	+	1437	FIG00858578; hypothetical protein	-none-	D23_1c2253	Neut_2063
fig/66666666.60966.peg.2261	CDS	2096049	2095090	-3	-	960	alpha/beta hydrolase fold	-none-	D23_1c2254	Neut_2064
fig/66666666.60966.peg.2262	CDS	2096035	2096187	1	+	153	hypothetical protein	-none-	D23_1c2255	NA
fig/66666666.60966.peg.2263	CDS	2097510	2096296	-3	-	1215	Argininosuccinate synthase (EC 6.3.4.5)	Arginine Biosynthesis-- glo; Arginine Biosynthesis extended	D23_1c2256	Neut_2065
fig/66666666.60966.peg.2264	CDS	2098473	2097550	-3	-	924	Ornithine carbamoyltransferase (EC 2.1.3.3)	Arginine Biosynthesis-- glo; Arginine Biosynthesis extended; Ornithine Degradation	D23_1c2257	Neut_2066
fig/66666666.60966.peg.2266	CDS	2099808	2099680	-3	-	129	hypothetical protein	-none-	D23_1c2259	Neut_2067
fig/66666666.60966.peg.2265	CDS	2099605	2098649	-1	-	957	Acetylornithine aminotransferase (EC 2.6.1.11)	Arginine Biosynthesis-- glo; Arginine Biosynthesis extended	D23_1c2259	Neut_2067
fig/66666666.60966.peg.2267	CDS	2100416	2100066	-2	-	351	FIG00858925; hypothetical protein	-none-	D23_1c2260	Neut_2068
fig/66666666.60966.peg.2268	CDS	2100762	2100517	-3	-	246	FIG00859242; hypothetical protein	-none-	D23_1c2261	Neut_2069
fig/66666666.60966.peg.2269	CDS	2102838	2100763	-3	-	2076	FIG00858999; hypothetical protein	-none-	D23_1c2262	Neut_2070
fig/66666666.60966.peg.2270	CDS	2103683	2102844	-2	-	840	Cysteine synthase B (EC 2.5.1.47)	Cysteine Biosynthesis	D23_1c2263	Neut_2071
fig/66666666.60966.peg.2272	CDS	2106646	2103854	-1	-	2793	FIG00860005; hypothetical protein	-none-	D23_1c2264	Neut_2072
fig/66666666.60966.peg.2274	CDS	2109519	2108932	-3	-	588	hypothetical protein	-none-	D23_1c2268	Neut_2074
fig/66666666.60966.peg.2275	CDS	2110331	2109591	-2	-	741	putative (U92432) ORF4 (Nitrosospira sp. NpAV)	-none-	D23_1c2269	Neut_2316
fig/66666666.60966.peg.2276	CDS	2111663	2110398	-2	-	1266	Particulate methane monooxygenase B-subunit (EC 1.14.13.25)	Particulate methane monooxygenase (pMMO)	D23_1c2270	Neut_2317
fig/66666666.60966.peg.2277	CDS	2112493	2111663	-1	-	831	Particulate methane monooxygenase A-subunit (EC 1.14.13.25)	Particulate methane monooxygenase (pMMO)	D23_1c2271	Neut_2318
fig/66666666.60966.peg.2278	CDS	2113482	2112667	-3	-	816	Particulate methane monooxygenase C-subunit (EC 1.14.13.25)	Particulate methane monooxygenase (pMMO)	D23_1c2272	Neut_2319
fig/66666666.60966.peg.2279	CDS	2114624	2114052	-2	-	573	CDP-diacylglycerol--glycerol-3-phosphate 3-	Glycerolipid and Glycerophospholipid	D23_1c2273	Neut_2079

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 GbkId	C91 Alias
fig/66666666.60966.peg.2280	CDS	2116254	2114755	-3	-	1500	phosphatidyltransferase (EC 2.7.8.5)	Metabolism in Bacteria	D23_1c2274	Neut_2080
							Glycerol kinase (EC 2.7.1.30)	Glycerol and Glycerol-3-phosphate Uptake and Utilization;		
								 Glycerolipid and Glycerophospholipid Metabolism in Bacteria		
								Transcription factors bacterial		
fig/66666666.60966.peg.2281	CDS	2116718	2116305	-2	-	414	Regulator of nucleoside diphosphate kinase	Transcription factors bacterial	D23_1c2275	Neut_2081
fig/66666666.60966.peg.2282	CDS	2116748	2116897	2	+	150	hypothetical protein	-none-	D23_1c2276	NA
fig/66666666.60966.peg.2283	CDS	2117727	2117005	-3	-	723	Phosphate regulon transcriptional regulatory protein PhoB (SphR)	High affinity phosphate transporter and control of PHO regulon;	D23_1c2277	Neut_2082
fig/66666666.60966.peg.2284	CDS	2119187	2118225	-2	-	963	Mobile element protein	 PhoR-PhoB two-component regulatory system; Phosphate metabolism	D23_1c2278	Neut_1746
							Aspartokinase (EC 2.7.2.4)	CBSS-216591.1.peg.168; Lysine Biosynthesis DAP Pathway, GJO scratch; Threonine and Homoserine Biosynthesis		
fig/66666666.60966.peg.2288	CDS	2122439	2121855	-2	-	585	Putative peptidoglycan binding domain 1	-none-	D23_1c2283	Neut_2085
fig/66666666.60966.peg.2289	CDS	2122914	2124089	3	+	1176	Acetate kinase (EC 2.7.2.1)	Fermentations: Lactate; Pyruvate metabolism II: acetyl-CoA, acetogenesis from pyruvate	D23_1c2284	Neut_2086
fig/66666666.60966.peg.2290	CDS	2124134	2126509	2	+	2376	Xylulose-5-phosphate phosphoketolase (EC 4.1.2.9); Fructose-6-phosphate phosphoketolase (EC 4.1.2.22)	Fermentations: Lactate; Fermentations: Lactate; Fermentations: Lactate; Pentose phosphate pathway; Pentose phosphate pathway	D23_1c2285	Neut_2087
fig/66666666.60966.peg.2291	CDS	2126813	2126977	2	+	165	hypothetical protein	-none-	D23_1c2287	NA
fig/66666666.60966.peg.2292	CDS	2128152	2127058	-3	-	1095	Transaldolase (EC 2.2.1.2)	Pentose phosphate pathway	D23_1c2288	Neut_2089
fig/66666666.60966.peg.2293	CDS	2128367	2128218	-2	-	150	hypothetical protein	-none-	D23_1c2289	NA
fig/66666666.60966.peg.2294	CDS	2128351	2128908	1	+	558	FIG006045; Sigma factor, ECF subfamily	Iron siderophore sensor & receptor system	D23_1c2290	Neut_2090
fig/66666666.60966.peg.2295	CDS	2128915	2129877	1	+	963	Iron siderophore sensor protein	Iron siderophore sensor & receptor system	D23_1c2291	Neut_2091
fig/66666666.60966.peg.2296	CDS	2129956	2132253	1	+	2298	TonB-dependent receptor	Ton and Tol transport systems	D23_1c2292	Neut_2092

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.									
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld C91 Alias
fig/66666666.60966.peg.2297	CDS	2132801	2132271	-2	-	531	FIG00858714: hypothetical protein	-none-	D23_1c2293 Neut_2093
fig/66666666.60966.peg.2298	CDS	2133157	2132813	-1	-	345	FIG00858447: hypothetical protein	-none-	D23_1c2294 Neut_2094
fig/66666666.60966.peg.2299	CDS	2134515	2133133	-3	-	1383	Xaa-Pro aminopeptidase (EC 3.4.11.9)	Aminopeptidases (EC 3.4.11.-); CBSS-87626.3.peg.3639	D23_1c2295 Neut_2095
fig/66666666.60966.peg.2301	CDS	2134633	2135313	1	+	681	Ribulose-phosphate 3-epimerase (EC 5.1.3.1)	Calvin-Benson cycle; Riboflavin synthesis cluster	D23_1c2296 Neut_2096
fig/66666666.60966.peg.2302	CDS	2135328	2136050	3	+	723	Phosphoglycolate phosphatase (EC 3.1.3.18)	2-phosphoglycolate salvage; Glycolate, glyoxylate interconversions; Photorespiration (oxidative C2 cycle)	D23_1c2297 Neut_2097
fig/66666666.60966.peg.2303	CDS	2136160	2137647	1	+	1488	Anthranelate synthase, aminase component (EC 4.1.3.27)	Chlorismate: Intermediate for synthesis of Tryptophan, PABA antibiotics, PABA, 3-hydroxyanthranilate and more; Tryptophan synthesis	D23_1c2298 Neut_2098
fig/66666666.60966.peg.2304	CDS	2138110	2137682	-1	-	429	Ferredoxin reductase	Anaerobic respiratory reductases	D23_1c2299 Neut_2099
fig/66666666.60966.peg.2305	CDS	2138482	2138138	-1	-	345	FIG00858997: hypothetical protein	-none-	D23_1c2300 Neut_2100
fig/66666666.60966.peg.2306	CDS	2138586	2139236	3	+	651	Iron-sulfur cluster regulator SufR	CBSS-196164.1.peg.1690	D23_1c2301 Neut_2101
fig/66666666.60966.peg.2307	CDS	2139353	2140534	2	+	1182	FIG00859085: hypothetical protein	-none-	D23_1c2302 Neut_2102
fig/66666666.60966.peg.2308	CDS	2142375	2140606	-3	-	1770	Dihydrolipoamide dehydrogenase of pyruvate dehydrogenase complex (EC 1.8.1.4)	Pyruvate metabolism II: acetyl-CoA, acetogenesis from pyruvate; TCA Cycle	D23_1c2304 Neut_2103
fig/66666666.60966.peg.2309	CDS	2143648	2142380	-1	-	1269	Gamma-glutamyl phosphate reductase (EC 1.2.1.41)	Proline Synthesis	D23_1c2305 Neut_2104
fig/66666666.60966.peg.2310	CDS	2144399	2143713	-2	-	687	InterPro IPR001440 COGs COG0457	-none-	D23_1c2306 Neut_2105
fig/66666666.60966.peg.2311	CDS	2145656	2144412	-2	-	1245	TPR/glycosyl transferase domain protein	-none-	D23_1c2307 Neut_2106
fig/66666666.60966.peg.2312	CDS	2146522	2145650	-1	-	873	Esterase/lipase/thioesterase family active site	-none-	D23_1c2308 Neut_2107

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 GbkId	C91 Alias
fig/66666666.60966.peg.2313	CDS	2147373	2146519	-3	-	855	Esterase/lipase/thioesterase family active site	-none-	D23_1c2309	Neut_2108
fig/66666666.60966.peg.2314	CDS	2147631	2147380	-3	-	252	FIG00858580: hypothetical protein	-none-	D23_1c2310	Neut_2109
fig/66666666.60966.peg.2315	CDS	2147817	2149007	3	+	1191	Tetraacyl-disaccharide 4'-kinase (EC 2.7.1.30)FIG002473: Protein YcaR in KDO2-Lipid A biosynthesis cluster	Broadly distributed proteins not in subsystems; KDO2-Lipid A biosynthesis cluster 2; KDO2-Lipid A biosynthesis cluster 2	D23_1c2311	Neut_2110
fig/66666666.60966.peg.2316	CDS	2149225	2149353	1	+	129	hypothetical protein	-none-	D23_1c2312	NA
fig/66666666.60966.peg.2318	CDS	2150705	2149527	-2	-	1179	Lipid carrier: UDP-N-acetylglactosaminyltransferase (EC 2.4.1.—) Alpha-1,3-N-acetylglactosamine transferase PglA (EC 2.4.1.—); Putative glycosyltransferase	CBSS-296591.1.peg.2330; CBSS-296591.1.peg.2330; CBSS-296591.1.peg.2330; N-linked Glycosylation in Bacteria; N-linked Glycosylation in Bacteria	D23_1c2313	Neut_2112
fig/66666666.60966.peg.2319	CDS	2151819	2150689	-3	-	1131	Glycosyl transferase, group 1 family protein	-none-	D23_1c2314	Neut_2113
fig/66666666.60966.peg.2320	CDS	2153132	2151816	-2	-	1317	hypothetical protein	-none-	D23_1c2315	NA
fig/66666666.60966.peg.2321	CDS	2154290	2153199	-2	-	1092	Alpha-1,4-N-acetylglactosamine transferase PglJ (EC 2.4.1.—)	N-linked Glycosylation in Bacteria	D23_1c2316	Neut_2115
fig/66666666.60966.peg.2322	CDS	2155270	2154290	-1	-	981	COGs COG0439	-none-	D23_1c2317	Neut_2116
fig/66666666.60966.peg.2323	CDS	2156578	2155337	-1	-	1242	Membrane protein involved in the export of O-antigen, teichoic acid lipoteichoic acids	-none-	D23_1c2318	Neut_2117
fig/66666666.60966.peg.2324	CDS	2156818	2157207	1	+	390	Low molecular weight protein-tyrosine-phosphatase Wzb (EC 3.1.3.48)	Capsular Polysaccharides Biosynthesis and Assembly	D23_1c2319	Neut_2118
fig/66666666.60966.peg.2325	CDS	2157230	2158477	2	+	1248	Capsule polysaccharide export protein	-none-	D23_1c2320	Neut_2119
fig/66666666.60966.peg.2326	CDS	2158531	2160774	1	+	2244	Tyrosine-protein kinase Wzc (EC 2.7.10.2)	Capsular Polysaccharides Biosynthesis and Assembly	D23_1c2321	Neut_2120
fig/66666666.60966.peg.2327	CDS	2160771	2162330	3	+	1560	Undecaprenyl-phosphate N-acetylglucosaminyl 1-phosphate transferase (EC 2.7.8.—)	-none-	D23_1c2322	Neut_2121

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld	C91 Alias
fig/66666666.60966.peg.2329	CDS	2165972	2162400	-2	-	3573	Chromosome partition protein smc	DNA structural proteins, bacterial -none-	D23_1c2323	Neut_2122
fig/66666666.60966.peg.2330	CDS	2166059	2166478	2	+	420	NADPH-dependent 7-cyano-7-deazaguanine reductase (EC 1.7.1.13)		D23_1c2324	Neut_2123
fig/66666666.60966.peg.2331	CDS	2166527	2167927	2	+	1401	Fumarate hydratase class II (EC 4.2.1.2)	TCA Cycle	D23_1c2325	Neut_2124
fig/66666666.60966.peg.2332	CDS	2168192	2168052	-2	-	141	hypothetical protein	-none-	D23_1c2326	NA
fig/66666666.60966.peg.2333	CDS	2168473	2168640	1	+	168	hypothetical protein	-none-	D23_1c2327	NA
fig/66666666.60966.peg.2334	CDS	2168888	2169181	2	+	294	Mobile element protein	-none-	D23_1c2328	Neut_1719
fig/66666666.60966.peg.2335	CDS	2169280	2170158	1	+	879	Mobile element protein	-none-	D23_1c2329	Neut_1720
fig/66666666.60966.peg.2337	CDS	2171884	2170655	-1	-	1230	diguanylate phosphodiesterase	-none-	D23_1c2332	Neut_2126
fig/66666666.60966.peg.2338	CDS	2172707	2171898	-2	-	810	Putative dihenne cytochrome c-553	Soluble cytochromes and functionally related electron carriers	D23_1c2333	Neut_2127
fig/66666666.60966.peg.2339	CDS	2172852	2172977	3	+	126	hypothetical protein	-none-	D23_1c2334	NA
fig/66666666.60966.peg.2340	CDS	2175331	2172950	-1	-	2382	Type II secretory pathway, AIPase Pule/Tip pilus assembly pathway, AIPase PilB pathway, AIPase PilC	-none-	D23_1c2335	Neut_2128
fig/66666666.60966.peg.2341	CDS	2176108	2175344	-1	-	765	CAMP phosphodiesterases class-II:Metallo-beta-lactamase superfamily	cAMP signaling in bacteria	D23_1c2336	Neut_2129
fig/66666666.60966.peg.2342	CDS	2176622	2176197	-2	-	426	Universal stress protein (Usp)	-none-	D23_1c2337	Neut_2130
fig/66666666.60966.peg.2343	CDS	2177459	2176734	-2	-	726	Membrane protein TerC, possibly involved in tellurium resistance	-none-	D23_1c2338	Neut_2131
fig/66666666.60966.peg.2344	CDS	2178386	2177462	-1	-	1125	Patatin	-none-	D23_1c2339	Neut_2132
fig/66666666.60966.peg.2345	CDS	2178900	2179850	3	+	951	Proline iminopeptidase (EC 3.4.11.5)	Proline, 4-hydroxyproline uptake and utilization	D23_1c2340	Neut_2133
fig/66666666.60966.peg.2346	CDS	2180553	2179870	-3	-	684	Dethiobiotin synthetase (EC 6.3.3.3)	Biotin biosynthesis; Biotin biosynthesis Experimental; Biotin synthesis cluster	D23_1c2341	Neut_2134
fig/66666666.60966.peg.2347	CDS	2181452	2180556	-2	-	897	Biotin synthesis protein BioC	 Biotin biosynthesis Experimental; Biotin synthesis cluster	D23_1c2342	Neut_2135
fig/66666666.60966.peg.2348	CDS	2182200	2181442	-3	-	759	Biotin synthesis protein BioH	 Biotin biosynthesis Experimental; Biotin synthesis cluster	D23_1c2343	Neut_2136

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.									
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld C91 Alias
fig/66666666.60966.peg.2349	CDS	2183365	2182202	-1	-	1164	8-amino-7-oxononanoate synthase (EC 2.3.1.47)	Biotin biosynthesis; Biotin biosynthesis Experimental; Biotin synthesis cluster	D23_1c2344 Neut_2137
fig/66666666.60966.peg.2350	CDS	2184393	2183371	-3	-	1023	Biotin synthase (EC 2.8.1.6)	Biotin biosynthesis; Biotin biosynthesis Experimental; Biotin synthesis cluster	D23_1c2345 Neut_2138
fig/66666666.60966.peg.2351	CDS	2184641	2184453	-2	-	189	hypothetical protein	-none-	D23_1c2346 NA
fig/66666666.60966.peg.2352	CDS	2184698	2185168	2	+	471	Competence protein F homolog, phosphoribosyltransferase domain; protein YhgH required for utilization of DNA as sole source of carbon and energy	Biotin biosynthesis Experimental; Biotin synthesis cluster; CBSS-216591.1.peg.168	D23_1c2347 Neut_2139
fig/66666666.60966.peg.2353	CDS	2185222	2185686	1	+	465	tRNA (cytidine(34)-2'-O)-methyltransferase (EC 2.1.1.207) ## TrmL protein of unknown function DUF1130 SAM-dependent methyltransferase (EC 2.1.1.-)	Biotin synthesis cluster; RNA methylation	D23_1c2348 Neut_2140
fig/66666666.60966.peg.2354	CDS	2185773	2186303	3	+	531	protein of unknown function DUF1130	-none-	D23_1c2349 Neut_2141
fig/66666666.60966.peg.2355	CDS	2186489	2187325	2	+	837	SAM-dependent methyltransferase (EC 2.1.1.-)	-none-	D23_1c2350 Neut_2142
fig/66666666.60966.peg.2356	CDS	2189013	2187523	-3	-	1491	Glucose-6-phosphate 1-dehydrogenase (EC 1.1.1.49)	Pentose phosphate pathway	D23_1c2351 Neut_2143
fig/66666666.60966.peg.2357	CDS	2189925	2189017	-3	-	909	6-phosphogluconate dehydrogenase, decarboxylating (EC 1.1.1.44)	D-gluconate and ketogluconates metabolism; Pentose phosphate pathway	D23_1c2352 Neut_2144
fig/66666666.60966.peg.2358	CDS	2191152	2189995	-3	-	1158	COGs COG1397	-none-	D23_1c2353 Neut_2145
fig/66666666.60966.peg.2359	CDS	2191494	2191195	-3	-	300	UPF0235 protein VC0458	CBSS-630.2.peg.3360	D23_1c2354 Neut_2146
fig/66666666.60966.peg.2360	CDS	2192054	2191494	-2	-	561	Integral membrane protein YggT, involved in response to extracytoplasmic stress (osmotic shock)	CBSS-630.2.peg.3360	D23_1c2355 Neut_2147
fig/66666666.60966.peg.2361	CDS	2192939	2192127	-2	-	813	Pyrroline-5-carboxylate reductase (EC 1.5.1.2)	A Hypothetical Protein Related to Proline Metabolism; CBSS-630.2.peg.3360; Proline Synthesis	D23_1c2356 Neut_2148

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld	C91 Alias
fig/66666666.60966.peg.2362	CDS	2193210	2193019	-3	-	192	FIG00859708: hypothetical protein	-none-	D23_1c2357	Neut_2149
fig/66666666.60966.peg.2364	CDS	2194302	2194580	3	+	279	Ribonuclease P protein component (EC 3.1.26.5)	Cell Division Subsystem including YidCD; RNA modification cluster; RNA processing	D23_1c2359	Neut_2152
fig/66666666.60966.peg.2365	CDS	2194877	2196745	2	+	1869	Inner membrane protein translocase component YidC, long form	CTP synthase (EC 6.3.4.2) cluster; Cell Division Subsystem including YidCD; RNA modification cluster	D23_1c2360	Neut_2154
fig/66666666.60966.peg.2366	CDS	2196792	2198171	3	+	1380	GTPase and tRNA-U34 5-formylation enzyme TrmE	Cell Division Subsystem including YidCD; RNA modification and chromosome partitioning cluster; RNA modification cluster; Universal GTPases; mm5U34 biosynthesis bacteria; RNA modification	D23_1c2361	Neut_2155
fig/66666666.60966.peg.2367	CDS	2198398	2198667	1	+	270	SSU ribosomal protein S15p (S13e)	Bacteria	D23_1c2363	Neut_2156
fig/66666666.60966.peg.2368	CDS	2198846	2200963	2	+	2118	Polyribonucleotide nucleotidyltransferase (EC 2.7.7.8)	-none-	D23_1c2364	Neut_2157
fig/66666666.60966.peg.2369	CDS	2201167	2201039	-1	-	129	hypothetical protein	Bacterial RNA-metabolizing Zn-dependent hydrolases; Polyadenylation bacterial	D23_1c2365	NA
fig/66666666.60966.peg.2371	CDS	2201553	2202926	3	+	1374	RND efflux system, outer membrane lipoprotein CmeC	Multidrug Resistance Efflux Pumps	D23_1c2367	Neut_2158
fig/66666666.60966.peg.2372	CDS	2203015	2204220	1	+	1206	Membrane fusion protein of RND family multidrug efflux pump	Multidrug Resistance Efflux Pumps	D23_1c2368	Neut_2159
fig/66666666.60966.peg.2373	CDS	2204238	2207432	3	+	3195	RND efflux system, inner membrane transporter CmeB	Multidrug Resistance Efflux Pumps	D23_1c2369	Neut_2160
fig/66666666.60966.peg.2375	CDS	2208043	2207702	-1	-	342	hypothetical protein	-none-	D23_1c2370	Neut_2161
fig/66666666.60966.peg.2376	CDS	2208812	2208072	-2	-	741	conserved hypothetical protein [<i>Pyrococcus horikoshii</i>]; COG2102: Predicted ATPases of PP-loop superfamily; IPR002761: Domain of	-none-	D23_1c2371	Neut_2162

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.									
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld C91 Alias
fig 66666666.60966.peg.2377	CDS	2209043	2208924	-2	-	120	unknown function DUF71	-none-	D23_1c2372 Neut_2163
fig 66666666.60966.peg.2379	CDS	2210878	2209472	-1	-	1407	FIG00859479: hypothetical protein GTP-binding protein EngA	CBSS- 290633.1.peg.1906; CBSS- 498211.3.peg.1415; Universal GTPases CBSS- 290633.1.peg.1906; CBSS- 498211.3.peg.1415; Lipopolysaccharide assembly	D23_1c2374 Neut_2164
fig 66666666.60966.peg.2380	CDS	2212159	2210930	-1	-	1230	Outer membrane protein YfgL, lipoprotein component of the protein assembly complex (forms a complex with YaeI, YfiO, and NlpB)	CBSS- 290633.1.peg.1906; CBSS- 498211.3.peg.1415; RNA aminoacylation, His	D23_1c2375 Neut_2165
fig 66666666.60966.peg.2381	CDS	2212800	2212162	-3	-	639	Mir7403 protein	CBSS- 290633.1.peg.1906; CBSS- 498211.3.peg.1415	D23_1c2376 Neut_2166
fig 66666666.60966.peg.2382	CDS	2214088	2212823	-1	-	1266	Histidinyl-tRNA synthetase (EC 6.1.1.21)	CBSS- 498211.3.peg.1415; RNA aminoacylation, His	D23_1c2377 Neut_2167
fig 66666666.60966.peg.2383	CDS	2215286	2214081	-2	-	1206	1-hydroxy-2-methyl-2- (E)-butenyl 4- diphosphate synthase (EC 1.1.7.1)	CBSS- 498211.3.peg.1415; CBSS- 8333.1.peg.3039; Isoprenoid Biosynthesis; Nonmevalonate Branch of Isoprenoid Biosynthesis CBSS-498211.3.peg.1415	D23_1c2378 Neut_2168
fig 66666666.60966.peg.2384	CDS	2216438	2215359	-2	-	1080	FIG021952: putative membrane protein	CBSS-498211.3.peg.1415	D23_1c2379 Neut_2169
fig 66666666.60966.peg.2385	CDS	2217306	2216428	-3	-	879	Type IV pilus biogenesis protein PilF	CBSS-498211.3.peg.1415	D23_1c2380 Neut_2170
fig 66666666.60966.peg.2386	CDS	2218414	2217275	-1	-	1140	Ribosomal RNA large subunit methyltransferase N (EC 2.1.1.—)	CBSS- 498211.3.peg.1415; RNA methylation	D23_1c2381 Neut_2171
fig 66666666.60966.peg.2387	CDS	2218877	2218452	-2	-	426	Nucleoside diphosphate kinase (EC 2.7.4.6)	CBSS- 498211.3.peg.1415; Purine conversions; pyrimidine conversions -none-	D23_1c2382 Neut_2172
fig 66666666.60966.peg.2389	CDS	2219591	2219175	-2	-	417	Chain A, Red Copper Protein Nitrosocyanin		D23_1c2383 Neut_2173

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld	C91 Alias
fig/66666666.60966.peg.2390	CDS	2220042	2221430	3	+	1389	Aminotransferase HpnO, required for aminobacteriohopanetriol	Hopanes	D23_1c2384	Neut_2174
fig/66666666.60966.peg.2391	CDS	2222146	2221442	-1	-	705	DNA polymerase III epsilon subunit (EC 2.7.7.7)	CBSS-228410.1.peg.134; CBSS-342610.3.peg.1536	D23_1c2385	Neut_2175
fig/66666666.60966.peg.2392	CDS	2222646	2222158	-3	-	489	Ribonuclease HI (EC 3.1.26.4)	CBSS-228410.1.peg.134; CBSS-342610.3.peg.1536; Ribonuclease H	D23_1c2386	Neut_2176
fig/66666666.60966.peg.2393	CDS	2223440	2222706	-2	-	735	FIG005121: SAM-dependent methyltransferase (EC 2.1.1.—)	CBSS-228410.1.peg.134; CBSS-342610.3.peg.1536; Glutathione: Non-redox reactions	D23_1c2387	Neut_2177
fig/66666666.60966.peg.2394	CDS	2223500	2224267	2	+	768	Hydroxyacylglutathione hydrolase (EC 3.1.2.6)	CBSS-228410.1.peg.134; CBSS-342610.3.peg.1536; Glutathione: Non-redox reactions; Methylglyoxal Metabolism	D23_1c2388	Neut_2178
fig/66666666.60966.peg.2396	CDS	2225143	2224571	-1	-	573	hypothetical protein	-none-	D23_1c2389	Neut_2179
fig/66666666.60966.peg.2397	CDS	2225287	2225162	-1	-	126	hypothetical protein	-none-	D23_1c2390	NA
fig/66666666.60966.peg.2398	CDS	2226167	2225250	-2	-	918	hypothetical protein	-none-	D23_1c2391	Neut_2180
fig/66666666.60966.peg.2399	CDS	2228864	2226171	-2	-	2694	Helicase, SNF2/RAD54 family	-none-	D23_1c2392	Neut_2181
fig/66666666.60966.peg.2401	CDS	2229650	2229480	-2	-	171	hypothetical protein	-none-	D23_1c2393	NA
fig/66666666.60966.peg.2402	CDS	2229693	2230814	3	+	1122	Oxidoreductase, FMN-binding	-none-	D23_1c2394	Neut_2182
fig/66666666.60966.peg.2403	CDS	2230925	2231368	2	+	444	Ornithine cyclodeaminase (EC 4.3.1.12)	Arginine and Ornithine Degradation	D23_1c2395	Neut_2183
fig/66666666.60966.peg.2404	CDS	2231434	2233005	1	+	1572	Phosphoglucomutase (EC 5.4.2.2)	-none-	D23_1c2396	Neut_2184
fig/66666666.60966.peg.2405	CDS	2233192	2233344	1	+	153	hypothetical protein	-none-	D23_1c2397	NA
fig/66666666.60966.peg.2406	CDS	2234195	2233347	-2	-	849	Mobile element protein	-none-	D23_1c2398	Neut_1888
fig/66666666.60966.peg.2407	CDS	2234437	2234252	-1	-	186	Mobile element protein	-none-	D23_1c2399	Neut_2500
fig/66666666.60966.peg.2408	CDS	2234636	2235049	2	+	414	Cobalt-zinc-cadmium resistance protein	Cobalt-zinc-cadmium resistance	D23_1c2400	Neut_2185
fig/66666666.60966.peg.2409	CDS	2235949	2235359	-1	-	591	Hydroxyacylglutathione hydrolase (EC 3.1.2.6)	CBSS-228410.1.peg.134; CBSS-342610.3.peg.1536; Glutathione: Non-redox reactions; Methylglyoxal Metabolism	D23_1c2401	Neut_2178
fig/66666666.60966.peg.2410	CDS	2236135	2236287	1	+	153	hypothetical protein	-none-	D23_1c2403	Neut_1255

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld	C91 Alias
fig/66666666.60966.peg.2411	CDS	2236295	2236462	2	+	168	hypothetical protein	-none-	D23_1c2404	Neut_2449
fig/66666666.60966.peg.2412	CDS	2236616	2236419	-2	-	198	Mobile element protein	-none-	D23_1c2405	Neut_2268
fig/66666666.60966.peg.2414	CDS	2236758	2237006	3	+	249	Mobile element protein	-none-	D23_1c2406	Neut_1756
fig/66666666.60966.peg.2417	CDS	2240037	2237515	-3	-	2523	Aconitate hydratase (EC 4.2.1.3)	TCA Cycle	D23_1c2409	Neut_2457
fig/66666666.60966.peg.2418	CDS	2240239	2240093	-1	-	147	Aconitate hydratase (EC 4.2.1.3)	TCA Cycle	D23_1c2410	NA
fig/66666666.60966.peg.2419	CDS	2240661	2241623	3	+	963	Mobile element protein	-none-	D23_1c2411	Neut_1278
fig/66666666.60966.peg.2421	CDS	2241835	2241987	1	+	153	Mobile element protein	-none-	D23_1c2413	NA
fig/66666666.60966.peg.2422	CDS	2242232	2242522	2	+	291	hypothetical protein	-none-	D23_1c2414	NA
fig/66666666.60966.peg.2423	CDS	2242506	2243450	3	+	945	Agmatinase (EC 3.5.3.11)	Arginine and Ornithine Degradation; Polyamine Metabolism	D23_1c2415	Neut_2187
fig/66666666.60966.peg.2424	CDS	2243462	2245420	2	+	1959	Biosynthetic arginine decarboxylase (EC 4.1.1.19)	Arginine and Ornithine Degradation; Polyamine Metabolism	D23_1c2416	Neut_2188
fig/66666666.60966.peg.2425	CDS	2246086	2245436	-1	-	651	Mobile element protein	-none-	D23_1c2417	Neut_2189
fig/66666666.60966.peg.2426	CDS	2247251	2246148	-2	-	1104	hypothetical protein	-none-	D23_1c2418	Neut_2191
fig/66666666.60966.peg.2427	CDS	2247251	2247436	2	+	186	Mobile element protein	-none-	D23_1c2419	Neut_2500
fig/66666666.60966.peg.2428	CDS	2247493	2248341	1	+	849	Mobile element protein	-none-	D23_1c2420	Neut_1375
fig/66666666.60966.peg.2430	CDS	2249188	2249352	1	+	165	Mobile element protein	-none-	D23_1c2422	Neut_2186
fig/66666666.60966.peg.2431	CDS	2249749	2251734	1	+	1986	Protein-L-isopartate O-methyltransferase (EC 2.1.1.77)	Protein-L-isopartate O-methyltransferase; Stationary phase repair cluster; Ton and Tol transport systems	D23_1c2423	Neut_2323
fig/66666666.60966.peg.2432	CDS	2252354	2253127	2	+	774	hypothetical protein	-none-	D23_1c2424	Neut_2196
fig/66666666.60966.peg.2433	CDS	2253955	2253515	-1	-	441	tRNA pseudouridine synthase A (EC 4.2.1.70)	Colicin V and Bacteriocin Production Cluster; RNA pseudouridine syntheses; tRNA modification Bacteria; tRNA processing	D23_1c2425	Neut_2203
fig/66666666.60966.peg.2434	CDS	2254408	2254247	-1	-	162	hypothetical protein	-none-	D23_1c2426	NA
fig/66666666.60966.peg.2435	CDS	2256251	2254536	-2	-	1716	Selenoprotein O and cysteine-containing homologs	Selenoprotein O	D23_1c2427	Neut_2204
fig/66666666.60966.peg.2436	CDS	2256898	2257503	1	+	606	GCN5-related N-acetyltransferase	-none-	D23_1c2429	Neut_2208
fig/66666666.60966.peg.2437	CDS	2258004	2257552	-3	-	453	probable multiple antibiotic resistance protein marC	-none-	D23_1c2430	Neut_2211
fig/66666666.60966.peg.2438	CDS	2258357	2258229	-2	-	129	hypothetical protein	-none-	D23_1c2431	Neut_2212
fig/66666666.60966.peg.2440	CDS	2259106	2259264	1	+	159	hypothetical protein	-none-	D23_1c2432	NA

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.									
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld C91 Alias
fig/66666666.60966.peg.2441	CDS	2259280	2260221	1	+	942	Ornithine cyclodeaminase (EC 4.3.1.12)	Arginine and Ornithine Degradation	D23_1c2433 Neut_2213
fig/66666666.60966.peg.2442	CDS	2260562	2261791	2	+	1230	hypothetical protein	-none-	D23_1c2434 Neut_2214
fig/66666666.60966.peg.2443	CDS	2262110	2262748	2	+	639	hypothetical protein	-none-	D23_1c2435 Neut_2232
fig/66666666.60966.peg.2444	CDS	2262936	2263406	3	+	471	FIG00858867: hypothetical protein	-none-	D23_1c2436 Neut_2233
fig/66666666.60966.peg.2446	CDS	2264766	2263630	-3	-	1137	N-succinyl-L,L-diaminopimelate desuccinylase (EC 3.5.1.18)	Arginine Biosynthesis--g/o; Arginine Biosynthesis extended; Lysine Biosynthesis DAP Pathway, GIO scratch	D23_1c2437 Neut_2234
fig/66666666.60966.peg.2447	CDS	2265597	2264815	-3	-	783	Methionine ABC transporter ATP-binding protein	Methionine Biosynthesis; Methionine Degradation	D23_1c2438 Neut_2235
fig/66666666.60966.peg.2448	CDS	2266742	2265597	-2	-	1146	ABC-type transport system involved in resistance to organic solvents, permease component USSDB6A	-none-	D23_1c2439 Neut_2236
fig/66666666.60966.peg.2449	CDS	2267603	2266767	-2	-	837	Prolipoprotein diacylglycerol transferase (EC 2.4.99.-)	Lipoprotein Biosynthesis	D23_1c2440 Neut_2237
fig/66666666.60966.peg.2450	CDS	2267740	2269413	1	+	1674	Dihydroxy-acid dehydratase (EC 4.2.1.9)	Branched-Chain Amino Acid Biosynthesis	D23_1c2441 Neut_2238
fig/66666666.60966.peg.2451	CDS	2269430	2271499	2	+	2070	Thymidylate kinase (EC 2.7.4.9)	pyrimidine conversions	D23_1c2442 Neut_2241
fig/66666666.60966.peg.2452	CDS	2271523	2271834	1	+	312	Cytochrome c551/c552	Soluble cytochromes and functionally related electron carriers	D23_1c2443 Neut_2242
fig/66666666.60966.peg.2454	CDS	2272004	2272972	2	+	969	D-3-phosphoglycerate dehydrogenase (EC 1.1.1.95)	Glycine and Serine Utilization; Pyridoxin (Vitamin B6) Biosynthesis; Serine Biosynthesis	D23_1c2444 Neut_2243
fig/66666666.60966.peg.2455	CDS	2273050	2273595	1	+	546	dTDP-4-dehydrothiamine 3,5-epimerase (EC 5.1.3.13)	Rhamnose containing glycans; dTDP-rhamnose synthesis	D23_1c2445 Neut_2244
fig/66666666.60966.peg.2456	CDS	2273700	2274983	3	+	1284	Permeases of the major facilitator superfamily	-none-	D23_1c2446 Neut_2245
fig/66666666.60966.peg.2457	CDS	2275758	2274997	-3	-	762	hypothetical protein	-none-	D23_1c2447 NA
fig/66666666.60966.peg.2458	CDS	2275968	2276633	3	+	666	Chromosomal replication initiator protein DnaA	Cell Division Subsystem including YidCD; DNA replication cluster 1	D23_1c2449 Neut_2247

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld	C91 Alias
fig/66666666.60966.peg.2459	CDS	2276770	2277435	1	+	666	Phosphoserine phosphatase (EC 3.1.3.3)	Glycine and Serine Utilization; Serine Biosynthesis; Serine Biosynthesis	D23_1c2450	Neut_2248
fig/66666666.60966.peg.2460	CDS	2277481	2277846	1	+	366	FIG00859424; hypothetical protein	-none-	D23_1c2451	Neut_2249
fig/66666666.60966.peg.2461	CDS	2277800	2277958	2	+	159	hypothetical protein	-none-	D23_1c2452	NA
fig/66666666.60966.peg.2462	CDS	2277971	2279434	2	+	1464	Inosine-5'-monophosphate dehydrogenase (EC 1.1.1.205)	Purine conversions; Purine salvage cluster	D23_1c2453	Neut_2250
fig/66666666.60966.peg.2463	CDS	2279447	2281006	2	+	1560	GMP synthase [glutamine-hydrolyzing], amidotransferase subunit (EC 6.3.5.2)/GMP synthase [glutamine-hydrolyzing], ATP pyrophosphatase subunit (EC 6.3.5.2)	GMP synthase; Purine conversions; Purine conversions; Purine salvage cluster; Purine salvage cluster	D23_1c2454	Neut_2251
fig/66666666.60966.peg.2465	CDS	2282716	2281352	-1	-	1365	hypothetical protein	-none-	D23_1c2455	NA
fig/66666666.60966.peg.2466	CDS	2282899	2282762	-1	-	138	Mobile element protein	-none-	D23_1c2456	Neut_2501
fig/66666666.60966.peg.2467	CDS	2283609	2282872	-3	-	738	Mobile element protein	-none-	D23_1c2457	Neut_1888
fig/66666666.60966.peg.2468	CDS	2283851	2283666	-2	-	186	Mobile element protein	-none-	D23_1c2458	Neut_2500
fig/66666666.60966.peg.2469	CDS	2285015	2284020	-2	-	996	hypothetical protein	-none-	D23_1c2459	NA
fig/66666666.60966.peg.2470	CDS	2285305	2285048	-1	-	258	hypothetical protein	-none-	D23_1c2460	NA
fig/66666666.60966.peg.2471	CDS	2285448	2285573	3	+	126	hypothetical protein	-none-	D23_1c2461	NA
fig/66666666.60966.peg.2472	CDS	2285631	2286293	3	+	663	Thiopurine S-methyltransferase (EC 2.1.1.67)	-none-	D23_1c2462	Neut_2272
fig/66666666.60966.peg.2473	CDS	2286615	2288795	3	+	2181	hypothetical protein	-none-	D23_1c2463	Neut_2273
fig/66666666.60966.peg.2474	CDS	2289016	2291046	1	+	2031	oligopeptide transporter	-none-	D23_1c2464	Neut_2274
fig/66666666.60966.peg.2475	CDS	2291342	2291145	-2	-	198	FIG00859558; hypothetical protein	-none-	D23_1c2465	Neut_2275
fig/66666666.60966.peg.2476	CDS	2291757	2291362	-3	-	396	FIG00859558; hypothetical protein	-none-	D23_1c2466	Neut_2275
fig/66666666.60966.peg.2477	CDS	2292959	2291901	-2	-	1059	Putative permease often clustered with de novo purine synthesis	-none-	D23_1c2467	Neut_2276
fig/66666666.60966.peg.2478	CDS	2293039	2294097	1	+	1059	Phosphoribosylformylglycinamide cyclo-ligase (EC 6.3.3.1)	De Novo Purine Biosynthesis	D23_1c2468	Neut_2277
fig/66666666.60966.peg.2479	CDS	2294109	2294741	3	+	633	Phosphoribosylglycinamide formyltransferase (EC 2.1.2.2)	5-FCL-like protein; De Novo Purine Biosynthesis	D23_1c2469	Neut_2278

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld	C91 Alias
fig/66666666.60966.peg.2480	CDS	2294738	2295460	2	+	723	FIG00859545; hypothetical protein	-none-	D23_1c2470	Neut_2279
fig/66666666.60966.peg.2481	CDS	2295477	2296751	3	+	1275	Fmt (Sum) eukaryotic nucleolar NOL1/Nop2p; tRNA and rRNA	-none-	D23_1c2471	Neut_2280
fig/66666666.60966.peg.2482	CDS	2297068	2296832	-1	-	237	cytosine-C5-methylases hypothetical protein PA0941	-none-	D23_1c2472	Neut_2281
fig/66666666.60966.peg.2483	CDS	2297785	2297237	-1	-	549	InterPro IPR000694;IPR001734	-none-	D23_1c2473	Neut_2282
fig/66666666.60966.peg.2484	CDS	2297997	2297800	-3	-	198	hypothetical protein	-none-	D23_1c2474	Neut_2283
fig/66666666.60966.peg.2485	CDS	2298574	2299293	1	+	720	possible transmembrane protein	-none-	D23_1c2475	Neut_2284
fig/66666666.60966.peg.2486	CDS	2299389	2301113	3	+	1725	Diguanylate cyclase/phosphodiesterase domain 2 (EAL)	-none-	D23_1c2476	Neut_2285
fig/66666666.60966.peg.2488	CDS	2301326	2301784	2	+	459	FKBP-type peptidyl-prolyl cis-trans isomerase	-none-	D23_1c2477	Neut_2286
fig/66666666.60966.peg.2489	CDS	2302743	2301856	-3	-	888	Ribosome small subunit-stimulated GTPase EngC	Universal GTPases	D23_1c2478	Neut_2287
fig/66666666.60966.peg.2490	CDS	2303060	2302782	-2	-	279	Pterin-4-alpha-carbinolamine dehydratase (EC 4.2.1.96)	Pterin carbinolamine dehydratase	D23_1c2479	Neut_2288
fig/66666666.60966.peg.2491	CDS	2304388	2303120	-1	-	1269	macromolecule metabolism; macromolecule degradation; degradation of proteins, peptides, glycopeptides	-none-	D23_1c2480	Neut_2289
fig/66666666.60966.peg.2492	CDS	2304540	2305085	3	+	546	3'-to-5' oligoribonuclease (orn)	RNA processing and degradation, bacterial	D23_1c2481	Neut_2290
fig/66666666.60966.peg.2493	CDS	2305123	2307678	1	+	2556	Glycogen phosphorylase (EC 2.4.1.1)	Glycogen metabolism	D23_1c2482	Neut_2291
fig/66666666.60966.peg.2494	CDS	2308460	2307702	-2	-	759	Pantoate-beta-alanine ligase (EC 6.3.2.1)	Coenzyme A Biosynthesis; Coenzyme A	D23_1c2483	Neut_2292
fig/66666666.60966.peg.2495	CDS	2309368	2308559	-1	-	810	3-methyl-2-oxobutanoate hydroxymethyltransferase (EC 2.1.2.11)	Biosynthesis cluster	D23_1c2484	Neut_2293
fig/66666666.60966.peg.2496	CDS	2310016	2309372	-1	-	645	Deoxyadenosine kinase (EC 2.7.1.76)	Biosynthesis cluster Purine conversions; Purine conversions	D23_1c2485	Neut_2294

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.									
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld C91 Alias
fig/66666666.609666.peg.2497	CDS	2310525	2310013	-3	-	513	Deoxyguanosine Kinase (EC 2.7.1.113) 2-amino-4-hydroxy-6-hydroxymethylidihydropteridine pyrophosphokinase (EC 2.7.6.3)	Folate Biosynthesis	D23_1c2486 Neut_2295
fig/66666666.609666.peg.2498	CDS	2311910	2310522	-2	-	1389	Poly(A) polymerase (EC 2.7.7.19)	Polyadenylation bacterial	D23_1c2487 Neut_2296
fig/66666666.609666.peg.2499	CDS	2313230	2312073	-2	-	1158	Cardiolipin synthetase (EC 2.7.8.—)	Cardiolipin synthesis; Glycerophospholipid Metabolism in Bacteria -none-	D23_1c2488 Neut_2297
fig/66666666.609666.peg.2500	CDS	2314054	2313227	-1	-	828	Endonuclease/exonuclease/phosphatase family protein		D23_1c2489 Neut_2298
fig/66666666.609666.peg.2501	CDS	2315160	2314060	-3	-	1101	Quinolinate synthetase (EC 2.5.1.72)	<i>Mycobacterium</i> virulence operon possibly involved in quinolate biosynthesis; NAD and NADP cofactor biosynthesis global -none-	D23_1c2490 Neut_2299
fig/66666666.609666.peg.2502	CDS	2315495	2316031	2	+	537	FIG00859627: hypothetical protein		D23_1c2491 Neut_2300
fig/66666666.609666.peg.2503	CDS	2316137	2316652	2	+	516	LptA, protein essential for LPS transport across the periplasm	Lipopolysaccharide assembly	D23_1c2492 Neut_2301
fig/66666666.609666.peg.2504	CDS	2316704	2317426	2	+	723	Lipopolysaccharide ABC transporter, ATP-binding protein LptB	Lipopolysaccharide assembly	D23_1c2493 Neut_2302
fig/66666666.609666.peg.2505	CDS	2317432	2318895	1	+	1464	RNA polymerase sigma-54 factor RpoN	Flagellar motility; Flagellum; Transcription initiation, bacterial sigma factors	D23_1c2494 Neut_2303
fig/66666666.609666.peg.2506	CDS	2319070	2319405	1	+	336	Ribosome hibernation protein YhbH	Ribosome activity modulation -none-	D23_1c2495 Neut_2304
fig/66666666.609666.peg.2507	CDS	2319646	2320116	1	+	471	PTS system nitrogen-specific IIA component, PtsN		D23_1c2496 Neut_2305
fig/66666666.609666.peg.2508	CDS	2320103	2321074	2	+	972	HPt kinase/phosphorylase (EC 2.7.1.—) (EC 2.7.4.—)	HPt catabolite repression system	D23_1c2497 Neut_2306
fig/66666666.609666.peg.2509	CDS	2321107	2321223	1	+	117	hypothetical protein	-none-	D23_1c2498 NA
fig/66666666.609666.peg.2510	CDS	2321260	2321874	1	+	615	3-polyphenyl-4-hydroxybenzoate carboxy-lyase UbiX (EC 4.1.1.—)	Ubiquinone Biosynthesis; Ubiquinone Biosynthesis-gio	D23_1c2499 Neut_2307

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 GbkId	C91 Alias
fig/66666666.60966.peg.2511	CDS	2322562	2321924	-1	-	639	5- formyltetrahydrofolate cyclo-ligase (EC 6.3.3.2)	5-FL-like protein; Folate Biosynthesis; One-carbon metabolism by tetrahydropterines	D23_1c2500	Neut_2308
fig/66666666.60966.peg.2512	CDS	2323682	2322555	-2	-	1128	A/G-specific adenine glycosylase (EC 3.2.2.-)	DNA repair, bacterial	D23_1c2501	Neut_2309
fig/66666666.60966.peg.2513	CDS	2324249	2323704	-2	-	546	Intracellular septation protein IspA	CBSS-211586.9.peg.2729	D23_1c2503	Neut_2310
fig/66666666.60966.peg.2514	CDS	2326043	2324346	-2	-	1698	Lipid A export ATP-binding/permease protein MsbA (EC 3.6.3.25)	KDO2-Lipid A biosynthesis cluster 2	D23_1c2504	Neut_2311
fig/66666666.60966.peg.2515	CDS	2327134	2326562	-1	-	573	hypothetical protein	-none-	D23_1c2505	Neut_2312
fig/66666666.60966.peg.2516	CDS	2329375	2327231	-1	-	2145	Copper resistance protein D	Copper homeostasis	D23_1c2506	Neut_2313
fig/66666666.60966.peg.2517	CDS	2329755	2329381	-3	-	375	Copper resistance protein CopC precursor	-none-	D23_1c2507	Neut_2314
fig/66666666.60966.peg.2518	CDS	2330496	2329909	-3	-	588	hypothetical protein	-none-	D23_1c2508	Neut_2074
fig/66666666.60966.peg.2519	CDS	2331308	2330568	-2	-	741	putative (U92432) ORF4 (<i>Vitrosospira</i> sp. NpAV)	-none-	D23_1c2509	Neut_2316
fig/66666666.60966.peg.2520	CDS	2332640	2331375	-2	-	1266	Particulate methane monooxygenase B-subunit (EC 1.14.13.25)	Particulate methane monooxygenase (pMMO)	D23_1c2510	Neut_2317
fig/66666666.60966.peg.2521	CDS	2333470	2332640	-1	-	831	Particulate methane monooxygenase A-subunit (EC 1.14.13.25)	Particulate methane monooxygenase (pMMO)	D23_1c2511	Neut_2318
fig/66666666.60966.peg.2522	CDS	2334459	2333644	-3	-	816	Particulate methane monooxygenase C-subunit (EC 1.14.13.25)	Particulate methane monooxygenase (pMMO)	D23_1c2512	Neut_2319
fig/66666666.60966.peg.2523	CDS	2334883	2335908	1	+	1026	TsaC protein (YrdC-Sua5 domains) required for threonylcarbamoyladenosine t(6)A37 modification in tRNA	-none-	D23_1c2513	Neut_2320
fig/66666666.60966.peg.2524	CDS	2335923	2336666	3	+	744	Hypothetical protein CbbY	-none-	D23_1c2514	Neut_2321
fig/66666666.60966.peg.2525	CDS	2337688	2336702	-1	-	987	Lipoprotein NlpD	Stationary phase repair cluster	D23_1c2515	Neut_2322
fig/66666666.60966.peg.2526	CDS	2338471	2337803	-1	-	669	Protein-L-isospartate O-methyltransferase (EC 2.1.1.77)	Protein-L-isospartate O-methyltransferase; Stationary phase repair cluster; Ton and Tol transport systems	D23_1c2516	Neut_2323
fig/66666666.60966.peg.2528	CDS	2339405	2338662	-2	-	744	5-nucleotidase SurE (EC 3.1.3.5) @	Housecleaning nucleoside triphosphate	D23_1c2517	Neut_2324

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 GbkId	C91 Alias
							Exopolysphatase (EC 3.6.1.11)	pyrophosphatases; Phosphate metabolism; Polypolysphate; Stationary phase repair cluster		
fig/66666666.60966.peg.2529	CDS	2340272	2339964	-2	-	309	Integration host factor alpha subunit	DNA structural proteins, bacterial	D23_1c2519	Neut_2325
fig/66666666.60966.peg.2530	CDS	2342785	2340485	-1	-	2301	Phenylalanyl-tRNA synthetase beta chain (EC 6.1.1.20)	tRNA aminoacylation, Phe	D23_1c2520	Neut_2326
fig/66666666.60966.peg.2531	CDS	2343901	2342882	-1	-	1020	Phenylalanyl-tRNA synthetase alpha chain (EC 6.1.1.20)	tRNA aminoacylation, Phe	D23_1c2521	Neut_2327
fig/66666666.60966.peg.2532	CDS	2344231	2343989	-1	-	243	LSU ribosomal protein L20p	-none-	D23_1c2522	Neut_2328
fig/66666666.60966.peg.2533	CDS	2345202	2344723	-3	-	480	Translation initiation factor 3	Translation initiation factors bacterial	D23_1c2524	Neut_2330
fig/66666666.60966.peg.2534	CDS	2347209	2345302	-3	-	1908	Threonyl-tRNA synthetase (EC 6.1.1.3)	tRNA aminoacylation, Thr	D23_1c2525	Neut_2331
fig/66666666.60966.peg.2535	CDS	2348256	2347537	-3	-	720	Cytochrome c-type protein TorY	-none-	D23_1c2526	Neut_1790
fig/66666666.60966.peg.2536	CDS	2348966	2348259	-2	-	708	Cytochrome c family protein	-none-	D23_1c2527	Neut_2333
fig/66666666.60966.peg.2537	CDS	2350169	2349048	-2	-	1122	FIG00859557: hypothetical protein	-none-	D23_1c2528	Neut_1792
fig/66666666.60966.peg.2538	CDS	2351878	2350166	-1	-	1713	Hydroxylamine oxidoreductase precursor (EC 1.7.3.4)	-none-	D23_1c2529	Neut_2335
fig/66666666.60966.peg.2539	CDS	2352170	2353255	2	+	1086	tRNA-specific 2-thiouridylase MhmA	RNA methylation	D23_1c2530	Neut_2336
fig/66666666.60966.peg.2540	CDS	2354395	2353256	-1	-	1140	Twitching motility protein PilT	-none-	D23_1c2531	Neut_2337
fig/66666666.60966.peg.2541	CDS	2355448	2354405	-1	-	1044	Twitching motility protein PilT	-none-	D23_1c2532	Neut_2338
fig/66666666.60966.peg.2543	CDS	2355643	2356359	1	+	717	Hypothetical protein YggS, proline synthase co-transcribed bacterial	A Hypothetical Protein Related to Proline Metabolism; CBSS-	D23_1c2533	Neut_2339
fig/66666666.60966.peg.2544	CDS	2356597	2356328	-1	-	270	homolog PROSC 4Fe-4S ferredoxin, iron-sulfur binding	630.2.peg.3360 Inorganic Sulfur Assimilation	D23_1c2534	Neut_2340
fig/66666666.60966.peg.2545	CDS	2357097	2356603	-3	-	495	Phosphopantetheine adenylyltransferase (EC 2.7.7.3)	CBSS-266117.6.peg.1260; CBSS-269801.1.peg.1715; Coenzyme A Biosynthesis	D23_1c2535	Neut_2341
fig/66666666.60966.peg.2546	CDS	2357644	2357090	-1	-	555	16S rRNA (guanine966)-N(2))-	CBSS-266117.6.peg.1260;	D23_1c2536	Neut_2342

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 GbkId	C91 Alias
fig 66666666.60966.peg.2547	CDS	2358960	2357659	-3	-	1302	methy/transferase (EC 2.1.1.171) ## SSU rRNA m(2)G966	 CBSS-269801.1.peg.1715; Heat shock Cell division Proteases and a Methyltransferase; RNA methylation Heat shock Cell division Proteases and a Methyltransferase	D23_1c2537	Neut_2343
fig 66666666.60966.peg.2548	CDS	2359437	2359126	-3	-	312	Transposase	-none-	D23_1c2538	Neut_2344
fig 66666666.60966.peg.2549	CDS	2360187	2359819	-3	-	369	Mobile element protein	-none-	D23_1c2539	Neut_1814
fig 66666666.60966.peg.2550	CDS	2360570	2360247	-2	-	324	Putative periplasmic protein	-none-	D23_1c2540	Neut_2357
fig 66666666.60966.peg.2551	CDS	2361073	2360771	-1	-	303	hypothetical protein	-none-	D23_1c2541	Neut_2349
fig 66666666.60966.peg.2552	CDS	2361307	2361110	-1	-	198	CsBd family protein	-none-	D23_1c2542	Neut_2359
fig 66666666.60966.peg.2553	CDS	2361551	2361384	-2	-	168	protein of unknown function DUF1328	-none-	D23_1c2543	NA
fig 66666666.60966.peg.2554	CDS	2361863	2362066	2	+	204	Mobile element protein	-none-	D23_1c2544	Neut_2365
fig 66666666.60966.peg.2555	CDS	2362071	2362247	3	+	177	hypothetical protein	-none-	D23_1c2545	NA
fig 66666666.60966.peg.2556	CDS	2362394	2362957	2	+	564	Alkyl hydroperoxide reductase protein C (EC 1.6.4.—)	Thioredoxin-disulfide reductase	D23_1c2546	Neut_2366
fig 66666666.60966.peg.2557	CDS	2363052	2363633	3	+	582	putative lipoprotein	-none-	D23_1c2547	Neut_2367
fig 66666666.60966.peg.2558	CDS	2364440	2363661	-2	-	780	Putative stomatin/prohibitin-family membrane protease subunit aq_911	-none-	D23_1c2548	Neut_2368
fig 66666666.60966.peg.2559	CDS	2365869	2364442	-3	-	1428	Putative membrane-bound ClpP-class protease associated with aq_911	-none-	D23_1c2549	Neut_2369
fig 66666666.60966.peg.2560	CDS	2366482	2365880	-1	-	603	ADP-ribose pyrophosphatase (EC 3.6.1.13)	CBSS-216591.1.peg.168; NADand NADP cofactor biosynthesis global; Nudix proteins (nucleoside triphosphate hydrolases)	D23_1c2550	Neut_2370
fig 66666666.60966.peg.2561	CDS	2367579	2366665	-3	-	915	putative membrane protein	-none-	D23_1c2551	Neut_2371
fig 66666666.60966.peg.2562	CDS	2367673	2368539	1	+	867	Protein YicC	CBSS-323097.3.peg.2594	D23_1c2552	Neut_2372
fig 66666666.60966.peg.2563	CDS	2369652	2368633	-3	-	1020	C4-dicarboxylate transporter/malic acid transport protein	-none-	D23_1c2553	Neut_2373
fig 66666666.60966.peg.2564	CDS	2371015	2370053	-1	-	963	Mobile element protein	-none-	D23_1c2554	Neut_1746
fig 66666666.60966.peg.2565	CDS	2371724	2371065	-2	-	660	Glucose-1-phosphate thymidyltransferase (EC 2.7.7.24)	Rhamnose containing glycans; dUDP-rhamnose synthesis	D23_1c2555	Neut_2374

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.									
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld C91 Alias
fig/66666666.60966.peg.2566	CDS	2372737	2371739	-1	-	999	COG3178: Predicted phosphotransferase related to Ser/Thr protein kinases	-none-	D23_1c2556 Neut_2375
fig/66666666.60966.peg.2567	CDS	2372902	2373210	1	+	309	Putative cytoplasmic protein	-none-	D23_1c2557 Neut_2376
fig/66666666.60966.peg.2568	CDS	2373276	2374775	3	+	1500	MG(C+) CHELATASE FAMILY PROTEIN/ComM-related protein	-none-	D23_1c2559 Neut_2377
fig/66666666.60966.peg.2569	CDS	2376191	2374794	-2	-	1398	Replicative DNA helicase (EC 3.6.1.—)	-none-	D23_1c2560 Neut_2378
fig/66666666.60966.peg.2571	CDS	2376752	2376297	-2	-	456	LSU ribosomal protein L9p	Primosomal replication protein N clusters with ribosomal proteins	D23_1c2561 Neut_2379
fig/66666666.60966.peg.2572	CDS	2377049	2376771	-2	-	279	SSU ribosomal protein S18p @ SSU ribosomal protein S18p, zinc-independent	Primosomal replication protein N clusters with ribosomal proteins	D23_1c2562 Neut_2380
fig/66666666.60966.peg.2573	CDS	2377399	2377091	-1	-	309	Primosomal replication protein N	Primosomal replication protein N clusters with ribosomal proteins	D23_1c2563 Neut_2381
fig/66666666.60966.peg.2574	CDS	2377721	2377401	-2	-	321	SSU ribosomal protein S6p	Primosomal replication protein N clusters with ribosomal proteins	D23_1c2564 Neut_2382
fig/66666666.60966.peg.2576	CDS	2378765	2378340	-2	-	426	hypothetical protein	ribosomal proteins	D23_1c2565 NA
fig/66666666.60966.peg.2578	CDS	2379069	2380031	3	+	963	Mobile element protein	-none-	D23_1c2566 Neut_1746
fig/66666666.60966.peg.2580	CDS	2380409	2380864	2	+	456	Mobile element protein	-none-	D23_1c2567 Neut_0883
fig/66666666.60966.peg.2581	CDS	2381862	2380975	-3	-	888	Acetylglutamate kinase (EC 2.7.2.8)	Arginine Biosynthesis--> Arginine	D23_1c2568 Neut_2384
fig/66666666.60966.peg.2582	CDS	2382398	2381919	-2	-	480	type IV pili signal transduction protein PilH	Biosynthesis extended	D23_1c2569 Neut_2385
fig/66666666.60966.peg.2583	CDS	2382774	2382427	-3	-	348	twirling motility protein PilH	-none-	D23_1c2570 Neut_2386
fig/66666666.60966.peg.2584	CDS	2383169	2383639	2	+	471	21 kDa hemolysin precursor	CBSS-160492.1.peg.550	D23_1c2571 Neut_2387
fig/66666666.60966.peg.2585	CDS	2383676	2385043	2	+	1368	Fe—S protein, homolog of lactate	-none-	D23_1c2572 Neut_2388
fig/66666666.60966.peg.2586	CDS	2386032	2385136	-3	-	897	dehydrogenase SO1521 Heme O synthase, protoheme IX	Biogenesis of cytochrome c oxidases; CBSS-196164.1.peg.1690; CBSS-316057.3.peg.563	D23_1c2573 Neut_2389
fig/66666666.60966.peg.2587	CDS	2386422	2386123	-3	-	300	Probable transmembrane protein	-none-	D23_1c2574 Neut_2390

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld	C91 Alias
fig/66666666.60966.peg.2588	CDS	2387398	2386679	-1	-	720	Cytochrome oxidase biogenesis protein Surf1, facilitates heme A insertion	Biogenesis of cytochrome c oxidases; CBSS-316057.3.peg.563	D23_1c2575	Neut_2391
fig/66666666.60966.peg.2589	CDS	2388336	2387491	-3	-	846	Cytochrome c oxidase polypeptide III (EC 1.9.3.1)	CBSS-316057.3.peg.563; Terminal cytochrome C oxidases	D23_1c2576	Neut_2392
fig/66666666.60966.peg.2590	CDS	2389047	2388529	-3	-	519	Cytochrome oxidase biogenesis protein Cox11-CtaG, copper delivery to Cox1	Biogenesis of cytochrome c oxidases; CBSS-316057.3.peg.563	D23_1c2577	Neut_2393
fig/66666666.60966.peg.2591	CDS	2390759	2389182	-2	-	1578	Cytochrome c oxidase polypeptide I (EC 1.9.3.1)	Terminal cytochrome C oxidases	D23_1c2578	Neut_2394
fig/66666666.60966.peg.2592	CDS	2391643	2390819	-1	-	825	Cytochrome c oxidase polypeptide II (EC 1.9.3.1)	CBSS-316057.3.peg.563; Terminal cytochrome C oxidases	D23_1c2579	Neut_2395
fig/66666666.60966.peg.2594	CDS	2392145	2392567	2	+	423	Putative TEGT family carrier/transport protein	CBSS-326442.4.peg.1852	D23_1c2580	NA
fig/66666666.60966.peg.2595	CDS	2392600	2392884	1	+	285	Putative TEGT family carrier/transport protein	CBSS-326442.4.peg.1852	D23_1c2581	Neut_1715
fig/66666666.60966.peg.2596	CDS	2393084	2393302	2	+	219	Copper chaperone	Copper homeostasis	D23_1c2582	Neut_2397
fig/66666666.60966.peg.2597	CDS	2393426	2394394	2	+	969	Acetyl-coenzyme A carboxyl transferase	Fatty Acid Biosynthesis FASII	D23_1c2583	Neut_2398
fig/66666666.60966.peg.2598	CDS	2394337	2395742	3	+	1386	alpha chain (EC 6.4.1.2) tRNA(Ile)-lysidine synthetase	-none-	D23_1c2584	Neut_2399
fig/66666666.60966.peg.2599	CDS	2395746	2396807	3	+	1062	dTDP-glucose 4,6-dehydratase (EC 4.2.1.46)	CBSS-296591.1.peg.2330; Rhamnose containing glycans; dTDP-rhamnose synthesis	D23_1c2585	Neut_2400
fig/66666666.60966.peg.2600	CDS	2396804	2397700	2	+	897	dTDP-4-dehydrorhamnose reductase (EC 1.1.1.133)	Rhamnose containing glycans; dTDP-rhamnose synthesis	D23_1c2586	Neut_2401
fig/66666666.60966.peg.2601	CDS	2398663	2397710	-1	-	954	InterPro IPR002142 COGs COG0616	-none-	D23_1c2587	Neut_2402
fig/66666666.60966.peg.2602	CDS	2399353	2398691	-1	-	663	Outer membrane lipoprotein omp16 precursor	Ton and Tol transport systems	D23_1c2588	Neut_2403
fig/66666666.60966.peg.2603	CDS	2400398	2399562	-2	-	837	SH3, type 3 domain protein	-none-	D23_1c2590	Neut_2404
fig/66666666.60966.peg.2604	CDS	2400602	2401477	2	+	876	esterase/lipase/thioesterase family active site	-none-	D23_1c2591	Neut_2408

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld	C91 Alias
fig/66666666.60966.peg.2605	CDS	2402010	2401492	-3	-	519	Mobile element protein	-none-	D23_1c2592	Neut_2502
fig/66666666.60966.peg.2606	CDS	2402097	2403560	3	+	1464	hypothetical protein	-none-	D23_1c2593	Neut_2493
fig/66666666.60966.peg.2607	CDS	2403699	2404946	3	+	1248	Lipoprotein releasing system transmembrane protein LolE	Lipopolysaccharide assembly; Lipoprotein sorting system	D23_1c2594	Neut_2492
fig/66666666.60966.peg.2608	CDS	2404939	2405613	1	+	675	Lipoprotein releasing system ATP-binding protein LolD	Lipopolysaccharide assembly; Lipoprotein sorting system	D23_1c2595	Neut_2491
fig/66666666.60966.peg.2609	CDS	2406957	2405641	-3	-	1317	FIG065221; Holliday junction DNA helicase	CBSS-83333.1.peg.876	D23_1c2596	Neut_2490
fig/66666666.60966.peg.2610	CDS	2407570	2406950	-1	-	621	Outer membrane lipoprotein carrier protein LolA	CBSS-83333.1.peg.876; Lipopolysaccharide assembly; Lipoprotein sorting system	D23_1c2597	Neut_2489
fig/66666666.60966.peg.2611	CDS	2409936	2407630	-3	-	2307	Cell division protein FtsK	Bacterial cell Division; Bacterial cytoskeleton; Bacterial RNA-metabolizing Zn-dependent hydrolases; CBSS-83333.1.peg.876	D23_1c2598	Neut_2488
fig/66666666.60966.peg.2612	CDS	2411206	2410151	-1	-	1056	InterPro IPR002110 COGs COG0666	-none-	D23_1c2600	Neut_2487
fig/66666666.60966.peg.2613	CDS	2412939	2411203	-3	-	1737	Succinate dehydrogenase flavoprotein subunit (EC 1.3.99.1)	Succinate dehydrogenase; TCA Cycle	D23_1c2601	Neut_2486
fig/66666666.60966.peg.2614	CDS	2413237	2412968	-1	-	270	Succinate dehydrogenase hydrophobic membrane anchor protein	Succinate dehydrogenase	D23_1c2602	Neut_2485
fig/66666666.60966.peg.2615	CDS	2413803	2413315	-3	-	489	Succinate dehydrogenase cytochrome b-556 subunit	Succinate dehydrogenase	D23_1c2603	Neut_2484
fig/66666666.60966.peg.2616	CDS	2413889	2415583	2	+	1695	CTP synthase (EC 6.3.4.2)	CTP synthase (EC 6.3.4.2) cluster; pyrimidine conversions	D23_1c2604	Neut_2483
fig/66666666.60966.peg.2617	CDS	2415872	2417158	2	+	1287	Enolase (EC 4.2.1.11)	Glycolysis and Gluconeogenesis	D23_1c2605	Neut_2482
fig/66666666.60966.peg.2618	CDS	2417185	2417436	1	+	252	Cell division protein DivC (FtsB), stabilizes	Bacterial cell Division; Bacterial	D23_1c2606	Neut_2481

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 GbkId	C91 Alias
							FtsL against RasP cleavage	cytoskeleton; repair cluster		
fig 66666666.60966.psg.2619	CDS	2417726	2417860	2	+	135	hypothetical protein	-none-	D23_1c2607	NA
fig 66666666.60966.psg.2621	CDS	2419576	2418317	-1	-	1260	Transcription termination factor Rho	Transcription factors	D23_1c2609	Neut_2479
fig 66666666.60966.psg.2622	CDS	2420006	2419791	-2	-	216	Thioredoxin	bacterial	D23_1c2610	Neut_2478
fig 66666666.60966.psg.2624	CDS	2420605	2421762	1	+	1158	Membrane-bound lytic murein transglycosylase B precursor (EC 3.2.1.—)	-none-	D23_1c2611	Neut_2477
fig 66666666.60966.psg.2625	CDS	2422278	2421835	-3	-	444	Universal stress protein	 Peptidoglycan Biosynthesis	D23_1c2612	Neut_2476
fig 66666666.60966.psg.2626	CDS	2424743	2422389	-2	-	2355	family COG0589 Partial urea carboxylase 2 (EC 6.3.4.6)	Urea carboxylase and Allophanate hydrolase cluster; Urea decomposition	D23_1c2613	Neut_2475
fig 66666666.60966.psg.2627	CDS	2425462	2424797	-1	-	666	Urea carboxylase-related aminomethyltransferase (EC 2.1.2.10)	Urea decomposition	D23_1c2614	Neut_2474
fig 66666666.60966.psg.2628	CDS	2426187	2425459	-3	-	729	Urea carboxylase-related aminomethyltransferase (EC 2.1.2.10)	Urea decomposition	D23_1c2615	Neut_2473
fig 66666666.60966.psg.2629	CDS	2427740	2426202	-2	-	1539	Urea carboxylase-related amino acid permease	Urea decomposition	D23_1c2616	Neut_2472
fig 66666666.60966.psg.2630	CDS	2427773	2427892	2	+	120	hypothetical protein	-none-	D23_1c2617	NA
fig 66666666.60966.psg.2631	CDS	2428128	2428319	3	+	192	hypothetical protein	-none-	D23_1c2618	NA
fig 66666666.60966.psg.2632	CDS	2428646	2428518	-2	-	129	hypothetical protein	-none-	D23_1c2620	NA
fig 66666666.60966.psg.2633	CDS	2428671	2432291	3	+	3621	Urea carboxylase (EC 6.3.4.6)	Urea carboxylase and Allophanate hydrolase cluster; Urea decomposition	D23_1c2621	Neut_2470
fig 66666666.60966.psg.2634	CDS	2433643	2432279	-1	-	1365	Membrane-bound lytic murein transglycosylase D precursor (EC 3.2.1.—)	CBSS-228410.1.psg.134; CBSS-342610.3.psg.1536; Murein Hydrolases	D23_1c2622	Neut_2469
fig 66666666.60966.psg.2635	CDS	2434924	2433791	-1	-	1134	Ribonucleotide reductase of class Ia (aerobic), beta subunit (EC 1.17.4.1)	Ribonucleotide reduction	D23_1c2623	Neut_2468
fig 66666666.60966.psg.2636	CDS	2437918	2434976	-1	-	2943	Ribonucleotide reductase of class Ia (aerobic), alpha subunit (EC 1.17.4.1)	Ribonucleotide reduction	D23_1c2624	Neut_2467
fig 66666666.60966.psg.2637	CDS	2438096	2441536	2	+	3441	Long-chain-fatty-acid-CoA ligase (EC 6.2.1.3)	Biotin biosynthesis; Biotin synthesis	D23_1c2625	Neut_2466

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld	C91 Alias
fig/66666666.60966.peg.2638	CDS	2441648	2441788	2	+	141	hypothetical protein	cluster: Fatty acid metabolism cluster -none- -none- Polyamine Metabolism	D23_1c2626	NA
fig/66666666.60966.peg.2639	CDS	2441958	2441845	-3	-	114	hypothetical protein		D23_1c2627	NA
fig/66666666.60966.peg.2640	CDS	2441984	2442337	2	+	354	S-adenosylmethionine decarboxylase proenzyme (EC 4.1.1.50), prokaryotic class 1B		D23_1c2628	Neut_2465
fig/66666666.60966.peg.2641	CDS	2442337	2443296	1	+	960	Spermidine synthase (EC 2.5.1.16)		D23_1c2629	Neut_2464
fig/66666666.60966.peg.2642	CDS	2444178	2443342	-3	-	837	Cytochrome c family protein	-none-	D23_1c2630	Neut_2463
fig/66666666.60966.peg.2643	CDS	2446065	2444275	-3	-	1791	ABC transporter, fused permease and ATPase domains	-none-	D23_1c2631	Neut_2462
fig/66666666.60966.peg.2644	CDS	2446735	2446268	-1	-	468	Single-stranded DNA-binding protein	DNA repair, bacterial	D23_1c2632	Neut_2461
fig/66666666.60966.peg.2645	CDS	2448142	2446736	-1	-	1407	Putative transport protein	-none-	D23_1c2633	Neut_2460
fig/66666666.60966.peg.2646	CDS	2448217	2451054	1	+	2838	Exonuclease ABC subunit A	DNA repair, UvrABC system	D23_1c2634	Neut_2459
fig/66666666.60966.peg.2647	CDS	2451051	2452322	3	+	1272	D-glycerate 2-kinase (EC 2.7.1.—)	Glycerate metabolism	D23_1c2635	Neut_2458
fig/66666666.60966.peg.2648	CDS	2455286	2452443	-2	-	2844	Aconitate hydratase (EC 4.2.1.3)	TCA Cycle	D23_1c2637	Neut_2457
fig/66666666.60966.peg.2649	CDS	2455429	2456334	1	+	906	Aldose 1-epimerase	-none-	D23_1c2638	Neut_2456
fig/66666666.60966.peg.2650	CDS	2456378	2456941	2	+	564	NADPH-dependent FMN reductase	-none-	D23_1c2639	Neut_2455
fig/66666666.60966.peg.2651	CDS	2457281	2457051	-2	-	231	HrgA protein	-none-	D23_1c2640	Neut_2454
fig/66666666.60966.peg.2652	CDS	2457998	2457480	-2	-	519	HrgA protein	-none-	D23_1c2641	NA
fig/66666666.60966.peg.2653	CDS	2459350	2458013	-1	-	1338	hypothetical protein	-none-	D23_1c2642	NA
fig/66666666.60966.peg.2654	CDS	2460249	2459347	-3	-	903	hypothetical protein	-none-	D23_1c2643	NA
fig/66666666.60966.peg.2655	CDS	2463305	2460264	-2	-	3042	Type I restriction-modification system, restriction subunit R (EC 3.1.21.3)	Restriction-Modification System: Type I Restriction-Modification	D23_1c2644	NA
fig/66666666.60966.peg.2656	CDS	2463943	2463341	-1	-	603	hypothetical protein	-none-	D23_1c2645	NA
fig/66666666.60966.peg.2657	CDS	2465348	2463960	-2	-	1389	Type I restriction-modification system, specificity subunit S (EC 3.1.21.3)	Restriction-Modification System: Type I Restriction-Modification	D23_1c2646	NA
fig/66666666.60966.peg.2658	CDS	2467567	2465348	-1	-	2220	Type I restriction-modification system, DNA-methyltransferase subunit M (EC 2.1.1.72)	Restriction-Modification System: Type I Restriction-Modification	D23_1c2647	Neut_0541

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.									
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 GbkId C91 Alias
fig 66666666.60966.peg.2659	CDS	2468029	2467751	-1	-	279	Type I restriction-modification system, DNA-methyltransferase subunit M (EC 2.1.1.72)	Restriction-Modification System; Type I Restriction-Modification	D23_1c2648 Neut_2448
fig 66666666.60966.peg.2660	CDS	2468945	2468235	-2	-	711	RNA polymerase sigma factor for flagellar operon	Flagellar motility; Flagellum; Transcription initiation, bacterial sigma factors	D23_1c2650 Neut_2447
fig 66666666.60966.peg.2661	CDS	2469847	2468954	-1	-	894	Flagellar synthesis regulator FlhN	Flagellar motility; Flagellum	D23_1c2651 Neut_2446
fig 66666666.60966.peg.2662	CDS	2471090	2469840	-2	-	1251	Flagellar biosynthesis protein FlhF	Flagellar motility; Flagellum	D23_1c2652 Neut_2445
fig 66666666.60966.peg.2663	CDS	2473171	2471087	-1	-	2085	Flagellar biosynthesis protein FlhA	Flagellar motility; Flagellum	D23_1c2653 Neut_2444
fig 66666666.60966.peg.2664	CDS	2474349	2473219	-3	-	1131	Flagellar biosynthesis protein FlhB	Flagellar motility; Flagellum	D23_1c2654 Neut_2443
fig 66666666.60966.peg.2665	CDS	2476208	2474601	-2	-	1608	Peptide chain release factor 3	Translation termination factors bacterial	D23_1c2655 Neut_2442
fig 66666666.60966.peg.2666	CDS	2478262	2476208	-1	-	2055	Dipeptide transport ATP-binding protein DppF (TC 3.A.1.5.2)	ABC transporter dipeptide (TC 3.A.1.5.2)	D23_1c2656 Neut_2441
fig 66666666.60966.peg.2667	CDS	2479752	2478259	-3	-	1494	Oligopeptide transport system permease protein	-none-	D23_1c2657 Neut_2440
fig 66666666.60966.peg.2669	CDS	2480032	2480334	1	+	303	Negative regulator of flagellin synthesis	Flagellum	D23_1c2658 Neut_2439
fig 66666666.60966.peg.2670	CDS	2480417	2480800	2	+	384	Flagellar biosynthesis protein FlgN	Flagellum	D23_1c2659 Neut_2438
fig 66666666.60966.peg.2671	CDS	2481184	2483109	1	+	1926	tRNA uridine 5-carboxymethylaminomethyl modification enzyme GldA	Cell Division Subsystem including YidCD; RNA modification and chromosome partitioning cluster; mmn5U34 biosynthesis bacteria; RNA modification Bacteria	D23_1c2660 Neut_2437
fig 66666666.60966.peg.2672	CDS	2483084	2483728	2	+	645	rRNA small subunit 7-methylguanosine (m7G) methyltransferase GldB	Cell Division Subsystem including YidCD; RNA methylation; RNA modification and chromosome partitioning cluster	D23_1c2661 Neut_2436
fig 66666666.60966.peg.2673	CDS	2483792	2484556	2	+	765	Chromosome (plasmid) partitioning protein ParA/Sporulation	Bacterial Cell Division; Bacterial Cytoskeleton;	D23_1c2662 Neut_2435

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.									
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld
fig/66666666.60966.psg.2674	CDS	2484637	2485440	1	+	804	initiation inhibitor protein Soj	 Bacterial Cytoskeleton; Cell Division Subsystem including YrdCD; RNA modification and chromosome partitioning cluster	D23_1c2663
							Chromosome (plasmid) partitioning protein ParB/Stage 0 sporulation protein J	Bacterial Cytoskeleton; Bacterial Cytoskeleton; Cell Division Subsystem including YrdCD; RNA modification and chromosome partitioning cluster	Neut_2434
fig/66666666.60966.psg.2675	CDS	2485777	2488083	1	+	2307	Glucose-6-phosphate isomerase (EC 5.3.1.9)	Glycolysis and Gluconeogenesis	D23_1c2664
fig/66666666.60966.psg.2676	CDS	2488076	2489131	2	+	1056	6-phosphogluconate dehydrogenase, decarboxylating (EC 1.1.1.44)	D-glucuronate and ketogluconates metabolism; Pentose phosphate pathway	D23_1c2665
fig/66666666.60966.psg.2677	CDS	2489149	2490591	1	+	1443	Glucose-6-phosphate 1-dehydrogenase (EC 1.1.1.49)	Pentose phosphate pathway	D23_1c2666
fig/66666666.60966.psg.2678	CDS	2490598	2490741	1	+	144	hypothetical protein	-none-	D23_1c2667
fig/66666666.60966.psg.2679	CDS	2491345	2490770	-1	-	576	putative membrane protein	-none-	D23_1c2668
fig/66666666.60966.psg.2680	CDS	2491701	2492270	3	+	570	DedA family inner membrane protein	DedA family of inner membrane proteins	D23_1c2669
fig/66666666.60966.psg.2681	CDS	2492664	2492897	3	+	234	Mobile element protein	-none-	D23_1c2671
fig/66666666.60966.psg.2682	CDS	2493050	2494789	2	+	1740	ABC-type anion transport system, duplicated permease component	-none-	D23_1c2672
fig/66666666.60966.psg.2683	CDS	2494811	2496100	2	+	1290	ABC-type nitrate/sulfonate/bicarbonate transport system, ATPase component	Alkanesulfonate assimilation	D23_1c2673
fig/66666666.60966.psg.2684	CDS	2499268	2496263	-1	-	3006	Exonuclease SbcC	DNA repair, bacterial; Rad50-Mre11 DNA repair cluster	D23_1c2676
fig/66666666.60966.psg.2685	CDS	2499230	2499520	2	+	291	hypothetical protein	-none-	D23_1c2677
fig/66666666.60966.psg.2686	CDS	2500770	2499526	-3	-	1245	Exonuclease SbcD	DNA repair, bacterial; Rad50-Mre11 DNA repair cluster	D23_1c2678

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld	C91 Alias
fig/66666666.60966.peg.2687	CDS	2501402	2500767	-2	-	636	FIG01057587: hypothetical protein	-none-	D23_1c2679	NA
fig/66666666.60966.peg.2688	CDS	2501615	2502037	2	+	423	Mobile element protein	-none-	D23_1c2680	Neut_2450
fig/66666666.60966.peg.2689	CDS	2502217	2502050	-1	-	168	hypothetical protein	-none-	D23_1c2681	NA
fig/66666666.60966.peg.2691	CDS	2503677	2502628	-3	-	1050	hypothetical protein	-none-	D23_1c2682	Neut_2415
fig/66666666.60966.peg.2692	CDS	2503893	2504105	3	+	213	hypothetical protein	-none-	D23_1c2683	NA
fig/66666666.60966.peg.2693	CDS	2505239	2504595	-2	-	645	hypothetical protein	-none-	D23_1c2685	NA
fig/66666666.60966.peg.2694	CDS	2506092	2505229	-3	-	864	Mobile element protein	-none-	D23_1c2686	Neut_2192
fig/66666666.60966.peg.2695	CDS	2506385	2506089	-2	-	297	Mobile element protein	-none-	D23_1c2687	Neut_2193
fig/66666666.60966.peg.2696	CDS	2506924	2506445	-1	-	480	DNA primase/helicase, phage-associated	Phage replication	D23_1c2688	NA
fig/66666666.60966.peg.2697	CDS	2507166	2506921	-3	-	246	hypothetical protein	-none-	D23_1c2689	NA
fig/66666666.60966.peg.2698	CDS	2507393	2507166	-2	-	228	hypothetical protein	-none-	D23_1c2690	NA
fig/66666666.60966.peg.2699	CDS	2507735	2508577	2	+	843	hypothetical protein	-none-	D23_1c2691	NA
fig/66666666.60966.peg.2700	CDS	2508650	2509111	2	+	462	Putative bacteriophage-related protein	-none-	D23_1c2692	NA
fig/66666666.60966.peg.2701	CDS	2509108	2510448	1	+	1341	probable DNA invertase	-none-	D23_1c2693	Neut_2563
fig/66666666.60966.peg.2702	CDS	2510478	2510897	3	+	420	elements of external origin; phage-related functions and prophages	-none-	D23_1c2694	NA
fig/66666666.60966.peg.2703	CDS	2512240	2511590	-1	-	651	Similar to phosphoglycolate phosphatase, clustered with ubiquinone biosynthesis SAM-dependent O-methyltransferase	2-phosphoglycolate salvage	D23_1c2696	Neut_2506
fig/66666666.60966.peg.2704	CDS	2512973	2512269	-2	-	705	3-demethylubiquinol 3-O-methyltransferase (EC 2.1.1.64)	Ubiquinone Biosynthesis; Ubiquinone	D23_1c2697	Neut_2507
fig/66666666.60966.peg.2705	CDS	2513079	2512963	-3	-	117	hypothetical protein	 Biosynthesis-gio	D23_1c2698	NA
fig/66666666.60966.peg.2706	CDS	2513892	2513197	-3	-	696	Outer membrane protein A precursor	-none-	D23_1c2699	Neut_2508
fig/66666666.60966.peg.2708	CDS	2514251	2515009	2	+	759	ABC-type multidrug transport system, ATPase component	Osmoregulation	D23_1c2700	Neut_2509
fig/66666666.60966.peg.2709	CDS	2515006	2515752	1	+	747	gliding motility protein GldF	-none-	D23_1c2701	Neut_2510
fig/66666666.60966.peg.2710	CDS	2515776	2517128	3	+	1353	Mucin 2 precursor	-none-	D23_1c2702	Neut_2511
fig/66666666.60966.peg.2711	CDS	2517144	2517959	3	+	816	Formamidopyrimidine-DNA glycosylase (EC 3.2.2.23)	DNA Repair Base Excision	D23_1c2703	Neut_2512
fig/66666666.60966.peg.2712	CDS	2518479	2517994	-3	-	486	Thioredoxin	-none-	D23_1c2704	Neut_2513
fig/66666666.60966.peg.2713	CDS	2520458	2518644	-2	-	1815	GTP-binding protein TypA/BipA	Universal GTPases	D23_1c2705	Neut_2514

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld	C91 Alias
fig/66666666.60966.peg.2714	CDS	2520591	2521196	3	+	606	Riboflavin synthase eubacterial/eukaryotic (EC 2.5.1.9)	Riboflavin, FMN and FAD metabolism; Riboflavin, FMN and FAD metabolism in plants; Riboflavin synthesis cluster;	D23_1c2706	Neut_2515
fig/66666666.60966.peg.2715	CDS	2521193	2522302	2	+	1110	3,4-dihydroxy-2-butanone 4-phosphate synthase (EC 4.1.99.12)/GTP cyclohydrolase II (EC 3.5.4.25)	Riboflavin, FMN and FAD metabolism; Riboflavin, FMN and FAD metabolism in plants; Riboflavin, FMN and FAD metabolism in plants; Riboflavin synthesis cluster; Riboflavin synthesis cluster;	D23_1c2707	Neut_2516
fig/66666666.60966.peg.2716	CDS	2522529	2523026	3	+	498	6,7-dimethyl-8-ribityllumazine synthase (EC 2.5.1.78)	Possible RNA degradation cluster; Riboflavin, FMN and FAD metabolism; Riboflavin, FMN and FAD metabolism in plants; Riboflavin synthesis cluster;	D23_1c2708	Neut_2517
fig/66666666.60966.peg.2717	CDS	2523023	2523526	2	+	504	Transcription termination protein NusB	Riboflavin synthesis cluster; Transcription factors bacterial	D23_1c2709	Neut_2518
fig/66666666.60966.peg.2719	CDS	2523801	2524787	3	+	987	Thiamine-monophosphate kinase (EC 2.7.4.16)	5-FLC-like protein; Riboflavin synthesis cluster; Thiamin biosynthesis	D23_1c2710	Neut_2519
fig/66666666.60966.peg.2720	CDS	2524864	2525307	1	+	444	Phosphatidylglycerophosphatase A (EC 3.1.3.27)	Glycerolipid and Glycerophospholipid Metabolism in Bacteria; Riboflavin synthesis cluster	D23_1c2711	Neut_2520
fig/66666666.60966.peg.2721	CDS	2525304	2525846	3	+	543	C-terminal domain of ClnA type S; Protein Implicated in DNA repair function with RecA and MutS	DNA repair system including RecA, MutS and a hypothetical protein; NAD and NADP cofactor biosynthesis global; NAD and NADP	D23_1c2712	Neut_2521

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld	C91 Alias
fig/66666666.60966.peg.2722	CDS	2527502	2526306	-2	-	1197	FIG00859610: hypothetical protein ATP-dependent DNA helicase RecQ	cofactor biosynthesis global; Possible RNA degradation cluster; Riboflavin, FMN and FAD metabolism in plants; Riboflavin synthesis cluster	D23_1c2713	Neut_2522
									D23_1c2714	Neut_2523
									D23_1c2715	Neut_2524
fig/66666666.60966.peg.2725	CDS	2530686	2529877	-3	-	810	InterPro IPR002781	-none-	D23_1c2716	Neut_2525
fig/66666666.60966.peg.2726	CDS	2532376	2530694	-1	-	1683	COGs COG0730 FIG003847: Oxidoreductase (flavoprotein)	CBSS-269482.4,peg.5018	D23_1c2717	Neut_2526
fig/66666666.60966.peg.2727	CDS	2532913	2532488	-1	-	426	Transcriptional regulator, ArsR family	-none-	D23_1c2718	Neut_2527
fig/66666666.60966.peg.2728	CDS	2533377	2532910	-3	-	468	GENE II AND X PROTEINS	-none-	D23_1c2719	Neut_2528
fig/66666666.60966.peg.2729	CDS	2533835	2533407	-2	-	429	Probable transmembrane protein	-none-	D23_1c2720	Neut_2529
fig/66666666.60966.peg.2730	CDS	2533940	2534806	2	+	867	FIG146518: Zn-dependent glyoxylases, including glyoxylases	CBSS-269482.4,peg.5018	D23_1c2721	Neut_2530
fig/66666666.60966.peg.2731	CDS	2535475	2534891	-1	-	585	FIG001587: exported protein	-none-	D23_1c2722	Neut_2531
fig/66666666.60966.peg.2732	CDS	2537039	2535534	-2	-	1506	FIG00859025: hypothetical protein	-none-	D23_1c2723	Neut_2532
fig/66666666.60966.peg.2733	CDS	2537388	2537077	-3	-	312	hypothetical protein	-none-	D23_1c2724	NA
fig/66666666.60966.mia.11	RNA	226254	226324	3	+	71	tRNA-Gly-CCC	tRNAs	NA	NA
fig/66666666.60966.mia.2	RNA	6795	6868	3	+	74	tRNA-Cys-GCA	tRNAs	NA	NA
fig/66666666.60966.mia.39	RNA	1810921	1810848	-1	-	74	tRNA-Gly-TCC	-none-	NA	NA
fig/66666666.60966.mia.23	RNA	969765	969839	3	+	75	tRNA-Gln-TTG	-none-	NA	NA
fig/66666666.60966.mia.36	RNA	1758968	1759042	2	+	75	tRNA-Val-CAC	tRNAs	NA	NA
fig/66666666.60966.mia.38	RNA	1810812	1810738	-3	-	75	tRNA-Thr-GGT	-none-	NA	NA
fig/66666666.60966.mia.1	RNA	6614	6689	2	+	76	tRNA-Gly-GCC	tRNAs	NA	NA
fig/66666666.60966.mia.7	RNA	120428	120503	2	+	76	tRNA-Ala-TGC	-none-	NA	NA
fig/66666666.60966.mia.12	RNA	284547	284472	-3	-	76	tRNA-Glu-TTC	-none-	NA	NA
fig/66666666.60966.mia.13	RNA	284648	284573	-2	-	76	tRNA-Ala-GGC	tRNAs	NA	NA
fig/66666666.60966.mia.14	RNA	493566	493641	3	+	76	tRNA-Thr-CGT	-none-	NA	NA
fig/66666666.60966.mia.16	RNA	561828	561903	3	+	76	tRNA-Val-TAC	-none-	NA	NA
fig/66666666.60966.mia.20	RNA	859357	859282	-1	-	76	tRNA-Lys-TTT	-none-	NA	NA

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld	C91 Alias
fig/66666666.60966.ma.21	RNA	948801	948876	3	+	76	tRNA-Thr-TGT	-none-	NA	NA
fig/66666666.60966.ma.24	RNA	1157662	1157587	-1	-	76	tRNA-Arg-CCT	-none-	NA	NA
fig/66666666.60966.ma.25	RNA	1199325	1199250	-3	-	76	tRNA-His-GTG	-none-	NA	NA
fig/66666666.60966.ma.29	RNA	1470510	1470435	-3	-	76	tRNA-Asn-GTT	-none-	NA	NA
fig/66666666.60966.ma.33	RNA	1618239	1618314	3	+	76	tRNA-Met-CAT	-none-	NA	NA
fig/66666666.60966.ma.34	RNA	1673498	1673423	-2	-	76	tRNA-Arg-CCG	tRNAs	NA	NA
fig/66666666.60966.ma.37	RNA	1809438	1809363	-3	-	76	tRNA-Tip-CCA	tRNAs	NA	NA
fig/66666666.60966.ma.42	RNA	2107281	2107356	3	+	76	tRNA-Phe-GAA	tRNAs	NA	NA
fig/66666666.60966.ma.6	RNA	120349	120425	1	+	77	tRNA-Ile-GAT	-none-	NA	NA
fig/66666666.60966.ma.10	RNA	196762	196838	1	+	77	tRNA-Met-CAT	-none-	NA	NA
fig/66666666.60966.ma.15	RNA	524553	524629	3	+	77	tRNA-Val-GAC	tRNAs	NA	NA
fig/66666666.60966.ma.17	RNA	561963	562039	3	+	77	tRNA-Asp-GTC	-none-	NA	NA
fig/66666666.60966.ma.18	RNA	728392	728316	-1	-	77	tRNA-Pro-CGG	tRNAs	NA	NA
fig/66666666.60966.ma.26	RNA	1199452	1199376	-1	-	77	tRNA-Arg-TCT	-none-	NA	NA
fig/66666666.60966.ma.27	RNA	1199573	1199497	-2	-	77	tRNA-Pro-TGG	-none-	NA	NA
fig/66666666.60966.ma.28	RNA	1427736	1427812	3	+	77	tRNA-Met-CAT	-none-	NA	NA
fig/66666666.60966.ma.41	RNA	2107137	2107213	3	+	77	tRNA-Arg-ACG	tRNAs	NA	NA
fig/66666666.60966.ma.44	RNA	2339644	2339568	-1	-	77	tRNA-Pro-GGG	tRNAs	NA	NA
fig/66666666.60966.ma.19	RNA	835520	835604	2	+	85	tRNA-Leu-GAG	tRNAs	NA	NA
fig/66666666.60966.ma.32	RNA	1597631	1597715	2	+	85	tRNA-Leu-CAG	tRNAs	NA	NA
fig/66666666.60966.ma.40	RNA	1811116	1811032	-1	-	85	tRNA-Tyr-GTA	-none-	NA	NA
fig/66666666.60966.ma.4	RNA	97072	96987	-1	-	86	tRNA-Leu-TAG	-none-	NA	NA
fig/66666666.60966.ma.31	RNA	1574940	1574854	-3	-	87	tRNA-Leu-CAA	tRNAs	NA	NA
fig/66666666.60966.ma.30	RNA	1550222	1550135	-2	-	88	tRNA-Ser-TGA	-none-	NA	NA
fig/66666666.60966.ma.3	RNA	6894	6982	3	+	89	tRNA-Leu-TAA	-none-	NA	NA
fig/66666666.60966.ma.22	RNA	956464	956374	-1	-	91	tRNA-Ser-GGA	tRNAs	NA	NA
fig/66666666.60966.ma.35	RNA	1757250	1757342	3	+	93	tRNA-Ser-CGA	tRNAs	NA	NA
fig/66666666.60966.ma.43	RNA	2120297	2120205	-2	-	93	tRNA-Ser-GCT	-none-	NA	NA
fig/66666666.60966.psg.111	CDS	108515	108628	2	+	114	hypothetical protein	-none-	NA	NA
fig/66666666.60966.psg.353	CDS	326760	326647	-3	-	114	hypothetical protein	-none-	NA	NA
fig/66666666.60966.psg.404	CDS	375917	375804	-2	-	114	hypothetical protein	-none-	NA	NA
fig/66666666.60966.psg.432	CDS	397539	397426	-3	-	114	hypothetical protein	-none-	NA	NA
fig/66666666.60966.psg.509	CDS	474482	474369	-2	-	114	hypothetical protein	-none-	NA	NA
fig/66666666.60966.psg.589	CDS	543008	543121	2	+	114	hypothetical protein	-none-	NA	NA
fig/66666666.60966.psg.838	CDS	770105	770218	2	+	114	hypothetical protein	-none-	NA	NA
fig/66666666.60966.psg.923	CDS	855159	855046	-3	-	114	hypothetical protein	-none-	NA	NA
fig/66666666.60966.psg.943	CDS	873193	873306	1	+	114	hypothetical protein	-none-	NA	NA
fig/66666666.60966.psg.1081	CDS	1010260	1010373	1	+	114	hypothetical protein	-none-	NA	NA
fig/66666666.60966.psg.1333	CDS	1235666	1235779	2	+	114	hypothetical protein	-none-	NA	NA
fig/66666666.60966.psg.1414	CDS	1326350	1326463	2	+	114	hypothetical protein	-none-	NA	NA
fig/66666666.60966.psg.1695	CDS	1575540	1575653	3	+	114	hypothetical protein	-none-	NA	NA
fig/66666666.60966.psg.1875	CDS	1749599	1749486	-2	-	114	Mobile element protein	-none-	NA	NA
fig/66666666.60966.psg.2082	CDS	1936491	1936378	-3	-	114	hypothetical protein	-none-	NA	NA
fig/66666666.60966.psg.2205	CDS	2047715	2047828	2	+	114	hypothetical protein	-none-	NA	NA
fig/66666666.60966.psg.2370	CDS	2201390	2201277	-2	-	114	hypothetical protein	-none-	NA	NA
fig/66666666.60966.psg.2413	CDS	2236764	2236651	-3	-	114	hypothetical protein	-none-	NA	NA
fig/66666666.60966.psg.2577	CDS	2379043	2378930	-1	-	114	hypothetical protein	-none-	NA	NA
fig/66666666.60966.psg.421	CDS	392634	392518	-3	-	117	hypothetical protein	-none-	NA	NA

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld	C91 Alias
fig/6666666.60966.psg.431	CDS	397234	397350	1	+	117	hypothetical protein	-none-	NA	NA
fig/6666666.60966.psg.1104	CDS	1033773	1033889	3	+	117	hypothetical protein	-none-	NA	NA
fig/6666666.60966.psg.1856	CDS	1730271	1730155	-3	-	117	hypothetical protein	-none-	NA	NA
fig/6666666.60966.psg.2044	CDS	1907891	1907775	-2	-	117	hypothetical protein	-none-	NA	NA
fig/6666666.60966.psg.2046	CDS	1908400	1908516	1	+	117	hypothetical protein	-none-	NA	NA
fig/6666666.60966.psg.2250	CDS	2086632	2086516	-3	-	117	hypothetical protein	-none-	NA	NA
fig/6666666.60966.psg.2378	CDS	2209362	2209246	-3	-	117	hypothetical protein	-none-	NA	NA
fig/6666666.60966.psg.28	CDS	31960	31841	-1	-	120	hypothetical protein	-none-	NA	NA
fig/6666666.60966.psg.536	CDS	496598	496717	2	+	120	hypothetical protein	-none-	NA	NA
fig/6666666.60966.psg.866	CDS	794273	794392	2	+	120	hypothetical protein	-none-	NA	NA
fig/6666666.60966.psg.1031	CDS	956342	956223	-2	-	120	hypothetical protein	-none-	NA	NA
fig/6666666.60966.psg.1238	CDS	1152349	1152230	-1	-	120	hypothetical protein	-none-	NA	NA
fig/6666666.60966.psg.1279	CDS	1184453	1184572	2	+	120	hypothetical protein	-none-	NA	NA
fig/6666666.60966.psg.1566	CDS	1462529	1462410	-2	-	120	hypothetical protein	-none-	NA	NA
fig/6666666.60966.psg.1604	CDS	1492079	1491960	-2	-	120	hypothetical protein	-none-	NA	NA
fig/6666666.60966.psg.1610	CDS	1495100	1495219	2	+	120	hypothetical protein	-none-	NA	NA
fig/6666666.60966.psg.1938	CDS	1820165	1820046	-2	-	120	hypothetical protein	-none-	NA	NA
fig/6666666.60966.psg.1940	CDS	1820392	1820511	1	+	120	hypothetical protein	-none-	NA	NA
fig/6666666.60966.psg.2194	CDS	2040715	2040596	-1	-	120	hypothetical protein	-none-	NA	NA
fig/6666666.60966.psg.2229	CDS	2069192	2069073	-2	-	120	hypothetical protein	-none-	NA	NA
fig/6666666.60966.psg.2542	CDS	2355482	2355601	2	+	120	hypothetical protein	-none-	NA	NA
fig/6666666.60966.psg.477	CDS	444630	444508	-3	-	123	hypothetical protein	-none-	NA	NA
fig/6666666.60966.psg.634	CDS	579242	579364	2	+	123	hypothetical protein	-none-	NA	NA
fig/6666666.60966.psg.927	CDS	859107	859229	3	+	123	hypothetical protein	-none-	NA	NA
fig/6666666.60966.psg.1015	CDS	941818	941940	1	+	123	hypothetical protein	-none-	NA	NA
fig/6666666.60966.psg.1388	CDS	1298114	1298236	2	+	123	hypothetical protein	-none-	NA	NA
fig/6666666.60966.psg.1755	CDS	1633195	1633073	-1	-	123	hypothetical protein	-none-	NA	NA
fig/6666666.60966.psg.2224	CDS	2067753	2067631	-3	-	123	hypothetical protein	-none-	NA	NA
fig/6666666.60966.psg.2271	CDS	2103666	2103788	3	+	123	hypothetical protein	-none-	NA	NA
fig/6666666.60966.psg.2336	CDS	2170340	2170462	2	+	123	hypothetical protein	-none-	NA	NA
fig/6666666.60966.psg.2415	CDS	2237130	2237008	-3	-	123	Hydroxyacylglutathione hydrolase (EC 3.1.2.6)	CBSS-228410.1.psg.134; CBSS-342610.3.psg.1536; Glutathione: Non-redox reactions; Methylglyoxal Metabolism	NA	NA
fig/6666666.60966.psg.82	CDS	79801	79676	-1	-	126	hypothetical protein	-none-	NA	NA
fig/6666666.60966.psg.585	CDS	541998	542123	3	+	126	hypothetical protein	-none-	NA	NA
fig/6666666.60966.psg.1268	CDS	1176741	1176866	3	+	126	hypothetical protein	-none-	NA	NA
fig/6666666.60966.psg.1473	CDS	1385674	1385799	1	+	126	hypothetical protein	-none-	NA	NA
fig/6666666.60966.psg.1510	CDS	1423155	1423030	-3	-	126	hypothetical protein	-none-	NA	NA
fig/6666666.60966.psg.1690	CDS	1573599	1573474	-3	-	126	hypothetical protein	-none-	NA	NA
fig/6666666.60966.psg.1758	CDS	1635456	1635331	-3	-	126	hypothetical protein	-none-	NA	NA
fig/6666666.60966.psg.1935	CDS	1816227	1816352	3	+	126	hypothetical protein	-none-	NA	NA
fig/6666666.60966.psg.1966	CDS	1832788	1832913	1	+	126	hypothetical protein	-none-	NA	NA
fig/6666666.60966.psg.2285	CDS	2119314	2119439	3	+	126	hypothetical protein	-none-	NA	NA
fig/6666666.60966.psg.21	CDS	28650	28522	-3	-	129	hypothetical protein	-none-	NA	NA

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.									
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	C91 Alias
fig/6666666.60966.psg.27	CDS	31632	31760	3	+	129	hypothetical protein	-none-	NA
fig/6666666.60966.psg.247	CDS	231537	231665	3	+	129	hypothetical protein	-none-	NA
fig/6666666.60966.psg.300	CDS	272364	272492	3	+	129	hypothetical protein	-none-	NA
fig/6666666.60966.psg.864	CDS	7922354	792226	-3	-	129	hypothetical protein	-none-	NA
fig/6666666.60966.psg.1166	CDS	1092637	1092509	-1	-	129	hypothetical protein	-none-	NA
fig/6666666.60966.psg.1314	CDS	12166530	1216658	3	+	129	hypothetical protein	-none-	NA
fig/6666666.60966.psg.1730	CDS	1612183	1612311	1	+	129	hypothetical protein	-none-	NA
fig/6666666.60966.psg.2328	CDS	2162434	2162306	-1	-	129	hypothetical protein	-none-	NA
fig/6666666.60966.psg.122	CDS	124194	124063	-3	-	132	hypothetical protein	-none-	NA
fig/6666666.60966.psg.154	CDS	155514	155645	3	+	132	hypothetical protein	-none-	NA
fig/6666666.60966.psg.874	CDS	801546	801677	3	+	132	hypothetical protein	-none-	NA
fig/6666666.60966.psg.894	CDS	823808	823939	2	+	132	hypothetical protein	-none-	NA
fig/6666666.60966.psg.960	CDS	886681	886550	-1	-	132	hypothetical protein	-none-	NA
fig/6666666.60966.psg.1188	CDS	1108861	1108730	-1	-	132	hypothetical protein	-none-	NA
fig/6666666.60966.psg.1578	CDS	1468706	1468837	2	+	132	hypothetical protein	-none-	NA
fig/6666666.60966.psg.1950	CDS	1824304	1824435	1	+	132	hypothetical protein	-none-	NA
fig/6666666.60966.psg.1983	CDS	1850338	1850469	1	+	132	hypothetical protein	-none-	NA
fig/6666666.60966.psg.2066	CDS	1923319	1923450	1	+	132	hypothetical protein	-none-	NA
fig/6666666.60966.psg.2445	CDS	2263603	2263472	-1	-	132	hypothetical protein	-none-	NA
fig/6666666.60966.psg.2570	CDS	2376200	2376331	2	+	132	hypothetical protein	-none-	NA
fig/6666666.60966.psg.296	CDS	267656	267790	2	+	135	hypothetical protein	-none-	NA
fig/6666666.60966.psg.1086	CDS	1016149	1016015	-1	-	135	hypothetical protein	-none-	NA
fig/6666666.60966.psg.1167	CDS	1092703	1092837	1	+	135	hypothetical protein	-none-	NA
fig/6666666.60966.psg.1560	CDS	1459112	1459246	2	+	135	hypothetical protein	-none-	NA
fig/6666666.60966.psg.1597	CDS	1486647	1486513	-3	-	135	hypothetical protein	-none-	NA
fig/6666666.60966.psg.1949	CDS	1824258	1824124	-3	-	135	hypothetical protein	-none-	NA
fig/6666666.60966.psg.2059	CDS	1916861	1916995	2	+	135	hypothetical protein	-none-	NA
fig/6666666.60966.psg.2151	CDS	1995848	1995982	2	+	135	FIG00858674; hypothetical protein	-none-	NA
fig/6666666.60966.psg.2230	CDS	2069201	2069335	2	+	135	hypothetical protein	-none-	NA
fig/6666666.60966.psg.2363	CDS	2194044	2194178	3	+	135	LSU ribosomal protein L34p	Cell Division Subsystem including YidCD; RNA modification cluster	NA
fig/6666666.60966.psg.2527	CDS	2338623	2338489	-3	-	135	hypothetical protein	-none-	NA
fig/6666666.60966.psg.1276	CDS	1182314	1182177	-2	-	138	hypothetical protein	-none-	NA
fig/6666666.60966.psg.1317	CDS	1218021	1218158	3	+	138	hypothetical protein	-none-	NA
fig/6666666.60966.psg.1606	CDS	1493209	1493072	-1	-	138	hypothetical protein	-none-	NA
fig/6666666.60966.psg.1867	CDS	1739233	1739370	1	+	138	Error-prone, lesion bypass DNA polymerase V(UmuC)	-none-	NA
fig/6666666.60966.psg.2117	CDS	1966387	1966250	-1	-	138	hypothetical protein	-none-	NA
fig/6666666.60966.psg.2227	CDS	2068598	2068735	2	+	138	hypothetical protein	-none-	NA
fig/6666666.60966.psg.163	CDS	160970	160830	-2	-	141	hypothetical protein	-none-	NA
fig/6666666.60966.psg.173	CDS	168825	168965	3	+	141	hypothetical protein	-none-	NA
fig/6666666.60966.psg.1286	CDS	1189481	1189621	2	+	141	hypothetical protein	-none-	NA
fig/6666666.60966.psg.1416	CDS	1328534	1328674	2	+	141	hypothetical protein	-none-	NA
fig/6666666.60966.psg.1422	CDS	1330733	1330873	2	+	141	hypothetical protein	-none-	NA

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld	C91 Alias
fig/66666666.60966.psg.1759	CDS	1635699	1635839	3	+	141	hypothetical protein	-none-	NA	NA
fig/66666666.60966.psg.1775	CDS	1649712	1649572	-3	-	141	Mobile element protein	-none-	NA	NA
fig/66666666.60966.psg.1984	CDS	1850712	1850572	-3	-	141	hypothetical protein	-none-	NA	NA
fig/66666666.60966.psg.2003	CDS	1866906	1867046	3	+	141	hypothetical protein	-none-	NA	NA
fig/66666666.60966.psg.2080	CDS	1935668	1935808	2	+	141	hypothetical protein	-none-	NA	NA
fig/66666666.60966.psg.2487	CDS	2301272	2301132	-2	-	141	hypothetical protein	-none-	NA	NA
fig/66666666.60966.psg.369	CDS	342367	342224	-1	-	144	hypothetical protein	-none-	NA	NA
fig/66666666.60966.psg.635	CDS	579509	579652	2	+	144	hypothetical protein	-none-	NA	NA
fig/66666666.60966.psg.1001	CDS	928837	928980	1	+	144	hypothetical protein	-none-	NA	NA
fig/66666666.60966.psg.1642	CDS	1524746	1524603	-2	-	144	hypothetical protein	-none-	NA	NA
fig/66666666.60966.psg.2114	CDS	1962886	1963029	1	+	144	hypothetical protein	-none-	NA	NA
fig/66666666.60966.psg.2593	CDS	2391861	2392004	3	+	144	hypothetical protein	-none-	NA	NA
fig/66666666.60966.psg.205	CDS	197036	196890	-2	-	147	Integrase	-none-	NA	NA
fig/66666666.60966.psg.284	CDS	260995	261141	1	+	147	hypothetical protein	-none-	NA	NA
fig/66666666.60966.psg.1885	CDS	1759287	1759141	-3	-	147	hypothetical protein	-none-	NA	NA
fig/66666666.60966.psg.2453	CDS	2272001	2271855	-2	-	147	hypothetical protein	-none-	NA	NA
fig/66666666.60966.psg.2620	CDS	2417974	2417828	-1	-	147	hypothetical protein	-none-	NA	NA
fig/66666666.60966.psg.2623	CDS	2420545	2420399	-1	-	147	hypothetical protein	-none-	NA	NA
fig/66666666.60966.psg.108	CDS	106117	106266	1	+	150	hypothetical protein	-none-	NA	NA
fig/66666666.60966.psg.813	CDS	746234	746085	-2	-	150	hypothetical protein	-none-	NA	NA
fig/66666666.60966.psg.1296	CDS	1205130	1205279	3	+	150	hypothetical protein	-none-	NA	NA
fig/66666666.60966.psg.1334	CDS	1235992	1235843	-1	-	150	hypothetical protein	-none-	NA	NA
fig/66666666.60966.psg.2300	CDS	2134634	2134485	-2	-	150	hypothetical protein	-none-	NA	NA
fig/66666666.60966.psg.35	CDS	36062	35910	-2	-	153	hypothetical protein	-none-	NA	NA
fig/66666666.60966.psg.1312	CDS	1215961	1216113	1	+	153	hypothetical protein	-none-	NA	NA
fig/66666666.60966.psg.2143	CDS	1990257	1990105	-3	-	153	hypothetical protein	-none-	NA	NA
fig/66666666.60966.psg.359	CDS	335237	335392	2	+	156	hypothetical protein	-none-	NA	NA
fig/66666666.60966.psg.546	CDS	506105	505950	-2	-	156	FIG00858972; hypothetical protein	-none-	NA	NA
fig/66666666.60966.psg.1174	CDS	1095978	1095823	-3	-	156	Mobile element protein	-none-	NA	NA
fig/66666666.60966.psg.1384	CDS	1296435	1296280	-3	-	156	hypothetical protein	-none-	NA	NA
fig/66666666.60966.psg.2575	CDS	2378324	2378169	-2	-	156	hypothetical protein	-none-	NA	NA
fig/66666666.60966.psg.89	CDS	85217	85059	-2	-	159	hypothetical protein	-none-	NA	NA
fig/66666666.60966.psg.556	CDS	516983	516825	-2	-	159	hypothetical protein	-none-	NA	NA
fig/66666666.60966.psg.562	CDS	523304	523462	2	+	159	hypothetical protein	-none-	NA	NA
fig/66666666.60966.psg.1099	CDS	1029431	1029589	2	+	159	hypothetical protein	-none-	NA	NA
fig/66666666.60966.psg.1571	CDS	1466563	1466405	-1	-	159	hypothetical protein	-none-	NA	NA
fig/66666666.60966.psg.613	CDS	562226	562065	-2	-	162	hypothetical protein	-none-	NA	NA
fig/66666666.60966.psg.886	CDS	814562	814401	-2	-	162	hypothetical protein	-none-	NA	NA
fig/66666666.60966.psg.1114	CDS	1043720	1043559	-2	-	162	hypothetical protein	-none-	NA	NA
fig/66666666.60966.psg.291	CDS	265661	265825	2	+	165	FIG00856904; hypothetical protein	-none-	NA	NA
fig/66666666.60966.psg.2077	CDS	1934012	1933848	-2	-	165	hypothetical protein	-none-	NA	NA
fig/66666666.60966.psg.620	CDS	568234	568401	1	+	168	Methyltransferase (EC 2.1.1.—)	-none-	NA	NA
fig/66666666.60966.psg.2317	CDS	2149354	2149521	1	+	168	Mobile element protein	-none-	NA	NA
fig/66666666.60966.psg.2420	CDS	2241817	2241650	-1	-	168	hypothetical protein	-none-	NA	NA
fig/66666666.60966.psg.856	CDS	785317	785147	-1	-	171	hypothetical protein	-none-	NA	NA

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld	C91 Alias
fig/66666666.609666.psg.1487	CDS	1401876	1401706	-3	-	171	FIG00858878: hypothetical protein	-none-	NA	NA
fig/66666666.609666.psg.2286	CDS	2119703	2119873	2	+	171	hypothetical protein	-none-	NA	NA
fig/66666666.609666.psg.2707	CDS	2514260	2514090	-2	-	171	hypothetical protein	-none-	NA	NA
fig/66666666.609666.psg.862	CDS	790885	790712	-1	-	174	hypothetical protein	-none-	NA	NA
fig/66666666.609666.psg.1450	CDS	1360555	1360382	-1	-	174	Mobile element protein	-none-	NA	NA
fig/66666666.609666.psg.1572	CDS	1466783	1466956	2	+	174	hypothetical protein	-none-	NA	NA
fig/66666666.609666.psg.1608	CDS	1494659	1494835	2	+	177	hypothetical protein	-none-	NA	NA
fig/66666666.609666.psg.2416	CDS	2237303	2237127	-2	-	177	hypothetical protein	-none-	NA	NA
fig/66666666.609666.psg.2668	CDS	2479940	2479764	-2	-	177	hypothetical protein	-none-	NA	NA
fig/66666666.609666.psg.748	CDS	694556	694377	-2	-	180	hypothetical protein	-none-	NA	NA
fig/66666666.609666.psg.1667	CDS	1550475	1550296	-3	-	180	hypothetical protein	-none-	NA	NA
fig/66666666.609666.psg.2116	CDS	1965779	1965600	-2	-	180	hypothetical protein	-none-	NA	NA
fig/66666666.609666.psg.2144	CDS	1990362	1990541	3	+	180	Mobile element protein	-none-	NA	NA
fig/66666666.609666.psg.2400	CDS	2229074	2228895	-2	-	180	hypothetical protein	-none-	NA	NA
fig/66666666.609666.psg.287	CDS	262264	262446	1	+	183	hypothetical protein	-none-	NA	NA
fig/66666666.609666.psg.422	CDS	392649	392831	3	+	183	hypothetical protein	-none-	NA	NA
fig/66666666.609666.psg.1562	CDS	1459862	1460044	2	+	183	hypothetical protein	-none-	NA	NA
fig/66666666.609666.psg.2388	CDS	2218996	2219178	1	+	183	hypothetical protein	-none-	NA	NA
fig/66666666.609666.psg.338	CDS	308508	308323	-3	-	186	hypothetical protein	-none-	NA	NA
fig/66666666.609666.psg.1842	CDS	1720173	1719988	-3	-	186	hypothetical protein	-none-	NA	NA
fig/66666666.609666.psg.286	CDS	262147	261959	-1	-	189	hypothetical protein	-none-	NA	NA
fig/66666666.609666.psg.1202	CDS	1125016	1124828	-1	-	189	hypothetical protein	-none-	NA	NA
fig/66666666.609666.psg.1304	CDS	1209919	1209731	-1	-	189	FIG00858878: hypothetical protein	-none-	NA	NA
fig/66666666.609666.psg.2395	CDS	2224520	2224332	-2	-	189	hypothetical protein	-none-	NA	NA
fig/66666666.609666.psg.250	CDS	233011	233202	1	+	192	hypothetical protein	-none-	NA	NA
fig/66666666.609666.psg.794	CDS	729662	729471	-2	-	192	hypothetical protein	-none-	NA	NA
fig/66666666.609666.psg.1764	CDS	1639415	1639606	2	+	192	hypothetical protein	-none-	NA	NA
fig/66666666.609666.psg.808	CDS	741018	740824	-3	-	195	hypothetical protein	-none-	NA	NA
fig/66666666.609666.psg.276	CDS	255673	255870	1	+	198	Mobile element protein	-none-	NA	NA
fig/66666666.609666.psg.749	CDS	694591	694788	1	+	198	hypothetical protein	-none-	NA	NA
fig/66666666.609666.psg.2718	CDS	2523733	2523536	-1	-	198	hypothetical protein	-none-	NA	NA
fig/66666666.609666.psg.217	CDS	203959	204159	1	+	201	Mobile element protein	-none-	NA	NA
fig/66666666.609666.psg.682	CDS	631759	631959	1	+	201	FIG00859622: hypothetical protein	-none-	NA	NA
fig/66666666.609666.psg.208	CDS	199189	199392	1	+	204	Cold shock protein CspA	Cold shock, CspA family of proteins	NA	NA
fig/66666666.609666.psg.380	CDS	348880	348677	-1	-	204	SSU ribosomal protein S16p	-none-	NA	NA
fig/66666666.609666.psg.2149	CDS	1995706	1995503	-1	-	204	dTDP-4-dehydrohamnose reductase (EC 1.1.1.133)	Rhamnose containing glycans; dTDP-rhamnose synthesis	NA	NA
fig/66666666.609666.psg.1148	CDS	1073050	1073259	1	+	210	hypothetical protein	-none-	NA	NA
fig/66666666.609666.psg.2464	CDS	2281343	2281134	-2	-	210	hypothetical protein	-none-	NA	NA
fig/66666666.609666.psg.522	CDS	484104	484316	3	+	213	LSU ribosomal protein L29p (L35e)	-none-	NA	NA

SUPPLEMENTARY TABLE 1-continued

Annotation of genes in SEQ ID NO: 1.										
Feature ID	Type	Start	Stop	Frame	Strand	Length (bp)	Function	Subsystem	D23 Gbkld	C91 Alias
fig/66666666.60966.psg.1481	CDS	1395113	1394901	-2	-	213	hypothetical protein	-none-	NA	NA
fig/66666666.60966.psg.292	CDS	265809	266024	3	+	216	hypothetical protein	-none-	NA	NA
fig/66666666.60966.psg.511	CDS	475254	475469	3	+	216	SSU ribosomal protein S12p (S23e)	<i>Mycobacterium</i> virulence operon involved in protein synthesis (SSU ribosomal proteins); Ribosomal protein S12p Asp methylthiotransferase	NA	NA
fig/66666666.60966.psg.529	CDS	490129	490347	1	+	219	Translation initiation factor 1	Translation initiation factors bacterial	NA	NA
fig/66666666.60966.psg.2374	CDS	2207422	2207640	1	+	219	hypothetical protein	-none-	NA	NA
fig/66666666.60966.psg.2105	CDS	1956719	1956940	2	+	222	hypothetical protein	-none-	NA	NA
fig/66666666.60966.psg.2690	CDS	2502317	2502538	2	+	222	hypothetical protein	-none-	NA	NA
fig/66666666.60966.psg.2153	CDS	1996354	1996578	1	+	225	hypothetical protein	-none-	NA	NA
fig/66666666.60966.psg.2138	CDS	1985283	1985510	3	+	228	hypothetical protein	-none-	NA	NA
fig/66666666.60966.psg.442	CDS	405529	405299	-1	-	231	hypothetical protein	-none-	NA	NA
fig/66666666.60966.psg.1716	CDS	1598000	1597770	-2	-	231	hypothetical protein	-none-	NA	NA
fig/66666666.60966.psg.2439	CDS	2259000	2258770	-3	-	231	hypothetical protein	-none-	NA	NA
fig/66666666.60966.psg.1111	CDS	1042874	1043110	2	+	237	Mobile element protein	-none-	NA	NA
fig/66666666.60966.psg.1999	CDS	1865704	1865468	-1	-	237	hypothetical protein	-none-	NA	NA
fig/66666666.60966.psg.1214	CDS	1136308	1136069	-1	-	240	hypothetical protein	-none-	NA	NA
fig/66666666.60966.psg.1733	CDS	1614273	1614031	-3	-	243	hypothetical protein	-none-	NA	NA
fig/66666666.60966.psg.48	CDS	47253	47498	3	+	246	Alpha-L-Rha alpha-1,3-L-rhamnosyltransferase (EC 2.4.1.—)	Rhamnose containing glycans	NA	NA
fig/66666666.60966.psg.423	CDS	393586	393338	-1	-	249	Mobile element protein	-none-	NA	NA
fig/66666666.60966.psg.429	CDS	396754	396506	-1	-	249	Transglycosylase-associated protein	-none-	NA	NA
fig/66666666.60966.psg.1147	CDS	1072594	1072842	1	+	249	hypothetical protein	-none-	NA	NA
fig/66666666.60966.psg.1951	CDS	1824656	1824408	-2	-	249	Plasmid stabilization system protein	-none-	NA	NA
fig/66666666.60966.psg.2579	CDS	2380195	2380446	1	+	252	Mobile element protein	-none-	NA	NA
fig/66666666.60966.psg.523	CDS	484294	484548	1	+	255	SSU ribosomal protein S17p (S11e)	-none-	NA	NA
fig/66666666.60966.psg.1022	CDS	948522	948782	3	+	261	Stringent starvation protein B	Carbon Starvation	NA	NA
fig/66666666.60966.psg.1957	CDS	1826664	1826936	3	+	273	hypothetical protein	-none-	NA	NA
fig/66666666.60966.psg.2273	CDS	2108052	2108336	3	+	285	putative DNA transport competence protein	-none-	NA	NA
fig/66666666.60966.psg.46	CDS	46807	46481	-1	-	327	Conserved domain protein	-none-	NA	NA
fig/66666666.60966.psg.2109	CDS	1959886	1960239	1	+	354	hypothetical protein	-none-	NA	NA
fig/66666666.60966.psg.2429	CDS	2248841	2248488	-2	-	354	hypothetical protein	-none-	NA	NA
fig/66666666.60966.psg.1579	CDS	1468834	1469334	1	+	501	hypothetical protein	-none-	NA	NA
fig/66666666.60966.psg.293	CDS	266114	266689	2	+	576	hypothetical protein	-none-	NA	NA

SUPPLEMENTARY TABLE 2

Selected D23 sequences		
SEQ ID and length	Description	Sequence
SEQ ID amoC1 No: 4 (D23_1c22 (271 72) aa) protein		MATTLGTSSASSVSSRGYDMSLWYDSKFYKFLITMLLVAIFWVWYQRYF AYSHGMDSMPEFDRVWMLWRVHMAIMPLFALVTVGWNIWKTRDTEEQLN NLDPKLEIKRYFYIYMMWLVGYIFGVYWGGSFFTEQDASWHQVIIRDTSFT PSHVVPYFGSFPYIVCGVATYLYAMTRLPLFHRGISFPLVMAIAGPLMI LPNVGLNEWGHAFWFMEELFSAPLHWGFVVLGWAGLFQGGVAAQIITRYS NLTDVVWNQSKIELNNRVVA
SEQ ID amoC1 No: 5 (D23_1c22 (816 72) DNA nt)		ATGGCAACTACGTTAGGAACGAGCAGTGCCTCATCAGTCTCATCAAGAGG CTATGACATGTCACTGTGGTATGACTCCAAATTTTATAAATTTGGTTTAA TAACCATGTTGTTGGTAGCGATATTCTGGGTATGGTATCAACGTTACTTT GCCTATTACACGGAATGGATTCAATGAACCAAGATTGACCGTGTATG GATGGGCTGTGGCGTGTGCACATGGCCATTATGCCGCTGTTTGCACCTGG TAACCTGGGGCTGGATCTGGAAAACACGTGATACAGAAGAGCAATTGAAT AACCTGGATCCGAACTGGAAATCAAACGCTACTTCTACTACATGATGTG GCTGGGTGTATACATTTTGGTGTCTACTGGGGTGGTAGCTTCTTTACGG AGCAAGATGCCTCCTGGCACCAGGTGATTATTCTGTACACCAGCTTTACA CCAAGTACGTCAGTCGTGTTTATGGATCATTCCGATGTACATCGTCTG CGGAGTTGCAACCTATCTGTATGCAATGACCCGCTGTCGCCGCTGTTTCATC GTGGAATTTCTTTCCCACTGGTGATGGCGATTGCAGGTCCCTCTGATGATT CTGCCAAACGTTGGTCTGAATGAATGGGGTCATGCTTCTGGTTCATGGA AGAGCTGTTTCAGCGCACCGCTGCATTGGGGTTTTGTAGTGTCTCGGTTGGG CTGGGTTATTCCAGGGTGGAGTTGCTGCCCAAATCATTACCCGTTATTCC AACCTGACTGACGTGGTCTGGAATAATCAAAGCAAAGAAATCTGAATAA CCGGGTTGTAGCTTAG
SEQ ID amoA1 No: 6 (D23_1c22 (276 71) aa) protein		VSIFRTEEILKAAKMPPPEAVHMSRLIDAVYFPIVLVLLVGTYHMHFLLA GDWDFWMDWKDRQWVPVVTPIVGITYCSAIMYYLWVNYRQPFGATLCVVC LLIGEWLTRYWGFYWWSHYPLNFVTPGIMLPGALMLDFTMYLTRNLVTA LVGGGFFGLLFYPGNWAI FGP THLP I VVEG TLLSMADYMGHLYVRTGTPE VYRHI EQGSLRTPFGGHTTVIAAFPAFVSMMLFAVWVWYLGKVCYTAFYFV KGRKRIVQRNDVTA FGE EGFPEGIK
SEQ ID amoA1 No: 7 (D23_1c22 (831 71) DNA nt)		GTGAGTATATTTAGAACAGAAGAGATCCTGAAAGCGGCCAAGATGCCGCC GGAAGCGGTCCATATGTCACGCCTGATTGATGCGGTTTATTTCCGATTCT TGGTGTCTCTGTTGGTAGGTACCTACCATATGCACTTCATGTTGTTGGCA GGTGACTGGGATTTCTGGATGGACTGGAAAGATCGTCAATGGTGGCCTGT AGTAACACCTATTGTAGGCATTACCTATTGCTCGGCAATTATGTATTACC TGTGGGTCAACTACCGTCAACCATTGGTGCAGCTCTGTGCGTAGTGTGT TTGCTGATAGGTGAGTGGCTGACACGTTACTGGGGTTTCTACTGGTGGTC ACACTATCCACTCAATTTGTAAACCCAGGTATCATGCTCCCGGGTGCAAT TGATGTTGGATTTCACATGTATCTGACACGTAACCTGGTGGTACTGCA TTGGTTGGGGTGGATTCTTTGGTCTGCTGTTTACCCGGGTAACCTGGGC AATCTTTGGTCCGACCATCTGCCAATCGTTGAGAAGGAACACTGTTGT CGATGGCTGACTATATGGGTACCTGTATGTTGTCACGGGTACACCTGAG TATGTTTCGTATATTGAACAGGTTTATTACGTACCTTTGGTGGTCACAC CACAGTTATTGCAGCATTTCTCGCTGCGTTTGTATCCATGCTGATGTTT CAGTCTGGTGGTATCTTGGAAAAGTTTACTGCACAGCTTCTTCTACGTT AAAGGTAAGAGGACGTATCTGTCAGCGCAATGATGTACGGCATTGAG TGAAGAAGGGTTTCCAGAGGGGATCAATAA
SEQ ID amoB1 No: 8 (D23_1c22 (421 70) aa) protein		MGIKNLYKRGMMGLCGVAVYAMAALMTVTLDVSTVAHGERSQEPFLRM RTVQWYDVWKGPVETKVNENAIQTKFHLAEDWPRAAARPDFAFFNVGSP SSVYVRLSTKINGHPWFISGLPLQIGRDYAFEVQLRARI PGRHHMHAMLN KDAGPIAGPGAWMNI TGSWDDFTNPLKLLTGETIDSETFNLSNGIFWHIL WMSITFIWIGIFVARPMFLPRSRVLLAYGDDLLLDPMDDKKITWVLAITL AIVWGGYRYTETKHPYTVPIQAGQSKVAPLPVAPNPVAIKITDANYDVPG RALRVSMEVNTMNGDTPVTFGEFTTAGIRFVNSTGRKYLDPQYPRELVAVG LNFDDDGAIQPGETKQLRMEAKDALWEIQRLMALLGDPESRFGGLMSWD SEGNRHINSIAGPVI PVFTKL
SEQ ID amoB1 No: 9 (D23_1c22 (1266 70) DNA nt)		ATGGGTATCAAGAACCTTTATAAAGCTGGAATGATGGGACTTTGTGGCGT TGCTGTTTATGCAATGGCGGCACGTGACCATGACAGTGACACTAGATGTCT CAACAGTAGCAGCCCATGGAGAACGATCCCAGGAACCGTTTCTTCCGATG CGTACAGTACAGTGGTACGATGTTAAGTGGGGTCCGGAAGTAACCAAAGT CAATGAGAATGCCAAATTACCGGCAATTTCACTTGGCTGAAGACTGGC CGCGTGGCGCAGCAAGACCGGATTTGCATTCTTTAACGTAGGTAGCCCA AGCTCGGTATACGTGCGTTTGTAGTACGAAGATTAATGGCCACCCATGGTT TATTTACAGTCCGCTGCAAAATGGTCTGTGACTATGCGTTCGAAGTTCAGC TGAGAGCACGTATTCAGGACGCCATCACATGCACGCCATGTTAAACGTT AAAGATGCAGGTCCAATTGCAGGACCGGGTGCATGGATGAACATTACCGG AAGCTGGGATGATTTTACTAATCCACTCAAGCTGCTGACAGGCGAAACAA

SUPPLEMENTARY TABLE 2-continued

Selected D23 sequences	
SEQ ID and length	Description Sequence
	TTGACTCAGAAACATTCAACCTGTCAAACGGTATTTTCTGGCATATTCTC TGGATGTCAATTGGTATATTTTGGATTGGTATCTTTGTAGCGCTCCGAT GTTCCTGCCACGTAGCCGGGTATTGCTCGCTTATGGTGATGATCTGTTGC TGGATCCGATGGATAAGAAAATCACCTGGGTACTTGCAATCCTGACCTTG GCTATAGTATGGGTGGATACCGCTATACAGAAACCAAGCATCCATACAC AGTACCTATCCAGGCTGGTCAATCCAAAGTTGCACCATTACCCGTAGCAC CAAATCCGGTAGCAATCAAATTACAGATGCTAACTATGACGTACCGGGA CGTGCACTGCGTGTATCGATGGAAGTAACCAACACGGTGATACACAGT CACATTTGGTGAATTTACCAAGCAGGTATTCTGTTTCGTTAACAGTACCG GCCGCAAGTACCTGGATCCACAGTATCCTCGTGAACGTGGTTGCAGTAGGC TTGAATTTTGATGATGATGGTGCAATTGAGCCAGGCGAGACCAAGCAATT GAGGATGGAAGCCAAAGATGCTCTGTGGGAAATCCAACGTCTGATGGCGT TGCTGGGTGACCCGGAAGCCGTTTGGTGGACTGTTAATGTCTTGGGAT TCAGAAGGTAATCGCCATATCAACAGTATTGCTGGTCCGGTGATTCCAGT CTTTACCAAGCTCTAA
SEQ ID amoC2 NO: 10 (D23_1c25 (271 12) aa) protein	MATTLGTSSASSVSRGYDMSLWYDSKFYKFLITMLLVAIFWVWYQRYF AYSHGMSMEPEFDRVWMLWRVHMAIMPLFALVTGWGIWIKTRDTEQLN NLDPKLEIKRYFYMMWLVYIFGVYWGGSFTEQDASWHQV IIRDTST PSHVVFVYGSFPMYIVCGVATYLYAMTRLPLFHRGISFPLVMAIAGPLMI LPNVGLNEWGHAFWFMEELFSAPLHWGFVVLGWAGLFQGGVAAQIITRYS NLTDVVWNNQSKIELNNRVVA
SEQ ID amoC2 NO: 11 (D23_1c25 (816 12) DNA nt)	ATGGCAACTACGTTAGGAACGAGCAGTGCCTCATCAGTCTCATCAAGAGG CTATGACATGTCACTGTGGTATGACTCCAAATTTTATAAATTGGTTTAA TAACCATGTTGTTGGTAGCGATATTTCTGGGTATGGTATCAACGTACTTT GCCTATTACACCGGAATGGATTCAATGGAACAGAGTTTGACCGTGTATG GATGGGCTGTGGCGTGTGCACATGGCCATTATGCCGCTGTTTGCACTGG TAACCTGGGGCTGGATCTGGAACACAGTGATCAGAAAGAGCAATTGAAT AACCTGGATCCGAACTGGAATCAAACGCTACTTCTACTACATGATGTG GCTGGGTGTATACATTTTGGTGTCTTACTGGGGTGGTAGCTTCTTTACGG AGCAAGATGCCTCCTGGCACCAGGTGATTATTCGTGACACCAGCTTTACA CCAAGTCACGTAGTCGTGTTTATGGATCATTTCCGATGTACATCGTCTG CGGAGTTGCAACCTATCTGTATGCAATGACCCGCTGCGCGCTGTTTCATC GTGGAATTTCTTTCCACTGGTGTATGGCGATTGCAAGTCTCTGATGATT CTGCCAAACGTTGGTCTGAATGAATGGGGTCATGCTTTCTGGTTTCATGGA AGAGCTGTTTCAGCGCACCGCTGCATTGGGGTTTGTAGTGCTCGGTTGGG CTGGGTATTCCAGGGTGGAGTTGCTGCCAAATCATTACCCGTTATTCC AACCTGACTGACGTGGTCTGGAATAATCAAAGCAAAGAAATCTGAATAA CCGGTTGTAGCTTAG
SEQ ID amoA2 NO: 12 (D23_1c25 (276 11) aa) protein	VSIFRTEEILKAAKMPPEAVHMSRLIDAVYFPILVVLVGYTHMHFMLLA GDWDFWMDWKDRQWVPVVTPIVGITYCSAIMYYLWVNYRQPFGLTLCVVC LLIGEWLTRYWGFYWWSHYPLNFVTPGIMLPALMLDFTMYLTRNLVTA LVGGGFFGLLFYPGNWAI FGPTHLPIVVEGTLMSADYMGHLYVRTGTPE YVRHIEQSLRTFGGHTTVIAAFPAFVSMMLFAVWWYLKVVYCTAFYFV KGKGRIVQRNDVTAFGEEGFPEGIK
SEQ ID amoA2 NO: 13 (D23_1c25 (831 11) DNA nt)	GTGAGTATATTTAGAACAGAAGAGATCCTGAAAGCGGCCAAGATGCCGCC GGAAGCGGTCCATATGTACGCCTGATGTATGCGGTTTATTTCCGATTCT TGGTTGTTCTGTTGGTAGGTACCTACCATATGCACTTCATGTTGTTGGCA GGTGACTGGGATTTCTGGATGGACTGGAAAGATCGTCAATGGTGGCCTGT AGTAACACCTATTGTAGGCATTACCTATTGCTCGGCAATTATGTATTACC TGTGGGTCAACTACCGTCAACCATTTGGTGGCACTCTGTGCGTAGTGTGT TTGCTGATAGGTGAGTGGCTGACACGTTACTGGGGTTTCTACTGGTGGTC ACATATCCACTCAATTTGTAAACCCAGGTATCATGCTCCCGGTGTCAT TGATGTTGGATTTACAAATGTATCTGACACGTAACCTGGTGGTACTGCA TTGTTGGGGTGGATTTCTTGGTCTGCTGTTTACC CGGTAACCTGGGC AATCTTGGTCCGACCATCTGCCAATCGTTGTAGAAGGAACACTGTTGT CGATGGCTGACTATATGGGTACCTGTATGTTCTGACGGGTACACCTGAG TATGTTCTGTCATATTGAACAAGGTTTATTACGTACCTTTGGTGGTCACAC CACAGTTATTGCAGCATTTCTCGCTGCGTTTGTATCCATGCTGATGTTGT CAGTCTGGTGGTATCTTGGAAAGTTTACTGCACAGCCTTCTTCTACGTT AAAGGTAAGAGAGGACGTATCGTGACGCGCAATGATGTACGGCATTTGG TGAAGAAGGGTTTCCAGAGGGGATCAAATAA
SEQ ID amoB2 NO: 14 (D23_1c25 (421 10) aa) protein	MGIKNLYKRMMLGCGVAVYAMAALMTVTLVDVSTVAHGRSQQEPFLRM RTVQWYDVWKWGEVTKVNENAQITGKFHLAEDWPRAAARPDFAFNVGSP SSVYVRLSTKINGHPWFISGPLQIGRDYAFEVQLRARI PGRHHMHAMLVN KDAGPIAGPGAWMNI TGSWDDFTNPLKLLTGETIDSETFNLSNGIFWHIL WMSIGIFWIGIFVARPMFLPRSRVLLAYGDDLLDPMDDKKTWVLAILTL

SUPPLEMENTARY TABLE 2-continued

Selected D23 sequences	
SEQ ID and length Description Sequence	
	AIVWGGYRYTETKHPYTVPIQAGQSKVAPLPVAPNPVAIKITDANYDVPGRALRVSMVETNNGDTPVTFGEFTTAGIRFVNSTGRKYLDPOYPRELVAVGLNFDDDGAIQPGETKQLRMEAKDALWEIQRLMALLGDPESRFGGLMSWDSEGNRHINSIAGPVIPTFKL
SEQ ID amoB2 NO: 15 (D23_1c25 (1266 10) DNA nt)	ATGGGTATCAAGAACCTTTATAAACGTGGAATGATGGGACTTTGTGGCGTTGCTGTTTATGCAATGGCGGCACTGACCATGACAGTGACACTAGATGTCTCAACAGTAGCAGCCCATGGAGAACGATCCCAGGAACCGTTTCTTCGGATGCGTACAGTACAGTGGTACGATGTTAAGTGGGGTCCGGAAGTAACCAAAGTCAATGAGAATGCCAAATTACCGGCAAAATTCACCTGGCTGAAGACTGGCCGCTGCGGCAGCAAGACCGGATTTCGCATTCTTAACGTAGGTAGCCCAAGCTCGGTATACGTGCGTTTGAGTACGAAGATTAATGGCCACCCATGGTTTATTCAGGTCCGCTGCAAATTGGTCTGACTATGCGTTCGAAGTTCAGCTGAGAGCAGTATTCCAGGACGCCATCACATGCACGCCATGTTAAACGTTAAAGATGCAGGTCCAATTGCAGGACCGGTGCATGGATGAACATTACCGGAAGCTGGGATGATTTTACTAATCCAACCTCAAGCTGCTGACAGGCGAAACAAATTGACTCAGAAACATTCAACCTGTCAAACCGGTATTTCTGGCATATTCTCTGGATGTCAATTGGTATATTTTGGATTGGTATCTTTGTAGCGCGTCCGATGTTCTTCCACGTAGCCCGGTATTGCTCGCTTATGGTGATGATCTGTTGCTGGATCCGATGGATAAGAAAATCACCTGGTACTTGCAATCCTGACCTTGCTATAGTATGGGTGGATACCGCTATACAGAAACCAAGCATCCATACACAGTACCTATCCAGGCTGGTCAATCCAAAGTTGACCATTACCGGTAGCACAAATCCGGTAGCAATCAAAATTACAGATGCTAATATGACGTACCGGGAAGTGCATGCGTGTATCGATGGAAGTAACCAACACCGGTGATACACCAAGTCAACATTTGGTGAATTTACCAAGCAGGATTCGTTTCGTTAACAGTACCGGCCGACAGTACCTGGATCCACAGTATCCTCGTGAACGGTTGCAGTAGGCTTGAATTTTGATGATGATGGTGCAATTGAGCCAGGCGAGACCAAGCAATTGAGGATGGAAGCCAAAGATGCTCTGTGGGAAATCCAACGCTGATGGCGTTGCTGGGTGACCCGGAAGCCGTTTGGTGAGCTGTTAATGTCTGGGATTCAGAAGGTAATCGCCATATCAACAGTATTGCTGGTCCGGTGATTCCAGTCTTTACCAAGCTCTAA
SEQ ID amoC3 NO: 16 (D23_1c16 (274 05) aa) protein	MATNLIKDKAAQQVADKPTYDKSEWFDKYYKFGLLPILAVAVMWVYFQRTYAYSHGMDSMPEFDRIMGLWRVQMAALPLIALFTWGWLYKTRNTAEQLANLTPKQEI KRYFYFLMWLVYIFAVYWGSSFFTEQDASWHQV IIRDTSFPSHIPLFYGSFPVYIIMGVSMIIYANRLPLYNKWSFPLIMTVAGPLMSLPNVGLNEWGHAFWFMEELFSAPLHWGFVILAWAALFQGGGLAVQI IARFNNLLDVEWNNKQDRAILDDVVTPA
SEQ ID amoC3 NO: 17 (D23_1c16 (825 05) DNA nt)	ATGGCTACAAATATATTAAAAGACAAAGCTGCACAGCAGGTTGCTGATAAACCAACTTATGATAAAATCCGAGTGGTTTGATGCTAAATACTATAAAATTCGGCTGCTACCTATCTTAGCTGTAGCTGTGATGTGGGTTTATTTCAGCGCACATACGCCCTATTCTCACGGCATGGATTCAATGGAACCGGAATTTGACCGGATCGATGGGCTTGCGGTGTTCAAATGGCCGCTCTGCCTCTTATAGCACTTTTACGTGGGGATGGTTATATAAAACCCGCAATACTGCAGAACAGCTTGCCAACTGACTCCAAAGCAGGAAATAAAGCGGTATTTCTATTTCTCATGTGGCTTGCGGCTATATATTGCAAGTTACTGGGGATCAAGCTTCTTTACCGAGCAGGACGCTTCATGGCACCAGGTGATTATCAGGGATACAAGTTTTACTCCTAGCCATATTCCTCTGTTTATGGTTCATTCCCGGTATACATCATCATGGGAGTATCGATGATTATTTACGCCAACACCCGGTTGCCGCTGTACAACAAGGGTGGTCATTCCTCTGATCATGACCGTAGCAGGACCGTTGATGAGTCTGCCTAACGTTGGCCTGAACGAGTGGGGACACGCCTTCTGGTTCATGGAAGAAGCTTTTCAGCGCACCGCTGCCTGGGGCTTCGTGATTCTGGCTTGGGCTGCCCTGTTCCAGGGTGGGCTTGCAAGTACAGATCATAGCTCGCTTTTCCAACTTGCTTGACGTGGAGTGAATAAACAAGACAGAGCCATATTGGACGATGTCGTAACCTGCTCTTAA
SEQ ID hao1 NO: 18 (D23_1c25 (570 29) aa) protein	MRLGEYLGKMLLCAGLLIGPVQADISTVPDETYEALKLDRSKATPKETYDALVKRYKDPAHGAGKGTMGDYWEPIALSIYMDPSTFYKPPVSPKEIAERKDCVECHSDETPVWVRAWKRSTHANLDKIRNLKPEDPLFYKKGKLEEVENLNRSMGLKEKEALKEVGCIDCHVDINAKKADHTKDVRMPTADVCGTCHLREFAERESERDTMIWPNGQWPDGRPSHALDYTANIETTVAAMPQREVAEGCTMCHTNQNKCDNCHTRHEFSAAESRKPEACATCHSGVDHNNYEAYIMSKHGKLAEMNRENWNWNVRLKDAFSKGGQTAPTCAACHMEYEYETHNITRKTRWANYPPVPGIAENITSDWSEARLDSWVVTCTQCHSERFARSYLDLMDKGTLEGLAKYQEANAIVHKMYEDGTLTGQKTNRPNPAPKEKPGFIFTQLFWSKGNPNASLELKVLEMAENNLAKMHVGLAHVNPNGGWTYTEGWPMNRAYVEIQDEYTKMQEMTALQARVNKLEGGKTSLLDLKGAGEKISLGGGGLGMLLAGAIALIGWRKRKQTQA

SUPPLEMENTARY TABLE 2-continued

Selected D23 sequences	
SEQ ID and length Description Sequence	
SEQ ID hao1 NO: 19 (D23_1c25 (1713 29) DNA nt)	ATGAGATTAGGGGAGTATTTGAAGGGGATGCTGCTGTGTGCGGGCCTGTT GTTGATTGGGCCGGTACAGGCGGATATATCGACGGTACCGGATGAGACGT ATGAAGCATTGAAGCTGGATCGCAGCAAAGCCACGCCGAAAGAGACCTAT GATGCGCTGGTGAAGCGTTACAAGGATCCTGCACATGGTGTGCGCAAGGG CACGATGGGAGACTACTGGGAACCGATAGCGCTTAGTATCTACATGGACC CGAGCACCTTTTACAAACCACCGGTTTCCCGAAAGAAATTGCTGAGCGC AAAGACTGCGTTGAATGCCACTCTGATGAAACGCCGGTTTGGGTAAGAGC ATGGAACCGCAGCACCCACGCCAACCTGGACAAAATACGCAACCTCAAGC CGGAAGATCCGCTTTTACAAAAAAGGCAAGCTGGAAGAAGTTGAGAAC AACCTGCGCTCCATGGGCAAACCTGGAGAGAAGGAAGCGCTCAAGGAAGT AGGCTGTATTGACTGTACGTTGACATCAACGCCAAAAAGAAAGCAGATC ACACCAAAGACGTACGATGCCTACAGCTGACGTTTGCAGAACCTGTAC CTGAGAGAATTGCGCAGCGTGAATCCGAGCGTGACACCATGATCTGGCC GAATGGCCAGTGGCCTGACGAGCTCCATCCCACGCACTGGACTACACAG CCAACATTGAAACCACCGTTTGGGCAGCCATGCCGCAACGTGAAGTGGCA GAAGGTTGCACCATGTGCCACCAACAGAACAAATGCGACAACTGCCA TACCCGCCATGAATTTTCGGCGGCAGAAATCCCGCAAACCGGAAGCCTGTG CCACCTGTACAGCGCGGTGGATCATAATAACTATGAAGCCTACATTATG TCCAAGCAGCGCAAACCTGGCTGAAATGAACCGGAGAACTGGAAGTGGAA TGTTCTGCTGAAAGACGCCCTTCTCCAAGGAGGTGACACCGCACCGACCT GTGACGCTGCCACATGGAATACGAAGGGGAATATACCCATAACATCACC CGTAAGACCCGCTGGGCAAACCTACCCGTTTGTTCGGGGATTGCGAGAAA CATCACCGAGCTGGTCAGAAGCTCGTCTGGATTCTGGGTTGTGACCT GTACCCCAATGTCACTCGGAACGGTTTGCCGCTCCTACCTGGATCTCATG GACAAAGGTACCTGGAAGGGTTGGCTAAATACCAGGAAGCCAATGCCAT TGTTTCAAAAATGTATGAAGACGGCACCCCTGACTGGTCAAAAAACCAATC GTCCGAATCCACCGCGCCAGAGAAACCGGGTTTGGTATCTTCAACCCAA CTGTTCTGGTCCAAGGGTAACAACCCGGCCTCACTTGAACGAAAGTGTCT GGAAATGGCAGAAAAACCTGGCCAAAATGCACGTAGGACTGGCACACG TTAATCCAGGTGGCTGGACATATACCGAAGGTGGGGCCCGATGAACCTG GCCATGTGTGAAATCCAGGATGAATACCAAGATGCAGGAAATGACAGC TCTGCAAGCGCGTGTTAACAACCTGGAAGGTAAAAAACCGACCTGCTTG ACCTCAAGGGAGCGGGAGAAAAGATCTCGCTGGGAGGACTGGGAGGTGGC ATGTTGCTGGCGGAGCGATTGCTCTGATTGGCTGGCGTAAACGTAAGCA AACGCAAGCTTGA
SEQ ID hao2 NO: 20 (D23_1c19 (570 26) aa) protein	MRLGEYLGKMLLCAGLLIGPVQADISTVPDETYEALKLDRSKATPKETY DALVKRYKDPAHGAGKGTMGDYWEPIALS IYMDPSTFYKPPVSPKEIAER KDCVCHSDETPVWVRAWKRSTHANLDKI RNLPEDPLFYKKGKLEEVEN NLRSMGLKEKEALKEVGCIDCHVD INAKKADHTKDV RMPADVC GTCH LREFAERESERD TMIWPNGQWPDGRPSHALDY TANIETTVAAMPQREVA EGCTMCHTNQNKCDNCHTRHEFSAAESRKPEACATCHSGVDHNNYEAYIM SKHGLKLEMNRENWNWNVRLKDAFSKGGQTAPTCAACHMEYEGEYTHNIT RKTRWANYPPFVPIAENITSDWSEARLDSWVVTCTQCHSERFARSYLDLM DKGTLEGLAKYQEANAIVHKMYEDGTLTGQKTNRPNPAPAEKPGFIFTQ LFWLKGNNPASLELKVLEMAENNLAKMHVGLAHVNPGGWYTEGWGPMNR AYVEIQDEYTKMQEMTALQARVNLKLGKTSLLDLKGAGEKI SLGGLGGG MLLAGAIALIGWRKRKQTQA
SEQ ID hao2 NO: 21 (D23_1c19 (1713 26) DNA nt)	ATGAGATTAGGGGAGTATTTGAAGGGGATGCTGCTGTGTGCGGGCCTGTT GTTGATTGGGCCGGTACAGGCGGATATATCGACGGTACCGGATGAGACGT ATGAAGCATTGAAGCTGGATCGCAGCAAAGCCACGCCGAAAGAGACCTAT GATGCGCTGGTGAAGCGTTACAAGGATCCTGCACATGGTGTGCGCAAGGG CACGATGGGAGACTACTGGGAACCGATAGCGCTTAGTATCTACATGGACC CGAGCACCTTTTACAAACCACCGGTTTCCCGAAAGAAATTGCTGAGCGC AAAGACTGCGTTGAATGCCACTCTGATGAAACGCCGGTTTGGGTAAGAGC ATGGAACCGCAGCACCCACGCCAACCTGGACAAAATACGCAACCTCAAGC CGGAAGATCCGCTTTTACAAAAAAGGCAAGCTGGAAGAAGTTGAGAAC AACCTGCGCTCCATGGGCAAACCTGGAGAGAAGGAAGCGCTCAAGGAAGT AGGCTGTATTGACTGTACGTTGACATCAACGCCAAAAAGAAAGCAGATC ACACCAAAGACGTACGATGCCTACAGCTGACGTTTGCAGAACCTGTAC CTGAGAGAATTGCGCAGCGTGAATCCGAGCGTGACACCATGATCTGGCC GAATGGCCAGTGGCCTGACGAGCTCCATCCCACGCACTGGACTACACAG CCAACATTGAAACCACCGTTTGGGCAGCCATGCCGCAACGTGAAGTGGCA GAAGGTTGCACCATGTGCCACCAACAGAACAAATGCGACAACTGCCA TACCCGCCATGAATTTTCGGCGGCAGAAATCCCGCAAACCGGAAGCCTGTG CCACCTGTACAGCGCGGTGGATCATAATAACTATGAAGCCTACATTATG TCCAAGCAGCGCAAACCTGGCTGAAATGAACCGGAGAACTGGAAGTGGAA TGTTCTGCTGAAAGACGCCCTTCTCCAAGGAGGTGACACCGCACCGACCT GTGACGCTGCCACATGGAATACGAAGGGGAATATACCCATAACATCACC

SUPPLEMENTARY TABLE 2-continued

Selected D23 sequences	
SEQ ID and length Description Sequence	
	CGTAAGACCCGCTGGGCAAACCTACCCGTTTGTTCGGGGATTGCAGAAAA CATCACCAGCGACTGGTCAGAAGCTCGTCTGGATTCTCTGGGTGTGACCT GTACCCAATGTCACTCGGAACGGTTTGCCCGCTCCTACCTGGATCTCATG GACAAAGGTACCTGGAAGGTTGGCTAAATACCAGGAAGCCAATGCCAT TGTTTCAAAAATGTATGAAGACGGCACCTGACTGGTCAAAAAACCAATC GTCCGAATCCACCGCGCCAGAGAAACCGGGTTTGGTATCTTCACCCAA CTGTTCTGGTCCAAGGGTAACAACCGGCCTCACTTGAAGTAAAGTGCT GGAAATGGCAGAAAACAACCTGGCCAAAATGCACGTAGGACTGGCACAG TTAATCCAGGTGGCTGGACATATACCGAAGGTTGGGGCCCGATGAACCGT GCCATATGTTGAAATCCAGGATGAATACCAAGATGCAGGAAATGACAGC TCTGCACGCGCTGTTAACAACCTGGAAGGTAAAAAACACGCTGCTTG ACCTCAAGGGAGCGGGAGAAAAGATCTCGCTGGGAGGACTGGGAGGTGGC ATGTTGCTGGCGGGAGCGATTGCTCTGATTGGCTGGCGTAAACGTAAGCA AACGCAAGCTTGA
SEQ ID hao3 NO: 22 (D23_1c17 (570 88) aa) protein	MRLGEYLGKMLLCAGLLLIGPVQADISTVPDETYEALKLDRSKATPKETY DALVKRYKDPAHGAGKGTMGDYWEPIALSIYMDPSTFYKPPVSPKEIAER KDCVECHSDETPVWVRAWKRSTHANLDKIRNLKPEDPLFYKKGKLEEVEN NLRSMGKLGEKEALKEVGCIDCHVDINAKKADHTKDVRMPTADVCGTCH LREFAERESERDAMIWPNQWPDGRPSHALDYTANIEETVWAAMPQREVA EGCTMCHTNQNKCDNCHTRHEFSAESRKPEACATCHSGVDHNNYEAYIM SKHKLAEEMNRENWNWNVRLKDAFSKGGQTAPTCAACHMEYEGEYTHNI RKTWANYPFVPGIAENITSDWSEARLDSWVVTCTQCHSERFARSYLDLM DKGTLEGLAKYQEAANIVHKMYEDGTLTGQKTNRPNPPEKPGFGIFTQ LFWSKGNPNPASLELKVLEMAENNLAKMHVGLAHVNPGGWTYTEGWPMNR AYVBIQDEYTKMQEMTALQARVNLKLGKTSLLDLKGAGEKISLGGG MLLAGAIALIGWRKRKQTQA
SEQ ID hao3 NO: 23 (D23_1c17 (1713 88) DNA nt)	ATGAGATTAGGGGAGTATTTGAAGGGGATGCTGCTGTGTGCGGGCCTGTT GTTGATTGGGCGGTACAGGCGGATATATCGACGGTACCGGATGAGACGT ATGAAGCATTGAAGCTGGATCGCAGCAAAGCCACGCCGAAAGAGACCTAT GATGCGCTGGTGAAGCGTTACAAGGATCCTGCGCATGGTGTGCGCAAGGG CACGATGGGAGACTACTGGGAACCGATAGCGCTTAGTATCTACATGGACC CGAGCACCTTTTACAAACACCGGTTTCCCGAAAGAAATGCTGAGCGC AAAGACTGCGTTGAATGCCACTCTGATGAAACGCGCGTTTGGGTAAAGAGC ATGGAAACGCGACCCACGCCAACCTGGACAAAAATACGCAACCTCAAGC CGGAAGATCCGCTTTTACAAAAAAGGCAAGCTGGAAGAAGTTGAGAAC AACCTGCGCTCCATGGGCAAACTTGGAGAGAAGGAAGCGCTCAAGGAAGT AGGCTGTATTGACTGTACGTTGACATCAACGCCAAAAAGAAAGCAGATC ACACAAAGACGTACGATGCCTACAGCTGACGTTTGCAGAACCTGTAC CTGAGAGAATTGCGGAGCGTGAATCCGAGCGTGACACCATGATCTGGCC GAATGGCCAGTGGCTGACGACGTCATCCACGCACTGACTACACAG CCAACATTGAAACACCGCTTTGGGCGAGCATGCCGCAACGTGAAGTGGCA GAAGGTTGCACCATGTGCCACACCAACAGAACAAATGCGACAACTGCCA TACCCGCCATGAATTTTCGGCGGCAGAAATCCCGCAACCGGAAGCCTGTG CCACCTGTCAAGCGCGTGGATCATAATAACTATGAAGCTACATTATG TCCAAGCACGGCAAACTGGCTGAAATGAACCGGAGAACTGGAAGTGGAA TGTTCTGCTGAAAGACGCCTTCTCCAAGGAGGTCAGACCGCACCGACCT GTGCAGCCTGCCACATGGAATACGAAGGGGAATATACCCATAACATCACC CGTAAGACCCGCTGGGCAAACCTACCGTTTGTTCGGGGATTGCAGAAAA CATCACCAGCGACTGGTCAGAAGCTCGTCTGGATTCTCTGGGTGTGACCT GTACCCAATGTCACTCGGAACGGTTTGCCCGCTCCTACCTGGATCTCATG GACAAAGGTACCTGGAAGGTTGGCTAAATACCAGGAAGCCAATGCCAT TGTTTCAAAAATGTATGAAGACGGCACCTGACTGGTCAAAAAACCAATC GTCCGAATCCACCGCGCCAGAGAAACCGGGTTTGGTATCTTCACCCAA CTGTTCTGGTCCAAGGGTAACAACCGGCCTCACTTGAAGTAAAGTGCT GGAAATGGCAGAAAACAACCTGGCCAAAATGCACGTAGGACTGGCACACG TTAATCCAGGTGGCTGGACATATACCGAAGGTTGGGGCCCGATGAACCGT GCCATATGTTGAAATCCAGGATGAATACCAAGATGCAGGAAATGACAGC TCTGCAAGCGCGTGTTAACAACCTGGAAGGTAAAAAACACGCTGCTTG ACCTCAAGGGAGCGGGAGAAAAGATCTCGCTGGGAGGACTGGGAGGTGGC ATGTTGCTGGCGGGAGCGATTGCTCTGATTGGCTGGCGTAAACGTAAGCA AACGCAAGCTTGA
SEQ ID c554 NO: 24 cycA1 (235 (D23_1c25 aa) 27) protein	MKIIACGLVAAALFTLTSGQSLAADAPFEGRRKKSSCHKPQAQSWKH HAKAMESLKPVKVEAKQKAKLDPKDYTDKDCVGVCHVDGFGQKGGYTI DSPKPMLTGVGCESCHGPRKRYGRDHRKAGQAFKESGKAPRKTLSKGG DNFNEERCACHLNIEGSPWKGAIPPYTPYTPVEDPKYTFKFDENVKDV AMHEHYKLDGVFDGEPKFKFHDEPQANAKTAKKKG

Selected D23 sequences		
SEQ ID		
and		
length	Description	Sequence
SEQ ID c554		ATGAAAAATAATAAGCCTGCGGACTGGTTGCTGCAGCCCTGTTACCCCT
NO: 25	cycA1	GACAAGTGGGCAGAGTCTGGCAGCGGATGCTCCGTTTGAAGGTCGGA AAAA
(708	(D23_1c25	AGTGCAGTTCCTGTCACAAAACCAAGCCCGAGTCGTGGAAACATACTGCC
nt)	27) DNA	CACGCCAAGGCGATGGAATCGCTCAAGCCCAATGTCAAAGTGAAGCCAA
		ACAAAAGGCCAAACTGGATCCTACCAAGGACTACACCCAGGACAAAGACT
		GCGTAGGTTGCCACGTGGATGGATTGGGCCAGAAAGGCGGTACACGATA
		GACTCCCCCAAACCTGCTGACTGGCGTAGGCTGTGAATCCTGCCACGG
		GCCTGGACGTAAATACCGGGGAGATCACCGCAAGGCCGGGCAAGCATTTG
		AGAAATCGGGCAAAAAGCGCCGCGCAAGACCTGGCAAGCAAGGGGCAA
		GACTTTAATTTTGAAGAACGTTGCAGCGCCTGCCATCTGAACATATGAAGG
		GTACCCCTGGAAGGAGCAAAACCTCCCTATAACCCGTATACACCGGAAG
		TGGATCCGAAATACACCTTCAAGTTTGACGAGATGGTAAAAGACGTCAAA
		GCCATGCACGAGCATTACAAACTGGATGGCGTATTGTACGGAGAGCCTTAA
		ATTCAAGTTCCATGACGAATTCCAGGCCAACGCTAAAAC TGCCAAAAAAG
		GAAAATAG
SEQ ID cycA2		MKII IACGLVAAALFTLTSGQSLAADAPFEGRKCCSSCHKPQAQSWKH TA
NO: 26	(D23_1c19	HA KAMESLKFNVKEAQKAKLDPTKDYTDQKDCVGVCHVDGFGQKGGYTI
(235	24)	DS PKPMLTGVGCESCHGPRGRKYRGDRHKAGQAF EKSGKAPRKT LASKGQ
aa)	protein	DFNFEERCSACHLNYEGSPWKHAKPPYTPYTPPEVDPKYTFKFDEMVKDVK
		AMHEHYKLDGVDPDGEPKFKFHDEFQANAKTAKKKG
SEQ ID cycA2		ATGAAAAATAATAAGCCTGCGGACTGGTTGCTGCAGCCCTGTTACCCCT
NO: 27	(D23_1c19	GACAAGTGGGCAGAGTCTGGCAGCGGATGCTCCGTTTGAAGGTCGGA AAAA
(708	24) DNA	AGTGCAGTTCCTGTCACAAAACCAAGCCCGAGTCGTGGAAACATACTGCC
nt)		CACGCCAAGGCGATGGAATCGCTCAAGCCCAATGTCAAAGTGAAGCCAA
		ACAAAAGGCCAAACTGGATCCTACCAAGGACTACACCCAGGACAAAGACT
		GCGTAGGTTGCCACGTGGATGGATTGGGCCAGAAAGGCGGTACACGATA
		GACTCCCCCAAACCATGCTGACTGGCGTAGGCTGTGAATCCTGCCACGG
		GCCTGGACGTAAATACCGGGGAGATCACCGCAAGGCTGGGCAAGCATTTG
		AGAAATCGGGCAAAAAGCGCCGCGCAAGACCTGGCAAGCAAGGGGCAA
		GACTTTAATTTTGAAGAACGTTGCAGCGCCTGCCATCTGAACATATGAAGG
		GTACCCCTGGAAGGAGCAAAAACCTCCCTATAACCCGTATACACCGGAAG
		TGGATCCGAAATACACCTTCAAGTTTGACGAGATGGTAAAAGACGTCAAA
		GCCATGCACGAGCATTACAAACTGGATGGCGTATTGTACGGAGAGCCTTAA
		ATTCAAGTTCCATGACGAATTCCAGGCCAACGCTAAAAC TGCCAAAAAAG
		GAAAATAG
SEQ ID cycA3		MKII IACGLVAAALFTLTSGQSLAADAPFEGRKCCSSCHKPQAQSWKH TA
NO: 28	(D23_1c17	HA KAMESLKFNVKEAQKAKLDPTKDYTDQKDCVGVCHVDGFGQKGGYTI
(235	86)	DS PKPMLTGVGCESCHGPRGRKYRGDRHKAGQAF EKSGKAPRKT LASKGQ
aa)	protein	DFNFEERCSACHLNYEGSPWKHAKPPYTPYTPPEVDPKYTFKFDEMVKDVK
		AMHEHYKLDGVDPDGEPKFKFHDEFQANAKTAKKKG
SEQ ID cycA3		ATGAAAAATAATAAGCCTGCGGACTGGTTGCTGCAGCCCTGTTACCCCT
NO: 29	(D23_1c17	GACAAGTGGGCAGAGTCTGGCAGCGGATGCTCCGTTTGAAGGTCGGA AAAA
(708	86) DNA	AGTGCAGTTCCTGTCACAAAACCAAGCCCGAGTCGTGGAAACATACTGCC
nt)		CACGCCAAGGCGATGGAATCGCTCAAGCCCAATGTCAAAGTGAAGCCAA
		ACAAAAGGCCAAACTGGATCCTACCAAGGACTACACCCAGGACAAAGACT
		GCGTAGGTTGCCACGTGGATGGATTGGGCCAGAAAGGCGGTACACGATA
		GACTCCCCCAAACCATGCTGACTGGCGTAGGCTGTGAATCCTGCCACGG
		GCCTGGACGTAAATACCGGGGAGATCACCGCAAGGCTGGGCAAGCATTTG
		AGAAATCGGGCAAAAAGCGCCGCGCAAGACCTGGCAAGCAAGGGGCAA
		GACTTTAATTTTGAAGAACGTTGCAGCGCCTGCCATCTGAACATATGAAGG
		GTACCCCTGGAAGGAGCAAAAACCTCCCTATAACCCGTATACACCGGAAG
		TGGATCCGAAATACACCTTCAAGTTTGACGAGATGGTAAAAGACGTCAAA
		GCCATGCACGAGCATTACAAACTGGATGGCGTATTGTACGGAGAGCCTTAA
		ATTCAAGTTCCATGACGAATTCCAGGCCAACGCTAAAAC TGCCAAAAAAG
		GAAAATAG

SUPPLEMENTARY TABLE 2-continued

Selected D23 sequences	
SEQ ID and length	Description Sequence
SEQ ID cM552 NO: 30 cycB1 (239 (D23_1c19 aa) 23) protein	MTRLQKGSIGTLLTGALLGIALVAVVFGGEAALSTEEFCTSCHSMSYPQS ELKESTHYGALGVNPTCKDCHIPQGIENFHLAVATHVVDGARELWLEMVN DYSTLEKFNERRELEMAHDARMNLKKWDSITCRTCHVKPAPPGESAQAEHK KMETEGATCIDCHQNLVHEEAPMTDLNASLAAGKLVLKPEEGDGDGDDDDV DVDDEEEDDEVEVEVEETETADSDSASSSNHDDSDSDE
SEQ ID cM552 NO: 31 cycB1 (720 (D23_1c19 nt) 23) DNA	ATGACTAGACTGCAAAAAGGATCAATTGGTACTTTACTGACAGGAGCTCT GCTGGGAATAGCATTGGTGGCTGTGGTTTTTGGTGGGGAAGCTGCGTTAT CGACCGAAGAGTTTTGTACCAGCTGTCTCATGTGCATACCCACAGAGT GAATTAAGAAGATCCACCCACTATGGTGCATTGGGGTTAATCCGACTTG TAAAGACTGT CATATTCACAAAGGGATAGAAAAATTCACCTGGCAGTAG CAACTCACGTGGTTGATGGTGCCAGAGAACTTTGGTTGGAGATGGTCAAT GACTACTCCACCTGGAGAAGTTCAACGAAAGAAGATTGGAATGGCGCA TGATGCCCGGATGAACCTCAAGAAATGGGACAGCATCACCTGCCGTACCT GTCATGTAAAACAGCTCCTCCGGGAGAAAGCGCCAGGCGGAACATAAG AAAATGGAACGGAAGGAGCAACCTGCATAGACTGT CATCAGAATCTGGT GCATGAAGAAGCGCCGATGACAGATTTGAATGCAAGTCTTGCTGCAGGCA AGCTGGTATTAAAGCCAGAAGAGGGTGACGGTGACGATGACGATGACGTT GACGTTGATGACGAGGAGGAGGATGAAGAAGTTCGAGGTGGAAGTTGAAGA AACTGAAACAGCTGACGACAGCGACTCCGCTTCCTCCAGCAACCATGATG ACGATAGTGATGATGAGTAA
SEQ ID cycB2 NO: 32 (D23_1c25 (239 26) aa) protein	MTRLQKGSIGTLLTGALLGIALVAVVFGGEAALSTEEFCTSCHSMSYPQS ELKESTHYGALGVNPTCKDCHIPQGIENFHLAVATHVVDGARELWLEMVN DYSTLEKFNERRELEMAHDARMNLKKWDSITCRTCHVKPAPPGESAQAEHK KMETEGATCIDCHQNLVHEEAPMTDLNASLAAGKLVLKPEEGDGDGDDDDV DVDDEEEDDEVEVEVEETETADSDSASSSNHDDSDSDE
SEQ ID cycB2 NO: 33 (D23_1c25 (720 26) DNA nt)	ATGACTAGACTGCAAAAAGGATCAATTGGCACTTTACTGACAGGAGCTCT GCTGGGAATAGCATTGGTGGCTGTGGTTTTTGGTGGGGAAGCTGCGTTAT CGACCGAAGAGTTTTGTACCAGCTGTCTCATGTGCATACCCACAGAGT GAATTAAGAAGATCCACCCACTATGGTGCATTGGGGTTAATCCGACTTG TAAAGACTGT CATATTCACAAAGGGATAGAAAAATTCACCTGGCAGTAG CAACTCACGTGGTTGATGGTGCCAGAGAACTTTGGTTGGAGATGGTCAAT GACTACTCCACCTGGAGAAGTTCAACGAAAGAAGATTGGAATGGCGCA TGATGCCCGGATGAACCTCAAGAAATGGGACAGCATCACCTGCCGTACCT GTCATGTAAAACAGCTCCTCCGGGAGAAAGCGCCAGGCGGAACATAAG AAAATGGAACGGAAGGAGCAACCTGCATAGACTGT CATCAGAATCTGGT GCATGAAGAAGCGCCGATGACAGATTTGAATGCAAGTCTTGCTGCAGGCA AGCTGGTATTAAAGCCAGAAGAGGGTGACGATGACGATGACGATGACGTT GACGTTGATGACGAGGAGGAGGATGAAGAAGTTCGAGGTGGAAGTTGAAGA AACTGAAACAGCTGACGACAGCGACTCCGCTTCCTCCAGCAACCATGATG ACGATAGTGATGATGAGTAA

INCORPORATION BY REFERENCE

[0764] All publications and patents mentioned herein are hereby incorporated by reference in their entirety as if each individual publication or patent was specifically and individually indicated to be incorporated by reference.

[0765] While specific embodiments of the subject invention have been discussed, the above specification is illustrative

and not restrictive. Many variations of the invention will become apparent to those skilled in the art upon review of this specification and the claims below. The full scope of the invention should be determined by reference to the claims, along with their full scope of equivalents, and the specification, along with such variations.

[0766] Certain embodiments are within the following claims.

SEQUENCE LISTING

The patent application contains a lengthy "Sequence Listing" section. A copy of the "Sequence Listing" is available in electronic form from the USPTO web site (<http://seqdata.uspto.gov/?pageRequest=docDetail&DocID=US20170037363A1>). An electronic copy of the "Sequence Listing" will also be available from the USPTO upon request and payment of the fee set forth in 37 CFR 1.19(b)(3).

What is claimed is:

1. A purified preparation of optimized *Nitrosomonas eutropha* (*N. eutropha*) bacterium having at least one property selected from:

- an optimized growth rate;
- an optimized NH_4^+ oxidation rate; and
- an optimized resistance to NH_4^+ .

2. The preparation of *N. eutropha* bacterium of claim 1, wherein the optimized growth rate is a rate allowing a continuous culture of *N. eutropha* at an OD600 (optical density at 600 nm) of about 0.15-0.18 to reach an OD600 of about 0.5-0.6 in about 1-2 days.

3. The preparation of *N. eutropha* bacterium of claim 1 or 2, wherein the optimized growth rate is a doubling time of about 8 hours when cultured under batch culture conditions.

4. The preparation of *N. eutropha* bacterium of any of claims 1-3, wherein the optimized NH_4^+ oxidation rate is a rate of at least about 125 micromoles per minute of oxidizing NH_4^+ to NO_2^- .

5. The preparation of *N. eutropha* bacterium of any of claims 1-4, wherein the optimized resistance to NH_4^+ is an ability to grow in medium comprising about 200 mM NH_4^+ for at least about 48 hours.

6. The preparation of *N. eutropha* bacterium of any of claims 1-5, which has at least two properties selected from an optimized growth rate, an optimized NH_4^+ oxidation rate, and an optimized resistance to NH_4^+ .

7. The preparation of *N. eutropha* bacterium of claim 1-6, which has an optimized growth rate, an optimized NH_4^+ oxidation rate, and an optimized resistance to NH_4^+ .

8. The preparation of *N. eutropha* bacterium of any of claims 1-7, which comprises a chromosome that hybridizes under very high stringency to SEQ ID NO: 1.

9. The preparation of *N. eutropha* bacterium of any of claim 1-8, which comprises an AmoA protein having an identity to SEQ ID NO: 6 or 12 selected from at least about 80%, 85%, 90%, 95%, 98%, 99%, 99.5%, and 100% identical, an AmoB protein having an identity to SEQ ID NO: 8 or 14 selected from at least about 80%, 85%, 90%, 95%, 98%, 99%, 99.5%, and 100% identical, an amoC gene having an identity to SEQ ID NO: 4, 10, or 16 selected from at least about 80%, 85%, 90%, 95%, 98%, 99%, 99.5%, and 100% identical, a hydroxylamine oxidoreductase protein having an identity to SEQ ID NO: 18, 20, or 22 selected from at least about 80%, 85%, 90%, 95%, 98%, 99%, 99.5%, and 100% identical, a cytochrome c554 protein having an identity to SEQ ID NO: 24, 26, or 28 selected from at least about 80%, 85%, 90%, 95%, 98%, 99%, 99.5%, and 100% identical, or a cytochrome c_M552 protein having an identity to SEQ ID NO: 30 or 32 selected from at least about 80%, 85%, 90%, 95%, 98%, 99%, 99.5%, and 100% identical.

10. The preparation of *N. eutropha* bacterium of any of claims 1-9, which comprises 1-5, 5-10, 10-15, 15-20, 20-25, 25-30, or all of the sequence characteristics of Table 2.

11. The preparation of *N. eutropha* bacterium of claim 10, which comprises an AmoA1 or AmoA2 protein having (or gene encoding) a mutation relative to *N. eutropha* strain C91 at position 1, e.g., a V at position 1.

12. The preparation of *N. eutropha* bacterium of any of claims 10-11, which comprises an AmoA1 or AmoA2 protein having (or gene encoding) a mutation relative to *N. eutropha* strain C91 at position 160, e.g., an L at position 160.

13. The preparation of *N. eutropha* bacterium of any of claims 10-12, which comprises an AmoA1 or AmoA2 protein having (or gene encoding) a mutation relative to *N. eutropha* strain C91 at position 167, e.g., an A at position 167.

14. The preparation of *N. eutropha* bacterium of any of claims 10-13, which comprises an AmoB1 or AmoB2 protein having (or gene encoding) a mutation relative to *N. eutropha* strain C91 at position 33, e.g., a V at position 33.

15. The preparation of *N. eutropha* bacterium of any of claims 10-14, which comprises an AmoB1 or AmoB2 protein having (or gene encoding) a mutation relative to *N. eutropha* strain C91 at position 165, e.g., an I at position 165.

16. The preparation of *N. eutropha* bacterium of any of claims 10-15, which comprises an AmoC3 protein having (or gene encoding) a mutation relative to *N. eutropha* strain C91 at position 79, e.g., an A at position 79.

17. The preparation of *N. eutropha* bacterium of any of claims 10-16, which comprises an AmoC3 protein having (or gene encoding) a mutation relative to *N. eutropha* strain C91 at position 271, e.g., a V at position 271.

18. The preparation of *N. eutropha* bacterium of any of claims 10-17, which comprises a Hao1, Hao2, or Hao3 protein having (or gene encoding) a mutation relative to *N. eutropha* strain C91 at position 85, e.g., an S at position 85.

19. The preparation of *N. eutropha* bacterium of any of claims 10-18, which comprises a Hao1, Hao2, or Hao3 protein having (or gene encoding) a mutation relative to *N. eutropha* strain C91 at position 312, e.g., an E at position 312.

20. The preparation of *N. eutropha* bacterium of any of claims 10-19, which comprises a Hao1 protein having (or gene encoding) a mutation relative to *N. eutropha* strain C91 at position 163, e.g., an A at position 163.

21. The preparation of *N. eutropha* bacterium of any of claims 10-20, which comprises a c554 CycA1, c554 CycA2, or c554 CycA3 protein having (or gene encoding) a mutation relative to *N. eutropha* strain C91 at position 65, e.g., a T at position 65.

22. The preparation of *N. eutropha* bacterium of any of claims 10-21, which comprises a c554 CycA1 protein having (or gene encoding) a mutation relative to *N. eutropha* strain C91 at position 186, e.g., a T at position 186.

23. The preparation of *N. eutropha* bacterium of any of claims 10-22, which comprises a c_M552 CycB1 or c_M552 CycB2 protein having (or gene encoding) a mutation relative to *N. eutropha* strain C91 at position 63, e.g., a V at position 63.

24. The preparation of *N. eutropha* bacterium of any of claims 10-23, which comprises a c_M552 CycB1 or c_M552 CycB2 protein having (or gene encoding) a mutation relative to *N. eutropha* strain C91 at position 189, e.g., a P at position 189.

25. The preparation of *N. eutropha* bacterium of any of claims 10-24, which comprises a c_M552 CycB1 or c_M552 CycB2 protein having (or gene encoding) a mutation relative to *N. eutropha* strain C91 at position 206, e.g., an insE at position 206.

26. The preparation of *N. eutropha* bacterium of any of claims 10-25, which comprises a c_M552 CycB1 or c_M552 CycB2 protein having (or gene encoding) a mutation relative to *N. eutropha* strain C91 at position 207, e.g., an insE at position 207.

27. The preparation of *N. eutropha* bacterium of any of claims 10-26, which comprises a c_M552 CycB1 protein having (or gene encoding) a mutation relative to *N. eutropha* strain C91 at position 195, e.g., an insD at position 195.

28. The preparation of *N. eutropha* bacterium of any of claims 10-27, which comprises a c_M552 CycB1 protein having (or gene encoding) a mutation relative to *N. eutropha* strain C91 at position 196, e.g., an insD at position 196.

29. The preparation of *N. eutropha* bacterium of any of claims 10-28, which comprises a c_M552 CycB1 protein having (or gene encoding) a mutation relative to *N. eutropha* strain C91 at position 197, e.g., an insD at position 197.

30. The preparation of *N. eutropha* bacterium of any of the preceding claims, which comprises at least one structural difference, e.g., at least one mutation, relative to a wild-type bacterium such as *N. eutropha* strain C91.

31. The preparation of *N. eutropha* bacterium of any of the preceding claims, which comprises a nucleic acid that can be amplified using a pair of primers described herein, e.g., a primer comprising a sequence of SEQ ID NO: 64 and a primer comprising a sequence of SEQ ID NO: 65.

32. The preparation of *N. eutropha* bacterium of any of the preceding claims, which comprises a nucleic acid or protein at least 80%, 85%, 90%, 95%, 98%, 99%, 99.5%, or 100% identical to a gene of FIG. 6 or a protein encoded by a gene of FIG. 6.

33. The preparation of *N. eutropha* bacterium of any of the preceding claims, which comprises a nucleic acid or protein at least 80%, 85%, 90%, 95%, 98%, 99%, 99.5%, or 100% identical to a sequence of any of SEQ ID NOS: 64-66 or a protein encoded by a sequence of any of SEQ ID NOS: 64-66.

34. An *N. eutropha* bacterium, or a purified preparation thereof, comprising a mutation in an ammonia monooxygenase gene, a hydroxylamine oxidoreductase gene, a cytochrome c554 gene, or a cytochrome c_M552 gene relative to a wild-type bacterium such as *N. eutropha* strain C91.

35. The *N. eutropha* bacterium, or a purified preparation thereof, of claim 34, wherein said mutation is in the amoA1 gene.

36. The *N. eutropha* bacterium, or a purified preparation thereof, of claim 34, wherein said mutation is in the amoA2 gene.

37. The *N. eutropha* bacterium, or a purified preparation thereof, of claim 34, wherein said mutation is in the amoB1 gene.

38. The *N. eutropha* bacterium, or a purified preparation thereof, of claim 34, wherein said mutation is in the amoB2 gene.

39. The *N. eutropha* bacterium, or a purified preparation thereof, of claim 34, wherein said mutation is in the amoC3 gene.

40. The *N. eutropha* bacterium, or a purified preparation thereof, of any of claims 34-39, wherein said mutation is at a position described herein, e.g., in Table 2.

41. The *N. eutropha* bacterium, or a purified preparation thereof, of any of claims 34-39, wherein said mutation is a mutation described herein, e.g., in Table 2.

42. The *N. eutropha* bacterium, or a purified preparation thereof, of claim 34, wherein said mutation is in the hao1 gene.

43. The *N. eutropha* bacterium, or a purified preparation thereof, of claim 34, wherein said mutation is in the hao2 gene.

44. The *N. eutropha* bacterium, or a purified preparation thereof, of claim 34, wherein said mutation is in the hao3 gene.

45. The *N. eutropha* bacterium, or a purified preparation thereof, of any of claims 42-44, wherein said mutation is at a position described herein, e.g., in Table 2.

46. The *N. eutropha* bacterium, or a purified preparation thereof, of any of claims 42-44, wherein said mutation is a mutation described herein, e.g., in Table 2.

47. The *N. eutropha* bacterium, or a purified preparation thereof, of claim 34, wherein said mutation is in the c554 cycA1 gene.

48. The *N. eutropha* bacterium, or a purified preparation thereof, of claim 34, wherein said mutation is in the c554 cycA2 gene.

49. The *N. eutropha* bacterium, or a purified preparation thereof, of claim 34, wherein said mutation is in the c554 cycA3 gene.

50. The *N. eutropha* bacterium, or a purified preparation thereof, of any of claims 47-49, wherein said mutation is at a position described herein, e.g., in Table 2.

51. The *N. eutropha* bacterium, or a purified preparation thereof, of any of claims 47-49, wherein said mutation is a mutation described herein, e.g., in Table 2.

52. The *N. eutropha* bacterium, or a purified preparation thereof, of claim 34, wherein said mutation is in the c_M552 cycB1 gene.

53. The *N. eutropha* bacterium, or a purified preparation thereof, of claim 34, wherein said mutation is in the c554 cycB2 gene.

54. The *N. eutropha* bacterium, or a purified preparation thereof, of any of claims 52-53, wherein said mutation is at a position described herein, e.g., in Table 2.

55. The *N. eutropha* bacterium, or a purified preparation thereof, of any of claims 52-53, wherein said mutation is a mutation described herein, e.g., in Table 2.

56. The purified preparation of optimized *N. eutropha* bacterium of claim 1, comprising a mutation in an ammonia monooxygenase gene, a hydroxylamine oxidoreductase gene, a cytochrome c554 gene, or a cytochrome c_M552 gene.

57. The purified preparation of optimized *N. eutropha* bacterium of claim 56, wherein said mutation is in the amoA1 gene.

58. The purified preparation of optimized *N. eutropha* bacterium of claim 56, wherein said mutation is in the amoA2 gene.

59. The purified preparation of optimized *N. eutropha* bacterium of claim 56, wherein said mutation is in the amoB1 gene.

60. The purified preparation of optimized *N. eutropha* bacterium of claim 56, wherein said mutation is in the amoB2 gene.

61. The purified preparation of optimized *N. eutropha* bacterium of claim 56, wherein said mutation is in the amoC3 gene.

62. The purified preparation of optimized *N. eutropha* bacterium of any of claims 57-61, wherein said mutation is at a position described herein, e.g., in Table 2.

63. The purified preparation of optimized *N. eutropha* bacterium of any of claims 57-61, wherein said mutation is a mutation described herein, e.g., in Table 2.

64. The purified preparation of optimized *N. eutropha* bacterium of claim 56, wherein said mutation is in the hao1 gene.

65. The purified preparation of optimized *N. eutropha* bacterium of claim 56, wherein said mutation is in the hao2 gene.

66. The purified preparation of optimized *N. eutropha* bacterium of claim 56, wherein said mutation is in the hao3 gene.

67. The purified preparation of optimized *N. eutropha* bacterium of any of claims 64-66, wherein said mutation is at a position described herein, e.g., in Table 2.

68. The purified preparation of optimized *N. eutropha* bacterium of any of claims 64-66, wherein said mutation is a mutation described herein, e.g., in Table 2.

69. The purified preparation of optimized *N. eutropha* bacterium of claim 56, wherein said mutation is in the c554 cycA1 gene.

70. The purified preparation of optimized *N. eutropha* bacterium of claim 56, wherein said mutation is in the c554 cycA2 gene.

71. The purified preparation of optimized *N. eutropha* bacterium of claim 56, wherein said mutation is in the c554 cycA3 gene.

72. The purified preparation of optimized *N. eutropha* bacterium of any of claims 69-71, wherein said mutation is at a position described herein, e.g., in Table 2.

73. The purified preparation of optimized *N. eutropha* bacterium of any of claims 69-71, wherein said mutation is a mutation described herein, e.g., in Table 2.

74. The purified preparation of optimized *N. eutropha* bacterium of claim 56, wherein said mutation is in the c_M552 cycB1 gene.

75. The purified preparation of optimized *N. eutropha* bacterium of claim 56, wherein said mutation is in the c_M552 cycB2 gene.

76. The purified preparation of optimized *N. eutropha* bacterium of any of claims 56, 74, and 75, wherein said mutation is at a position described herein, e.g., in Table 2.

77. The purified preparation of optimized *N. eutropha* bacterium of any of claims 56, 74, and 75, wherein said mutation is a mutation described herein, e.g., in Table 2.

78. The *N. eutropha* bacterium, or a purified preparation thereof, of claim 34, which has a mutation in at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, or 35 positions of one or more of amoA1 gene, amoA2 gene, amoB1 gene, amoB2 gene, amoC3 gene, hao1 gene, hao2 gene, hao3 gene, c554 cycA1 gene, c554 cycA2 gene, c554 cycA3 gene, c_M552 cycB1 gene, and c554 cycB2 gene.

79. The purified preparation of optimized *N. eutropha* bacterium of claim 56, which has a mutation in at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, or 35 positions of one or more of amoA1 gene, amoA2 gene, amoB1 gene, amoB2 gene, amoC3 gene, hao1 gene, hao2 gene, hao3 gene, c554 cycA1 gene, c554 cycA2 gene, c554 cycA3 gene, c_M552 cycB1 gene, and c554 cycB2 gene.

80. A *N. eutropha* bacterium comprising a chromosome that hybridizes at high stringency to SEQ ID NO: 1.

81. The *N. eutropha* bacterium of claim 80, wherein the chromosome hybridizes at very high stringency to SEQ ID NO: 1.

82. The *N. eutropha* bacterium of claim 80 or 81, which comprises at least one of the genes of FIGS. 6-8, or a gene with at least 80% identity thereto.

83. The *N. eutropha* bacterium of claim 80 or 81, which comprises at least one of the genes of FIGS. 6-8.

84. The *N. eutropha* bacterium of any of claims 80-82, which lacks any plasmid that is at least about 80% identical to SEQ ID NO: 2 or SEQ ID NO: 3.

85. The *N. eutropha* bacterium of claim any of claims 80-84, which lacks any plasmid.

86. A *N. eutropha* bacterium comprising one or more of an AmoA1 gene at least about 98.9% identical to SEQ ID NO: 7 and an amoA2 gene at least about 98.8% identical to SEQ ID NO: 13.

87. A *N. eutropha* bacterium comprising one or more of an AmoA1 protein at least about 99.0% identical to SEQ ID NO: 6 and an AmoA2 protein at least about 99.0% identical to SEQ ID NO: 12.

88. A *N. eutropha* bacterium comprising one or more of an amoB1 gene at least about 99.2% identical to SEQ ID NO: 9 and an amoB2 gene at least about 99.2% identical to SEQ ID NO: 15.

89. The *N. eutropha* bacterium of claim 88, further comprising one or more of an AmoA1 or amoA2 gene at least about 98.9% identical to SEQ ID NO: 7 or 13.

90. An *N. eutropha* bacterium comprising one or more of an AmoB1 protein at least about 99.6% identical to SEQ ID NO: 8 or an AmoB1 protein at least about 99.6% identical to SEQ ID NO: 14.

91. The *N. eutropha* bacterium of claim 90, further comprising one or more of an AmoA1 protein at least about 99.0% identical to SEQ ID NO: 6 and an AmoA2 protein at least about 99.0% identical to SEQ ID NO: 12.

92. An *N. eutropha* bacterium comprising one or more of an amoC1 gene at least about 99.9% identical to SEQ ID NO: 5, an amoC2 gene at least about 99.9% identical to SEQ ID NO: 11, and an amoC3 gene at least about 99.0% identical to SEQ ID NO: 17.

93. The *N. eutropha* bacterium of claim 92, further comprising one or more of an amoA1 gene at least about 98.9% identical to SEQ ID NO: 7, an amo2 gene at least about 98.9% identical to SEQ ID NO: 13, an amoB1 gene at least about 99.2% identical to SEQ ID NO: 9, and an amoB2 gene at least about 99.2% identical to SEQ ID NO: 15.

94. A *N. eutropha* bacterium comprising an AmoC3 protein at least about 99.4% identical to SEQ ID NO: 16.

95. The *N. eutropha* bacterium of claim 94, further comprising one or more of an AmoA1 protein at least about 99.0% identical to SEQ ID NO: 6, an AmoA2 protein at least about 99.0% identical to SEQ ID NO: 12, an AmoB1 protein at least about 99.6% identical to SEQ ID NO: 8, and an AmoB1 protein at least about 99.6% identical to SEQ ID NO: 14.

96. A *N. eutropha* bacterium comprising one or more of a hao1 gene at least about 99.1% identical to SEQ ID NO: 19, a hao2 gene at least about 99.5% identical to SEQ ID NO: 21, and a hao3 gene at least about 99.3% identical to SEQ ID NO: 23.

97. The *N. eutropha* bacterium of claim 96, further comprising one or more of an amoA1 gene at least about 98.9% identical to SEQ ID NO: 7, an amo2 gene at least about 98.9% identical to SEQ ID NO: 13, an amoB1 gene at least about 99.2% identical to SEQ ID NO: 9, an amoB2 gene at least about 99.2% identical to SEQ ID NO: 15, an amoC1 gene at least about 99.9% identical to SEQ ID NO:

5, an amoC2 gene at least about 99.9% identical to SEQ ID NO: 11, and an amoC3 gene at least about 99.0% identical to SEQ ID NO: 17.

98. A *N. eutropha* bacterium comprising one or more of a Hao1 protein at least about 99.6% identical to SEQ ID NO: 18, a Hao2 protein at least about 99.7% identical to SEQ ID NO: 20, and a Hao3 protein at least about 99.7% identical to SEQ ID NO: 22.

99. The *N. eutropha* bacterium of claim **98**, further comprising an AmoA1 protein at least about 99.0% identical to SEQ ID NO: 6, an AmoA2 protein at least about 99.0% identical to SEQ ID NO: 12, an AmoB1 protein at least about 99.6% identical to SEQ ID NO: 8, an AmoB1 protein at least about 99.6% identical to SEQ ID NO: 14, or an AmoC3 protein at least about 99.4% identical to SEQ ID NO: 16.

100. An *N. eutropha* bacterium comprising one or more of a cycA1 gene at least about 98.1% identical to SEQ ID NO: 25, a cycA2 gene at least about 98.8% identical to SEQ ID NO: 27, and a cycA3 gene at least about 99.4% identical to SEQ ID NO: 28.

101. The *N. eutropha* bacterium of claim **100**, further comprising one or more of an amoA1 gene at least about 98.9% identical to SEQ ID NO: 7, an amo2 gene at least about 98.9% identical to SEQ ID NO: 13, an amoB1 gene at least about 99.2% identical to SEQ ID NO: 9, an amoB2 gene at least about 99.2% identical to SEQ ID NO: 15, an amoC1 gene at least about 99.9% identical to SEQ ID NO: 5, an amoC2 gene at least about 99.9% identical to SEQ ID NO: 11, an amoC3 gene at least about 99.0% identical to SEQ ID NO: 17, a hao1 gene at least about 99.1% identical to SEQ ID NO: 19, a hao2 gene at least about 99.5% identical to SEQ ID NO: 21, and a hao3 gene at least about 99.3% identical to SEQ ID NO: 23.

102. An *N. eutropha* bacterium comprising one or more of a CycA1 protein at least about 99.2% identical to SEQ ID NO: 24, a CycA2 protein at least about 99.7% identical to SEQ ID NO: 26, and a CycA3 protein at least about 99.7% identical to SEQ ID NO: 28.

103. The *N. eutropha* bacterium of claim **102**, further comprising one or more of an AmoA1 protein at least about 99.0% identical to SEQ ID NO: 6, an AmoA2 protein at least about 99.0% identical to SEQ ID NO: 12, an AmoB1 protein at least about 99.6% identical to SEQ ID NO: 8, an AmoB1 protein at least about 99.6% identical to SEQ ID NO: 14, an AmoC3 protein at least about 99.4% identical to SEQ ID NO: 16, a Hao1 protein at least about 99.6% identical to SEQ ID NO: 18, a Hao2 protein at least about 99.7% identical to SEQ ID NO: 20, and a Hao3 protein at least about 99.7% identical to SEQ ID NO: 22.

104. A *N. eutropha* bacterium comprising one or more of a cycB1 gene at least about 96.8% identical to SEQ ID NO: 31 and a cycB2 gene at least about 97.2% identical to SEQ ID NO: 33.

105. The *N. eutropha* bacterium of claim **104**, further comprising one or more of an amoA1 gene at least about 98.9% identical to SEQ ID NO: 7, an amo2 gene at least about 98.9% identical to SEQ ID NO: 13, an amoB1 gene at least about 99.2% identical to SEQ ID NO: 9, an amoB2 gene at least about 99.2% identical to SEQ ID NO: 15, an amoC1 gene at least about 99.9% identical to SEQ ID NO: 5, an amoC2 gene at least about 99.9% identical to SEQ ID NO: 11, an amoC3 gene at least about 99.0% identical to SEQ ID NO: 17, a hao1 gene at least about 99.1% identical to SEQ ID NO: 19, a hao2 gene at least about 99.5%

identical to SEQ ID NO: 21, a hao3 gene at least about 99.3% identical to SEQ ID NO: 23, a cycA1 gene at least about 98.1% identical to SEQ ID NO: 25, a cycA2 gene at least about 98.8% identical to SEQ ID NO: 27, and a cycA3 gene at least about 99.4% identical to SEQ ID NO: 28.

106. A *N. eutropha* bacterium comprising one or more of a CycB1 protein at least about 97.2% identical to SEQ ID NO: 30 or a CycB2 protein at least about 98.8% identical to SEQ ID NO: 32.

107. The *N. eutropha* bacterium of claim **106**, further comprising one or more of an AmoA1 protein at least about 99.0% identical to SEQ ID NO: 6, an AmoA2 protein at least about 99.0% identical to SEQ ID NO: 12, an AmoB1 protein at least about 99.6% identical to SEQ ID NO: 8, an AmoB1 protein at least about 99.6% identical to SEQ ID NO: 14, an AmoC3 protein at least about 99.4% identical to SEQ ID NO: 16, a Hao1 protein at least about 99.6% identical to SEQ ID NO: 18, a Hao2 protein at least about 99.7% identical to SEQ ID NO: 20, a Hao3 protein at least about 99.7% identical to SEQ ID NO: 22, a CycA1 protein at least about 99.2% identical to SEQ ID NO: 24, a CycA2 protein at least about 99.7% identical to SEQ ID NO: 26, and a CycA3 protein at least about 99.7% identical to SEQ ID NO: 28.

108. A *N. eutropha* bacterium comprising one or more genes according to SEQ ID NOS: 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, and 33.

109. A *N. eutropha* bacterium comprising one or more proteins according to SEQ ID NOS: 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, and 32.

110. A *N. eutropha* bacterium comprising a protein that is mutant relative to *N. eutropha* strain C91 at at least 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, or all of the amino acid positions listed in Table 2.

111. A *N. eutropha* bacterium comprising proteins that are mutant relative to *N. eutropha* strain C91 at all of the amino acid positions listed in Table 2.

112. The *N. eutropha* bacterium of any one of claims **1-111**, which is transgenic.

113. The *N. eutropha* bacterium of any one of claims **80-112**, having at least one property selected from an optimized growth rate, an optimized NH_4^+ oxidation rate, and an optimized resistance to NH_4^+ .

114. The *N. eutropha* bacterium of claim **113**, which has at least two properties selected from an optimized growth rate, an optimized NH_4^+ oxidation rate, and an optimized resistance to NH_4^+ .

115. The *N. eutropha* bacterium of claim **113**, which has an optimized growth rate, an optimized NH_4^+ oxidation rate, and an optimized resistance to NH_4^+ .

116. A composition comprising the *N. eutropha* bacterium of any one of claims **1-115**, wherein the composition is substantially free of other organisms.

117. A composition comprising the *N. eutropha* bacterium of one any of claims **1-115** and further comprising a second organism, wherein the composition is substantially free of other organisms.

118. The composition of claim **117**, wherein the second organism is an ammonia oxidizing bacterium.

119. The composition of claim **117**, wherein the second organism is selected from the group consisting of *Nitrosomonas*, *Nitrosococcus*, *Nitrosospria*, *Nitrosocystis*, *Nitrosolobus*, *Nitrosovibrio*, *Lactobacillus*, *Streptococcus*, and *Bifidobacter*, and combinations thereof.

120. A composition comprising a cell suspension of an actively dividing culture of *N. eutropha* bacteria having an OD600 of at least about 0.2, wherein the composition is substantially free of other organisms.

121. A composition for topical administration, comprising the *N. eutropha* bacterium of any one of claims 1-115 and a pharmaceutically or cosmetically acceptable excipient suitable for topical administration.

122. The composition of claim 121, which is substantially free of other organisms.

123. The composition of claim 121, further comprising a second organism.

124. The composition of claim 123, wherein the second organism is an ammonia oxidizing bacterium.

125. The composition of claim 123, wherein the second organism is selected from the group consisting of *Nitrosomonas*, *Nitrosococcus*, *Nitrosospria*, *Nitrosocystis*, *Nitrosolobus*, *Nitrosovibrio*, *Lactobacillus*, *Streptococcus*, and *Bifidobacter*, and combinations thereof.

126. The composition of any of claims 121-125, which is provided as or disposed in a powder, cosmetic, cream, stick, aerosol, salve, wipe, or bandage.

127. The composition of any of claims 121-126, which further comprises a moisturizing agent, deodorizing agent, scent, colorant, insect repellent, cleansing agent, or UV-blocking agent.

128. The composition of any of claims 121-127, wherein the excipient comprises an anti-adherent, binder, coat, disintegrant, filler, flavor, color, lubricant, glidant, sorbent, preservative, or sweetener.

129. The composition of any of claims 121-128, wherein the concentration of *N. eutropha* in the composition is about 10^{11} - 10^{12} CFU/L.

130. The composition of any of claims 121-129, wherein the concentration of *N. eutropha* in the composition is about 10^9 CFU/ml.

131. The composition of any of claims 121-130, wherein the mass ratio of *N. eutropha* to pharmaceutical excipient is in a range of about 0.1 grams per liter to about 1 gram per liter.

132. A composition comprising at least about 1,000 L at about 10^{12} CFUs/L of the *N. eutropha* bacterium of any one of claims 1-115.

133. A composition comprising at least about 1, 2, 5, 10, 20, 50, 100, 200, or 500 g of the *N. eutropha* bacterium of any one of claims 1-115, e.g., as a dry formulation such as a powder.

134. An article of clothing comprising the *N. eutropha* of any one of claims 1-115, e.g., at a concentration that provides one or more of a treatment or prevention of a skin disorder, a treatment or prevention of a disease or condition associated with low nitrite levels, a treatment or prevention of body odor, a treatment to supply nitric oxide to a subject, or a treatment to inhibit microbial growth.

135. The article of clothing of claim 134, which is packaged.

136. The article of clothing of claim 134-135, which is packaged in a material that is resistant to gaseous exchange or resistant to water.

137. A cloth comprising the *N. eutropha* of any one of claims 1-115.

138. A yarn comprising the *N. eutropha* of any one of claims 1-115.

139. A thread comprising the *N. eutropha* of any one of claims 1-115.

140. A method of obtaining, e.g., manufacturing, an *N. eutropha* bacterium having an optimized growth rate, an optimized NH_4^+ oxidation rate, or an optimized resistance to NH_4^+ , comprising:

(a) culturing the bacterium under conditions that select for one or more of an optimized growth rate, an optimized NH_4^+ oxidation rate, or an optimized resistance to NH_4^+ , thereby producing a culture;

(b) testing a sample from the culture for an optimized growth rate, an optimized NH_4^+ oxidation rate, or an optimized resistance to NH_4^+ ; and

(c) repeating the culturing and testing steps until a bacterium having an optimized growth rate, an optimized NH_4^+ oxidation rate, or an optimized resistance to NH_4^+ is obtained.

141. The method of claim 140, further comprising a step of obtaining an *N. eutropha* bacterium from a source.

142. The method of claim 141, wherein the source is soil or the skin of an individual.

143. The method of claim 142, wherein culturing the bacterium under conditions that select for one or more of an optimized growth rate, an optimized NH_4^+ oxidation rate, or an optimized resistance to NH_4^+ comprises culturing the bacterium in *N. europae* medium that comprises about 200 mM NH_4^+ .

144. The method of claim 143, which comprises a step of creating an axenic culture.

145. The method of claim 143 or 144, which comprises a step of co-culturing the *N. eutropha* together with at least one other type of ammonia oxidizing bacteria.

146. The method of any of claims 140-145, wherein the *N. eutropha* of step (a) lack an optimized growth rate, an optimized NH_4^+ oxidation rate, and an optimized resistance to NH_4^+ .

147. The method of any of claims 140-146, wherein step (c) comprises repeating the culturing and testing steps until a bacterium having at least two of an optimized growth rate, an optimized NH_4^+ oxidation rate, and an optimized resistance to NH_4^+ is obtained.

148. An *N. eutropha* bacterium produced by the method of any of claims 140-147.

149. A method of testing a preparation of *N. eutropha*, comprising:

assaying the *N. eutropha* for one or more of an optimized growth rate, an optimized NH_4^+ oxidation rate, or an optimized resistance to NH_4^+ ; and

if the *N. eutropha* has one or more of an optimized growth rate, an optimized NH_4^+ oxidation rate, or an optimized resistance to NH_4^+ , classifying the *N. eutropha* as accepted.

150. The method of claim 149, further comprising a step of testing the preparation for contaminating organisms.

151. The method of any of claims 149-150, further comprising a step of removing a sample from the preparation and conducting testing on the sample.

152. The method of any of claims 149-151, further comprising testing medium in which the *N. eutropha* is cultured.

153. The method of any of claims 149-152, further comprising packaging *N. eutropha* from the preparation into a package.

154. The method of any of claims **149-153**, further comprising placing *N. eutropha* from the preparation into commerce.

155. A method of producing, e.g., manufacturing *N. eutropha*, comprising contacting *N. eutropha* with culture medium and culturing the *N. eutropha* until an OD600 of at least about 0.5 is reached.

156. The method of claim **155**, further comprising assaying the *N. eutropha* and culture medium for contaminating organisms.

157. The method of any of claims **155-156**, further comprising assaying the *N. eutropha* for one or more of an optimized growth rate, an optimized NH_4^+ oxidation rate, or an optimized resistance to NH_4^+ .

158. The method of any of claims **155-157**, which comprises producing at least at least about 1,000 L per day at about 10^{12} CFUs/L of *N. eutropha*.

159. A method of producing, e.g., manufacturing, *N. eutropha*, comprising contacting *N. eutropha* with culture medium and culturing the *N. eutropha* until about at least about 1,000 L at about 10^{12} CFUs/L *N. eutropha* are produced.

160. The method of claim **159**, further comprising a step of assaying the *N. eutropha* for one or more of an optimized growth rate, an optimized NH_4^+ oxidation rate, or an optimized resistance to NH_4^+ .

161. The method of any of claims **159-160**, further comprising a step of testing the *N. eutropha* or culture medium for contaminating organisms.

162. The method of any of claims **159-161**, wherein the *N. eutropha* brought into contact with the culture medium is an *N. eutropha* having one or more of an optimized growth rate, an optimized NH_4^+ oxidation rate, or an optimized resistance to NH_4^+ .

163. A method of producing, e.g., manufacturing *N. eutropha*, comprising:

- (a) contacting *N. eutropha* with a culture medium; and
- (b) culturing the *N. eutropha* for 1-2 days, thereby creating a culture, until the culture reaches an OD600 of about 0.5-0.6.

164. The method of claim **163**, further comprising a step of assaying the *N. eutropha* for one or more of an optimized growth rate, an optimized NH_4^+ oxidation rate, or an optimized resistance to NH_4^+ .

165. The method of any of claims **163-164**, further comprising a step of testing the culture for contaminating organisms.

166. The method of any of claims **163-165**, wherein the *N. eutropha* of step (a) is an *N. eutropha* having one or more of an optimized growth rate, an optimized NH_4^+ oxidation rate, or an optimized resistance to NH_4^+ .

167. The method of any of claims **163-166**, which comprises producing at least at least about 1,000 L per day at about 10^{12} CFUs/L of *N. eutropha*.

168. An *N. eutropha* bacterium produced by the method of any of claims **155-167**.

169. A preparation of *N. eutropha* made by the method of any of claims **155-167**.

170. The preparation of claim **169**, wherein the preparation comprises at least about 0.1 to about 100 milligrams (mg) of *N. eutropha*.

171. A reaction mixture comprising *N. eutropha* at an optical density of about 0.5 to about 0.6.

172. A method of producing *N. eutropha*-bearing clothing, comprising contacting an article of clothing with of the *N. eutropha* of any one of claims **1-115**.

173. The method of claim **172**, which comprises producing at least 10, 100, or 1000 articles of clothing.

174. The method of claim **172**, which comprises contacting the article of clothing with at least 10^{10} CFUs of *N. eutropha*.

175. The method of claim **172**, further comprising packaging the clothing.

176. A method of obtaining a formulation of *N. eutropha*, combining contacting *N. eutropha* of any of claims **1-115** with a pharmaceutically or cosmetically acceptable excipient.

177. The method of claim **176**, further comprising mixing the *N. eutropha* and the excipient.

178. The method of claim **176**, which is performed under conditions that are substantially free of contaminating organisms.

179. A method of packaging *N. eutropha*, comprising assembling *N. eutropha* of any of claim **1-115** into a package.

180. The method of claim **179**, wherein the package is resistant to gaseous exchange or resistant to water.

181. The method of claim **179**, wherein the package is permeable to gaseous exchange, NH_3 , NH_4^+ , or NO_2^- .

182. A method of inhibiting microbial growth on a subject's skin, comprising topically administering to a subject in need thereof an effective dose of the *N. eutropha* bacteria of any one of claims **1-115**.

183. The method of claim **182**, wherein the effective dose is approximately 1.5×10^{10} CFU.

184. The method of any of claims **182-183**, wherein the administration is performed twice per day.

185. The method of any of claims **182-184**, wherein the subject is a human.

186. The method of any of claims **182-185**, wherein the microbial growth to be inhibited is growth of *Pseudomonas aeruginosa*, *Staphylococcus aureus*, *Streptococcus pyogenes*, or *Acinetobacter baumannii*.

187. A method of supplying nitric oxide to a subject, comprising positioning an effective dose of the *N. eutropha* bacteria of any one of claims **1-115** in close proximity to the subject.

188. A method of reducing body odor, comprising topically administering to a subject in need thereof an effective dose of the *N. eutropha* bacteria of any one of claims **1-115**.

189. A method of treating a disease associated with low nitrite levels, comprising topically administering to a subject in need thereof a therapeutically effective dose of the *N. eutropha* bacteria of any of claims **1-115**.

190. The method of claim **189**, wherein the disease is HIV dermatitis, infection in a diabetic foot ulcer, atopic dermatitis, acne, e.g., acne vulgaris, eczema, contact dermatitis, allergic reaction, psoriasis, skin infections, vascular disease, vaginal yeast infection, a sexually transmitted disease, heart disease, atherosclerosis, baldness, leg ulcers secondary to diabetes or confinement to bed, angina, particularly chronic, stable angina pectoris, ischemic diseases, congestive heart failure, myocardial infarction, ischemia reperfusion injury, laminitis, hypertension, hypertrophic organ degeneration, Raynaud's phenomenon, fibrosis, fibrotic organ degeneration, allergies, autoimmune sensitization, end stage renal disease, obesity, impotence, or cancer.

191. A method of treating a skin disorder, comprising topically administering to a subject in need thereof a therapeutically effective dose of the *N. eutropha* bacteria of any of claims 1-115.

192. The method of claim 191, wherein the skin disorder is acne, e.g., acne vulgaris, rosacea, eczema, or psoriasis.

193. The method of claim 191, wherein the skin disorder is an ulcer, e.g., venous ulcer, e.g., leg ulcer, e.g., venous leg ulcer, e.g., infection in a diabetic foot ulcer.

194. The method of any one of claims 191-193, wherein topically administering comprises pre-treating the subject with *N. eutropha*, e.g., an *N. eutropha* of any of claims 1-115.

195. The method of any one of claims 191-194, wherein topically administering comprises topically administering prior to occurrence of the skin disorder.

196. The method of any one of claims 191-195, wherein topically administering comprises topically administering subsequent to occurrence of the skin disorder.

197. A method of promoting wound healing or closure, comprising administering to a wound an effective dose of the *N. eutropha* bacteria of any of claims 1-115.

198. The method of claim 197, wherein the wound comprises one or more undesirable bacteria, e.g., pathogenic bacteria.

199. The method of claim 197, wherein the wound comprises *Staphylococcus aureus*, *Pseudomonas aeruginosa*, or *Acinetobacter baumannii*.

200. The method of claim 197, wherein the *N. eutropha* is administered to the subject prior to occurrence of the wound.

201. The method of claim 197, wherein administering to the wound comprises administering to the subject prior to occurrence of the wound.

202. The method of any one of claims 197-201, further comprising administering *N. eutropha* to the wound subsequent to occurrence of the wound.

203. A method of killing or inhibiting growth of pathogenic bacteria comprising contacting, e.g., applying, *N. eutropha* bacteria, e.g., an *N. eutropha* bacterium of any of claims 1-115, to the skin.

204. The method of claim 203, wherein the pathogenic bacteria contribute to one or more of the following conditions: HIV dermatitis, an ulcer, e.g., venous ulcer, e.g., leg ulcer, e.g., venous leg ulcer, e.g., infection in a diabetic foot ulcer, atopic dermatitis, acne, e.g., acne vulgaris, eczema, contact dermatitis, allergic reaction, psoriasis, urticaria, rosacea, skin infections, vascular disease, vaginal yeast infection, a sexually transmitted disease, heart disease, atherosclerosis, baldness, leg ulcers secondary to diabetes or confinement to bed, angina, particularly chronic, stable angina pectoris, ischemic diseases, congestive heart failure, myocardial infarction, ischemia reperfusion injury, laminitis, hypertension, hypertrophic organ degeneration, Raynaud's phenomenon, fibrosis, fibrotic organ degeneration, allergies, autoimmune sensitization, end stage renal disease, obesity, impotence, pneumonia, primary immunodeficiency, epidermal lysis bulosa, or cancer.

205. The method of claim 204, wherein the condition is an ulcer, e.g., venous ulcer, e.g., leg ulcer, e.g., venous leg ulcer, e.g., infection in a diabetic foot ulcer.

206. The method of claim 204, wherein the condition is a venous leg ulcer.

207. The method of claim 204, wherein the condition is acne, e.g., acne vulgaris.

208. The method of claim 204, wherein the condition is acne vulgaris.

209. The method of any one of claims 203-208, wherein the pathogenic bacteria is one or more of *Propionibacterium acnes*, *Pseudomonas aeruginosa*, *Staphylococcus aureus*, *Streptococcus pyogenes*, or *Acinetobacter baumannii*.

210. The method of any one of claims 203-209, further comprising determining whether the subject is in need of killing or inhibiting growth of pathogenic bacteria, e.g., determining that the subject is in need of killing or inhibiting growth of pathogenic bacteria.

211. The method of any one of claims 203-210, further comprising selecting the subject in need of killing or inhibiting growth of pathogenic bacteria.

212. A method of changing a composition of a skin microbiome of a subject comprising:

administering, e.g., applying, a preparation comprising ammonia oxidizing bacteria to a surface of the skin, wherein the amount and frequency of administration, e.g., application, is sufficient to reduce the proportion of pathogenic bacteria on the surface of the skin.

213. The method of claim 212, further comprising, selecting the subject on the basis of the subject being in need of a reduction in the proportion of pathogenic bacteria on the surface of the skin.

214. The method of any one of claims 212-213, wherein the preparation comprises at least one of ammonia, ammonium salts, and urea.

215. The method of any one of claims 212-214, wherein the preparation comprises a controlled release material, e.g., slow release material.

216. The method of any one of claims 212-215, wherein the preparation of ammonia oxidizing bacteria, further comprises an excipient, e.g., one of a pharmaceutically acceptable excipient or a cosmetically acceptable excipient.

217. The method of claim 216, wherein the excipient, e.g., one of the pharmaceutically acceptable excipient and the cosmetically acceptable excipient, is suitable for one of topical, nasal, pulmonary, and gastrointestinal administration.

218. The method of any one of claims 216-217, wherein the excipient, e.g., one of the pharmaceutically acceptable excipient and the cosmetically acceptable excipient, is a surfactant.

219. The method of claim 218, wherein the surfactant is selected from the group consisting of cocamidopropyl betaine (ColaTeric COAB), polyethylene sorbitol ester (e.g., Tween 80), ethoxylated lauryl alcohol (RhodaSurf 6 NAT), sodium laureth sulfate/lauryl glucoside/cocamidopropyl betaine (Plantapon 611 L UP), sodium laureth sulfate (e.g., RhodaPex ESB 70 NAT), alkyl polyglucoside (e.g., Plantaren 2000 N UP), sodium laureth sulfate (Plantaren 200), Dr. Bronner's Castile soap, Lauramine oxide (ColaLux Lo), sodium dodecyl sulfate (SDS), polysulfonate alkyl polyglucoside (PolySufanate 160 P), sodium lauryl sulfate (Stepanol-WA Extra K), and any combination thereof.

220. The method of any one of claims 212-219, wherein the preparation is substantially free of other organisms.

221. The method of any one of claims 212-220, wherein the preparation is disposed in a powder, cosmetic, cream, stick, aerosol, salve, wipe, or bandage.

222. The method of any one of claims **212-221**, wherein the preparation is provided as a powder, cosmetic, cream, stick, aerosol, salve, wipe, or bandage.

223. The method of any one of claims **212-222**, wherein the preparation comprises a moisturizing agent, deodorizing agent, scent, colorant, insect repellent, cleansing agent, or UV-blocking agent.

224. The method of any one of claims **216-217**, wherein the excipient, e.g., the pharmaceutically acceptable excipient or the cosmetically acceptable excipient, comprises an anti-adherent, binder, coat, disintegrant, filler, flavor, color, lubricant, glidant, sorbent, preservative, or sweetener.

225. The method of any one of claims **212-224**, wherein the preparation comprising ammonia oxidizing bacteria comprises about 10^8 to about 10^{14} CFU/L.

226. The method of claim **225**, wherein the preparation comprises between about 1×10^9 CFU/L to about 10×10^9 CFU/L.

227. The method of any one of claims **212-226**, wherein the preparation comprising ammonia oxidizing bacteria comprises between about 50 milligrams (mg) and about 1000 mg of ammonia oxidizing bacteria.

228. The method of any one of claims **216-227**, wherein the mass ratio of ammonia oxidizing bacteria to the excipient, e.g., the pharmaceutically acceptable excipient or the cosmetically acceptable excipient is in a range of about 0.1 grams per liter to about 1 gram per liter.

229. The method of any one of claims **212-228**, wherein the preparation of ammonia oxidizing bacteria are useful for treating or preventing a skin disorder, a treatment or prevention of a disease or condition associated with low nitrite levels, a treatment or prevention of body odor, a treatment to supply nitric oxide to a subject, or a treatment to inhibit microbial growth, e.g., pathogenic bacterial growth.

230. The method of any one of claims **212-229**, wherein the ammonia oxidizing bacteria is selected from the group consisting of *Nitrosomonas*, *Nitrosococcus*, *Nitrosospria*, *Nitrosocystis*, *Nitrosolobus*, *Nitrosovibrio*, and combinations thereof.

231. The method of any one of claims **212-230**, wherein the preparation comprises an organism selected from the group consisting of *Lactobacillus*, *Streptococcus*, *Bifidobacter*, and combinations thereof.

232. The method of any one of claims **212-230**, wherein the preparation is substantially free of organisms other than ammonia oxidizing bacteria.

233. The method of any one of claims **212-232**, wherein the preparation of ammonia oxidizing bacteria comprises ammonia oxidizing bacteria in a growth state.

234. The method of any one of claims **212-232**, wherein the preparation of ammonia oxidizing bacteria comprises ammonia oxidizing bacteria in a storage state.

235. The method of any one of claims **212-234**, to deliver a cosmetic product.

236. The method of any one of claims **212-234**, to deliver a therapeutic product.

237. The method of any one of claims **212-236**, wherein the preparation is useful for treatment of at least one of HIV dermatitis, infection in a diabetic foot ulcer, atopic dermatitis, acne, e.g., acne vulgaris, eczema, contact dermatitis, allergic reaction, psoriasis, urticaria, rosacea, skin infections, vascular disease, vaginal yeast infection, a sexually transmitted disease, heart disease, atherosclerosis, baldness, leg ulcers secondary to diabetes or confinement to bed, angina,

particularly chronic, stable angina pectoris, ischemic diseases, congestive heart failure, myocardial infarction, ischemia reperfusion injury, laminitis, hypertension, hypertrophic organ degeneration, Raynaud's phenomenon, fibrosis, fibrotic organ degeneration, allergies, autoimmune sensitization, end stage renal disease, obesity, impotence, pneumonia, primary immunodeficiency, epidermal lysis bulosa, or cancer.

238. The method of claim **237**, wherein the preparation is useful for treatment of at least one of acne, e.g., acne vulgaris, eczema, psoriasis, urticaria, rosacea, and skin infections.

239. The method of any one of claims **212-238**, wherein the preparation is provided in a container, the preparation and the container having a weight of less than about 50, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1500, or 2000 grams.

240. The method of any one of claims **212-239**, wherein the preparation has less than about 0.1% to about 10% of surfactant.

241. The method of any one of claims **212-240**, wherein the preparation is substantially free of surfactant.

242. The method of any one of claims **212-241**, wherein the preparation comprises a chelator.

243. The method of any one of claims **212-242**, wherein the preparation is applied about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, or 24 times per day.

244. The method of any one of claims **212-243**, wherein the preparation is applied one time per day.

245. The method of any one of claims **212-243**, wherein the preparation is applied two times per day.

246. The method of any one of claims **212-245**, wherein the preparation is applied for about 1-3, 3-5, 5-7, 7-9, 5-10, 10-14, 12-18, 12-21, 21-28, 28-35, 35-42, 42-49, 49-56, 46-63, 63-70, 70-77, 77-84, or 84-91 days.

247. The method of any one of claims **212-246**, wherein the preparation is applied for about 16 days.

248. The method of any one of claims **212-247** further comprising obtaining a sample from the surface of the skin.

249. The method of claim **248**, further comprising isolating DNA of bacteria in the sample.

250. The method of any one of claims **248-249**, further comprising sequencing DNA of bacteria in the sample.

251. The method of any one of claims **212-250**, wherein administering the ammonia oxidizing bacteria provides for an increase in the proportion of non-pathogenic bacteria on the surface.

252. The method of claim **251**, wherein the non-pathogenic bacteria is commensal non-pathogenic bacteria.

253. The method of claim **252**, wherein the non-pathogenic bacteria is commensal non-pathogenic bacteria of the genus *Staphylococcus*.

254. The method of claim **253**, wherein the non-pathogenic bacteria is commensal non-pathogenic bacteria *Staphylococcus epidermidis*.

255. The method of any one of claims **252-254**, wherein the non-pathogenic bacteria of the genus *Staphylococcus* is, or is identified as being, increased after about 2 weeks.

256. The method of claim **255**, wherein the non-pathogenic bacteria *Staphylococcus epidermidis* is, or is identified as being, increased after about 2 weeks.

257. The method of any one of claims **212-256**, wherein potentially pathogenic or disease associated *Propionibacteria* is, or is identified as being, reduced after about 2 weeks.

258. The method of any one of claims **212-257**, wherein potentially pathogenic or disease associated *Stenotrophomonas* is, or is identified as being, reduced after about 2 weeks.

259. The method of any one of claims **212-258**, wherein the surface of the skin comprises a wound.

260. A method of treating acne e.g., acne vulgaris, by the method of any one of claims **212-259**.

261. A method of treating eczema by the method of any one of claims **212-259**.

262. A method of treating psoriasis by the method of any one of claims **212-259**.

263. A method of treating urticaria by the method of any one of claims **212-259**.

264. A method of treating rosacea by the method of any one of claims **212-259**.

265. A method of treating a skin infection by the method of any one of claims **212-259**.

266. A method of reducing an amount of undesirable bacteria on a surface of a subject by the method of any one of claims **212-258**.

267. A nucleic acid comprising a sequence of 15-100 consecutive nucleotides from within SEQ ID NO: 66, or a reverse complement thereof, provided that the nucleic acid has a non-naturally occurring sequence, or another modification, e.g., a label, or both.

268. The nucleic acid of claim **267**, wherein the sequence of 15-100 consecutive nucleotides is a sequence not found in *N. Eutropha* strain C91.

269. The nucleic acid of claim **267** or **268**, further comprising a heterologous sequence 5' to the sequence of 15-100 consecutive nucleotides from within SEQ ID NO: 66.

270. The nucleic acid of claim **267** or **268**, further comprising a heterologous sequence 3' to the sequence of 15-100 consecutive nucleotides from within SEQ ID NO: 66.

271. The nucleic acid of claim **267** or **268**, further comprising a first heterologous sequence 5' to the sequence of 15-100 consecutive nucleotides from within SEQ ID NO: 66 and a second heterologous sequence 3' to the sequence of 15-100 consecutive nucleotides from within SEQ ID NO: 66.

272. The nucleic acid of any of claims **267-271**, which has a length of 15-20, 20-25, 25-30, 30-40, or 25-40 nucleotides.

273. The nucleic acid of any of claims **267-272**, which is bound, e.g., covalently bound, to a detectable label, e.g., a fluorescent label.

274. A composition comprising:

a first nucleic acid comprising 15-100 consecutive nucleotides from within SEQ ID NO: 66; and

a second nucleic acid comprising 15-100 consecutive nucleotides from within a reverse complement of SEQ ID NO: 66,

provided that the first nucleic acid or the second nucleic acid or both has a non-naturally occurring sequence or a modification such as a label, or both.

275. The composition of claim **274**, wherein the first nucleic acid, the second nucleic acid, or each of the first nucleic acid and second nucleic acid does not comprise a sequence found in *N. Eutropha* strain C91.

276. The composition of claim **274** or **275**, wherein the first nucleic acid, the second nucleic acid, or each of the first nucleic acid and second nucleic acid further comprises a heterologous sequence 5' to the sequence of 15-100 consecutive nucleotides from within SEQ ID NO: 66.

277. The composition of any of claims **274-276**, wherein the first nucleic acid, the second nucleic acid, or each of the first nucleic acid and second nucleic acid further comprises a heterologous sequence 3' to the sequence of 15-100 consecutive nucleotides from within SEQ ID NO: 66.

278. The composition of any of claims **274-277**, wherein the first nucleic acid, the second nucleic acid, or each of the first nucleic acid and second nucleic acid has a length of 15-20, 20-25, 25-30, 30-40, or 25-40 nucleotides.

279. The composition of any of claims **274-278**, wherein the first nucleic acid, the second nucleic acid, or each of the first nucleic acid and second nucleic acid is bound, e.g., covalently bound, to a detectable label, e.g., a fluorescent label.

280. A nucleic acid consisting of the sequence AATCTGTCTCCACAGGCAGC (SEQ ID NO: 64).

281. A nucleic acid consisting of the sequence TATACCCACCCACGCTA (SEQ ID NO: 65).

282. A molecule comprising the nucleic acid of claim **280** or **281** and a detectable label, e.g., a fluorescent label.

283. A composition comprising a first nucleic acid consisting of the sequence AATCTGTCTCCACAGGCAGC (SEQ ID NO: 64) and a second nucleic acid consisting of the sequence TATACCCACCCACGCTA (SEQ ID NO: 65).

284. A composition comprising:

a first molecule comprising (i) a first nucleic acid consisting of the sequence AATCTGTCTCCACAGGCAGC (SEQ ID NO: 64) and optionally comprising (ii) a detectable label, e.g., a fluorescent label; and

a second molecule comprising (i) a second nucleic acid consisting of the sequence TATACCCACCCACGCTA (SEQ ID NO: 65) and optionally comprising (ii) a detectable label, e.g., a fluorescent label.

285. A method of detecting whether a D23 *N. eutropha* nucleic acid is present in a sample, comprising:

performing a polymerase chain reaction (PCR) on the sample using primers specific to *N. eutropha* D23, and determining whether a PCR product is produced, wherein the presence of a PCR product indicates that the D23 *N. eutropha* nucleic acid was present in the sample.

* * * * *