



(12) 发明专利

(10) 授权公告号 CN 1756843 B

(45) 授权公告日 2012.03.21

(21) 申请号 200480005628.5

C12R 1/46 (2006.01)

(22) 申请日 2004.03.02

(56) 对比文件

(30) 优先权数据

03450061.1 2003.03.04 EP

W0 02059148 A, 2002.08.01, 全文.

W0 0234771 A2, 2002.05.02, 摘要、实施例
44.

(85) PCT申请进入国家阶段日

2005.08.31

Ferrtti JJ. et al..Complete
genome sequence of an M1strain
of Streptococcus pyogenes. PNAS98
8. 2001, 98(8), 4658-4663.

(86) PCT申请的申请数据

PCT/EP2004/002087 2004.03.02

Smoot JC. et al..Genome sequence and
comparative microarray analysis ofserotype
M18 group A Streptococcus strains
associated withacute rheumatic fever
outbreaks. PNAS99 7. 2002, 99(7), 4668-4673.

(87) PCT申请的公布数据

W02004/078907 EN 2004.09.16

(73) 专利权人 英特塞尔股份公司

地址 奥地利维也纳

审查员 王莉

(72) 发明人 A·迈因克 E·纳吉 B·温克勒

D·格尔布曼

(74) 专利代理机构 中科专利商标代理有限责任
公司 11021

代理人 程金山

(51) Int. Cl.

C12N 15/31 (2006.01)

C07K 14/315 (2006.01)

C07K 16/12 (2006.01)

C12Q 1/68 (2006.01)

A61K 39/09 (2006.01)

G01N 33/68 (2006.01)

权利要求书 2 页 说明书 150 页
序列表 144 页 附图 6 页

(54) 发明名称

化脓链球菌抗原

(57) 摘要

本发明公开了编码超免疫血清反应性抗原或其片段的分离的核酸分子以及来自化脓链球菌的超免疫血清反应性抗原或其片段、分离所述抗原的方法和它们的特殊应用。

1. 一种用于治疗或预防化脓链球菌感染的药物组合物,其包含至少一种抗原并且任选地包含药用载体或赋形剂,所述抗原是一种超免疫血清反应性化脓链球菌 (*S. pyogenes*) 抗原,其与来自具有化脓链球菌感染的人的血清具有免疫反应性,所述抗原包含分离的化脓链球菌肽片段,其中所述肽片段是由 SEQ ID NO. 215 的 31-345 或 247-260 区段的氨基酸组成。

2. 按照权利要求 1 的药物组合物,其另外包含免疫刺激物质。

3. 按照权利要求 2 的药物组合物,其中所述免疫刺激物质是聚阳离子聚合体,免疫刺激性脱氧核苷酸 (ODN),含至少两个 Lys-Leu-Lys 基序的肽,刺激神经组织的化合物,清蛋白,或弗氏完全或不完全佐剂。

4. 按照权利要求 3 的药物组合物,其中所述聚阳离子聚合体是聚阳离子肽。

5. 按照权利要求 1 的药物组合物,其为疫苗。

6. 抗体或其免疫学活性片段,所述抗体对于权利要求 1 中所述的抗原是免疫学特异的,所述免疫学活性片段是 Fab 片段。

7. 按照权利要求 6 的抗体,其为单克隆抗体。

8. 按照权利要求 6 的抗体,其为人抗体或人源化抗体。

9. 按照权利要求 6,7,或 8 的抗体在制备用于诊断动物体内化脓链球菌感染的诊断剂中的应用,所述诊断剂与包括在所述动物的组织或样品中的包含多肽或片段的链球菌抗原反应。

10. 按照权利要求 1 中所述的抗原在制备用于诊断动物体内化脓链球菌感染的诊断剂中的应用,所述诊断剂与在所述动物的组织或样品中的抗体反应,该抗体对于化脓链球菌抗原是免疫学特异的。

11. 一种用于治疗或预防化脓链球菌感染的药物组合物,其包含一种或多种按照权利要求 6,7,或 8 的抗体并且任选地包含药用载体或赋形剂。

12. 按照权利要求 11 的药物组合物在制备用于治疗动物中化脓链球菌感染的药物中的应用。

13. 按照权利要求 5 的疫苗在制备用于对动物免疫以防止化脓链球菌感染的药物中的应用。

14. 按照权利要求 5 的疫苗在制备用于在动物中刺激针对化脓链球菌的免疫应答的药物中的应用。

15. 按照权利要求 13 的应用,其中所述药物另外包括免疫刺激性物质。

16. 权利要求 15 的应用,其中所述免疫刺激物质是聚阳离子聚合体,免疫刺激性脱氧核苷酸 (ODN),含至少两个 Lys-Leu-Lys 基序的肽,刺激神经组织的化合物,清蛋白,或弗氏完全或不完全佐剂。

17. 权利要求 16 的应用,其中所述聚阳离子聚合体是聚阳离子肽。

18. 权利要求 14 的应用,其中所述药物另外包括免疫刺激性物质。

19. 权利要求 18 的应用,其中所述免疫刺激物质是聚阳离子聚合体,免疫刺激性脱氧核苷酸 (ODN),含至少两个 Lys-Leu-Lys 基序的肽,刺激神经组织的化合物,清蛋白,或弗氏完全或不完全佐剂。

20. 权利要求 19 的应用,其中所述聚阳离子聚合体是聚阳离子肽。

21. 权利要求 9 的应用,其中所述动物是人类。
22. 权利要求 10 的应用,其中所述动物是人类。
23. 权利要求 12 的应用,其中所述动物是人类。
24. 权利要求 13 的应用,其中所述动物是人类。
25. 权利要求 14 的应用,其中所述动物是人类。

化脓链球菌抗原

[0001] 本发明涉及分离的核酸分子,它编码关于化脓链球菌 (*Streptococcus pyogenes*) 的抗原,适用于制备药物以预防和治疗由化脓链球菌引起的细菌感染。

[0002] 化脓链球菌,又称为 A 族链球菌 (GAS),是一种重要的革兰氏阳性细胞外细菌病原体且常常感染人。GAS 集群于咽喉或皮肤处造成许多化脓性感染和非化脓性后遗症。它主要是儿童疾病且在人体中引起包括细菌性咽炎、猩红热、脓疱病和脓毒病在内的多种感染。数十年的流行病学研究产生了显著不同的咽喉和皮肤菌株的概念,其中某些血清型通常分别与咽喉或皮肤感染有关 {Cunningham, M., 2000}。已发现 GAS 造成了最近在美国复苏的链球菌中毒休克综合征相关性坏死性筋膜炎 {Cone, L. 等, 1987 ; Stevens, D., 1992} 并被描述为侵入皮肤和软组织导致组织或肢体损伤的“食肉”细菌。

[0003] 感染后人体中可出现几种链球菌后遗症,诸如急性风湿热、急性肾小球肾炎和反应性关节炎。急性风湿热和风湿性心脏病是其中最严重的自身免疫后遗症并已在世界范围内导致了儿童的残疾和死亡。化脓链球菌还可引起诸如猩红热和坏死性筋膜炎等严重的急性疾病并且已与 Tourette 氏综合征以及运动和注意力失调症相关。

[0004] A 族链球菌是咽喉疼痛和咽炎的最常见的细菌诱因且至少占季节性依赖的普通门诊的 16% {Hope-Simpson, R., 1981}。它主要侵袭 5-15 周岁的学龄儿童 {Cunningham, M., 2000}。在例如学校等拥挤条件下所有年龄段都容易受该生物体传播的侵袭。尽管 GAS 不被认为是正常的菌群,但 A 族链球菌的咽运载现象存在时可以无临床症状。

[0005] A 族链球菌可用血清学分类的 Lancefield 分类方案基于其碳水化合物进行区分或基于通过与盐酸一起煮沸细菌可提取的表面蛋白划分成 M 蛋白血清型。这样已产生了 80 个以上的血清型标识,它还可通过分子方法进行分型 (emm 基因)。化脓链球菌的某些 M 蛋白血清型主要与咽炎和风湿热相关,而其它的似乎主要引起了脓皮病和急性肾小球肾炎 {Cunningham, M., 2000}。

[0006] 此外,在引起咽炎及偶尔的中毒性休克中涉及的是 C 型和 G 型链球菌,它必须在咽喉培养后被识别出来 {Hope-Simpson, R., 1981 ; Bisno, A. 等, 1987}。

[0007] 目前,链球菌感染只能用抗生素疗法治疗。不过,用抗生素治疗的案例中有 25-30% 显示出疾病复发和 / 或在粘膜分泌物中流出了所述的生物体。目前尚未有预防疗法 (疫苗) 可供避免链球菌感染。

[0008] 因此,仍然需要寻找一种有效的疗法用于预防或减轻链球菌感染。疫苗不仅能预防链球菌的感染,还可更特异地预防或改善对宿主组织的定植,从而降低咽炎和其它化脓性感染的发生率。诸如风湿热、急性肾小球肾炎、脓毒病、中毒性休克和坏死性筋膜炎等非化脓性后遗症的消除是降低该生物体急性感染和携带发生率的直接后果。显示出能起交叉保护作用以抵抗其它链球菌的疫苗也可用于预防或减轻由所有其它 β -溶血性链球菌引起的感染,即 A、B、C 和 G 族。

[0009] 疫苗可包含全部各种不同抗原。例如,抗原的实例是被完全杀死或减毒的生物体、这些生物体 / 组织的亚组分、蛋白质,或者,它们最简单的形式 - 肽。抗原还可以糖基化的蛋白质或肽的形式被免疫系统所识别,而且还可是或者包含多糖或脂类。可以使用短肽,因

为例如细胞毒性 T 细胞 (CTL) 识别通常形式短至 8-11 个氨基酸长的肽且与主要组织相容性复合体 (MHC) 结合的抗原。B- 细胞可识别 4-5 个氨基酸短的线性表位以及三维结构 (构象表位)。为了获得持续不变的抗原特异性免疫应答,需要用佐剂触发涉及所有必需的免疫系统细胞的免疫级联反应。根本上,佐剂是作用于所谓的抗原呈递细胞 (APCs),但不局限于其作用形式。这些细胞通常首先遇到抗原,然后将加工过或未修饰过的抗原呈递于免疫效应器细胞上。还可涉及中间细胞类型。只有具有适当特异性的效应器细胞可在生产性免疫应答中被激活。佐剂还可局部保留抗原并与其它因子共同注射。此外,佐剂可用作其它免疫细胞的化学引诱物或可局部的和 / 或全身性地用作免疫系统的刺激剂。

[0010] 开发 A 族链球菌疫苗的方法主要集中于化脓链球菌的细胞表面 M 蛋白 {Bessen, D. 等,1988 ;Bronze, M. 等,1988}。由于存在 80 多种不同的化脓链球菌 M 血清型且新的血清型不断出现 {Fischetti, V.,1989},用有限数目的血清型特异的 M 蛋白或 M 蛋白衍生肽接种将不可能起到有效抵抗所有其它 M 血清型的保护作用。此外,已显示 M 蛋白包含一段氨基酸序列,它与人的心脏组织具有免疫学交叉反应性,被认为是风湿热相关性心脏瓣膜损伤的起因 {Fenderson, P. 等,1989}。

[0011] 疫苗开发还可考虑其它蛋白质,诸如红斑毒素、链球菌热原性外毒素 A 和链球菌热原性外毒素 B {Lee, P. K.,1989}。对这些毒素的免疫性可能预防链球菌中毒性休克的致命症状,但可能不预防 A 族链球菌的定植。

[0012] 利用上述蛋白质以及许多另外的候选分子作为抗原用于潜在疫苗 {Ji, Y. 等,1997 ;Guzman, C. 等,1999} 主要是通过基于鉴定的方便性或获得的机会进行选择决定的。需要鉴定针对化脓链球菌的有效和相关抗原。

[0013] 本发明的发明人研发了一种方法用于鉴定、分离和产生来自具体病原体的超免疫血清反应性抗原,特别是来自金黄色葡萄球菌 (*Staphylococcus aureus*) 和表皮葡萄球菌 (*Staphylococcus epidermidis*) (WO 02/059148)。不过,由于生物学特性、致病功能和遗传背景的差异,化脓链球菌与葡萄球菌菌株是显著不同的。重要的是,用于鉴定来自化脓链球菌的抗原的血清选择是与应用于金黄色葡萄球菌筛选的血清选择不同的。为此目的收集三种主要类型的人血清。首先,针对化脓链球菌的鼻咽携带,检测年龄小于 45 周岁且优选家庭中有幼儿的健康成人。大部分幼儿是化脓链球菌的携带者,且就其家庭成员而言他们被认为是一暴露源。基于相关资料,保护性 (集群中和) 抗体有可能存在于非化脓链球菌携带者的被暴露的个体中 (在家庭中具有高携带率的儿童)。为了选择相应的血清来源,用细菌裂解物和培养物上清蛋白质进行检测抗化脓链球菌 IgG 和 IgA 抗体水平的一系列 ELISA。在基于基因组的抗原鉴定中包括来自高效价非携带者的血清。用于人血清选择的这种方法基本上与针对金黄色葡萄球菌的方法非常不同,其中携带和非携带状态不能与抗体水平联系起来。

[0014] 其次,以相同方式表征和选择来自咽炎患者的血清样品。获自患有链球菌后遗症 - 诸如急性风湿热和肾小球肾炎的个体的第三组血清样品主要用于确认的目的。此较后的一组帮助排除表位,它在这些患者中诱发了高水平的抗体,因为链球菌感染后疾病是与 GAS 所诱导的抗体有关的且具有抗诸如心肌等人组织的反应性或者涉及肾小球炎中有害的免疫复合物的形成。两个细菌种属化脓链球菌和金黄色葡萄球菌自身的基因组显示出了许多重要的差异。化脓链球菌含大约 1.85Mb,而金黄色葡萄球菌含 2.85Mb。它们的平均 GC

含量分别为 38.5 和 33% 且大约 30-45% 的编码基因在这两个病原体之间不是共有的。此外,这两个细菌种属繁殖需要不同的生长条件和培养基。而化脓链球菌是严格的人类病原体,金黄色葡萄球菌还被发现可感染一系列温血动物。下文中提供了可由所述的两种病原体造成的一系列重要疾病。金黄色葡萄球菌主要引起医院内的、条件性病原微生物感染:脓疱病、滤泡炎、脓肿、疖子、被感染的伤口、心内膜炎、脑膜炎、脓毒性关节炎、肺炎、骨髓炎、皮肤烫伤综合征(SSS)、中毒性休克综合征。化脓链球菌主要引起群落获得性感染:链球菌咽喉疼痛(发热、渗出性扁桃体炎、咽炎)、链球菌皮肤感染、猩红热、产褥热、败血病、丹毒、肛周蜂窝织炎、乳突炎、中耳炎、肺炎、腹膜炎、伤口感染、急性肾小球肾炎、急性风湿热、中毒性休克样综合征、坏死性筋膜炎。

[0015] 关于本发明的问题是提供用于开发诸如抗化脓链球菌感染的疫苗等药物的方法。更具体而言,此问题是提供可用于所述药物制备的来自化脓链球菌的一套有效、相关和全面的核酸分子或超免疫血清反应性抗原。

[0016] 因此,本发明提供了编码超免疫血清反应性抗原或其片段的分离的核酸分子,该核酸包含选自由下列各项组成的组的核酸序列:

[0017] a) 与选自下列各项的核酸分子具有至少 70% 序列同一性的核酸分子:Seq ID No 1、4-8、10-18、20、22、24-32、34-35、38-40、43-46、49-51、53-54、57-61、63、65-71、73、75-77、81-82、88、91-94 和 96-150,

[0018] b) 与 a) 的核酸分子互补的核酸分子,

[0019] c) 包含 a) 或 b) 的核酸分子的至少 15 个连续碱基的核酸分子,

[0020] d) 在严谨杂交条件下与 a)、b) 或 c) 的核酸分子退火的核酸分子,

[0021] e) 若非遗传密码的简并性,可与 a)、b)、c) 或 d) 中所定义的核酸分子杂交的核酸分子。

[0022] 依照本发明优选的实施方案,序列同一性至少是 80%,优选至少 95%,尤其是 100%。

[0023] 此外,本发明提供了编码超免疫血清反应性抗原或其片段的分离的核酸分子,该核酸包含选自由下列各项组成的组的核酸序列:

[0024] a) 与选自 Seq ID No 64 的核酸分子具有至少 96% 的序列同一性的核酸分子,

[0025] b) 与 a) 的核酸分子互补的核酸分子,

[0026] c) 包含 a) 或 b) 的核酸分子的至少 15 个连续碱基的核酸分子,

[0027] d) 在严谨杂交条件下与 a)、b) 或 c) 的核酸分子退火的核酸分子,

[0028] e) 若非遗传密码的简并性,可与 a)、b)、c) 或 d) 中所定义的核酸杂交的核酸分子。

[0029] 根据另一方面,本发明提供了包含核酸序列的分离的核酸分子,所述核酸序列选自由下列各项组成的组:

[0030] a) 选自 Seq ID No 3、36、47-48、55、62、72、80、84、95 的核酸分子,

[0031] b) 与 a) 的核酸互补的核酸分子,

[0032] c) 若非遗传密码的简并性,可与 a)、b)、c) 或 d) 中所定义的核酸杂交的核酸分子。

[0033] 优选地,核酸分子是 DNA 或 RNA。

[0034] 依照本发明的优选实施方案,核酸分子分离自基因组 DNA,尤其是来自化脓链球菌的基因组 DNA。

[0035] 依照本发明提供了含有按照本发明任一方面的核酸分子的载体。

[0036] 在优选的实施方案中,改造所述载体使其适用于由按照本发明的核酸分子编码的超免疫血清反应性抗原或其片段的重组表达。

[0037] 本发明还提供了含有按照本发明的所述载体的宿主细胞。

[0038] 根据另一方面,本发明进一步提供了所含氨基酸序列由按照本发明的核酸分子编码的超免疫血清反应性抗原。

[0039] 在优选的实施方案中,氨基酸序列(多肽)选自:Seq ID No 151、154-158、160-168、170、172、174-182、184-185、188-190、193-196、199-201、203-204、207-211、213、215-221、223、225-227、231-232、238、241-244 和 246-300。

[0040] 在另一优选的实施方案中,氨基酸序列(多肽)选自 SEq ID No 214。

[0041] 在另一优选的实施方案中,氨基酸序列(多肽)选自 Seq ID No 153、186、197-198、205、212、222、230、234、245。

[0042] 根据另一方面,本发明提供了超免疫血清反应性抗原的片段,其选自包含下列各项的氨基酸的肽:表 1 栏“预测免疫原性氨基酸”和“已鉴定的免疫原性区域的位置”的氨基酸序列、表 2 血清反应性表位,尤其是含以下氨基酸序列的肽:Seq ID No 151 的第 4-44 位、57-65 位、67-98 位、101-107 位、109-125 位、131-144 位、146-159 位、168-173 位、181-186 位、191-200 位、206-213 位、229-245 位、261-269 位、288-301 位、304-317 位、323-328 位、350-361 位、374-384 位、388-407 位、416-425 和 1-114 位;Seq ID No 152 的第 5-17 位、49-64 位、77-82 位、87-98 位、118-125 位、127-140 位、142-150 位、153-159 位、191-207 位、212-218 位、226-270 位、274-287 位、297-306 位、325-331 位、340-347 位、352-369 位、377-382 位、390-395 和 29-226 位;Seq ID No 153 的第 4-16 位、20-26 位、32-74 位、76-87 位、93-108 位、116-141 位、148-162 位、165-180 位、206-219 位、221-228 位、230-236 位、239-245 位、257-268 位、313-328 位、330-335 位、353-359 位、367-375 位、394-403 位、414-434 位、437-444 位、446-453 位、456-464 位、478-487 位、526-535 位、541-552 位、568-575 位、577-584 位、589-598 位、610-618 位、624-643 位、653-665 位、667-681 位、697-718 位、730-748 位、755-761 位、773-794 位、806-821 位、823-831 位、837-845 位、862-877 位、879-889 位、896-919 位、924-930 位、935-940 位、947-955 位、959-964 位、969-986 位、991-1002 位、1012-1036 位、1047-1056 位、1067-1073 位、1079-1085 位、1088-1111 位、1130-1135 位、1148-1164 位、1166-1173 位、1185-1192 位、1244-1254 和 919-929 位;Seq ID No 154 的第 5-44 位、62-74 位、78-83 位、99-105 位、107-113 位、124-134 位、161-174 位、176-194 位、203-211 位、216-237 位、241-247 位、253-266 位、272-299 位、323-349 位、353-360 和 145-305 位;Seq ID No 155 的第 15-39 位、52-61 位、72-81 位、92-97 和 71-81 位;Seq ID No 156 的第 13-19 位、21-31 位、40-108 位、115-122 位、125-140 位、158-180 位、187-203 位、210-223 位、235-245 和 173-186 位;Seq ID No 157 的第 5-12 位、19-27 位、29-39 位、59-67 位、71-78 位、80-88 位、92-104 位、107-124 位、129-142 位、158-168 位、185-191 位、218-226 位、230-243 位、256-267 位、272-277 位、283-291 位、307-325 位、331-344 位、346-352 和 316-331 位;Seq ID No 158 的第 6-28 位、43-53 位、60-76 位、93-103 和 21-99 位;Seq ID No 159 的第 10-30 位、120-126 位、145-151 位、159-169 位、174-182 位、191-196 位、201-206 位、214-220 位、222-232 位、

254-272 位、292-307 位、313-323 位、332-353 位、361-369 位、389-396 位、401-415 位、428-439 位、465-481 位、510-517 位、560-568 和 9-264 位 ;Seq ID No 160 的第 5-29 位、39-45 位、107-128 和 1-112 位 ;Seq ID No 161 的第 4-38 位、42-50 位、54-60 位、65-71 位、91-102 和 21-56 位 ;Seq ID No 162 的第 4-13 位、19-25 位、41-51 位、54-62 位、68-75 位、79-89 位、109-122 位、130-136 位、172-189 位、192-198 位、217-224 位、262-268 位、270-276 位、281-298 位、315-324 位、333-342 位、353-370 位、376-391 和 23-39 位 ;Seq ID No 163 的第 6-41 位、49-58 位、62-103 位、117-124 位、147-166 位、173-194 位、204-211 位、221-229 位、255-261 位、269-284 位、288-310 位、319-325 位、348-380 位、383-389 位、402-410 位、424-443 位、467-479 位、496-517 位、535-553 位、555-565 位、574-581 位、583-591 和 474-489 位 ;Seq ID No 164 的第 8-35 位、52-57 位、66-73 位、81-88 位、108-114 位、125-131 位、160-167 位、174-180 位、230-235 位、237-249 位、254-262 位、278-285 位、308-314 位、321-326 位、344-353 位、358-372 位、376-383 位、393-411 位、439-446 位、453-464 位、471-480 位、485-492 位、502-508 位、523-529 位、533-556 位、558-563 位、567-584 位、589-597 位、605-619 位、625-645 位、647-666 位、671-678 位、690-714 位、721-728 位、741-763 位、766-773 位、777-787 位、792-802 位、809-823 位、849-864 和 37-241 位、409-534 位、582-604 位、743-804 位 ;Seq ID No 165 的第 4-17 位、24-36 位、38-44 位、59-67 位、72-90 位、92-121 位、126-149 位、151-159 位、161-175 位、197-215 位、217-227 位、241-247 位、257-264 位、266-275 位、277-284 位、293-307 位、315-321 位、330-337 位、345-350 位、357-366 位、385-416 和 202-337 位 ;Seq ID No 166 的第 4-20 位、22-46 位、49-70 位、80-89 位、96-103 位、105-119 位、123-129 位、153-160 位、181-223 位、227-233 位、236-243 位、248-255 位、261-269 位、274-279 位、283-299 位、305-313 位、315-332 位、339-344 位、349-362 位、365-373 位、380-388 位、391-397 位、402-407 和 1-48 位 ;Seq ID No 167 的第 18-37 位、41-63 位、100-106 位、109-151 位、153-167 位、170-197 位、199-207 位、212-229 位、232-253 位、273-297 和 203-217 位 ;Seq ID No 168 的第 20-26 位、54-61 位、80-88 位、94-101 位、113-119 位、128-136 位、138-144 位、156-188 位、193-201 位、209-217 位、221-229 位、239-244 位、251-257 位、270-278 位、281-290 位、308-315 位、319-332 位、339-352 位、370-381 位、388-400 位、411-417 位、426-435 位、468-482 位、488-497 位、499-506 位、512-521 和 261-273 位 ;Seq ID No 169 的第 6-12 位、16-36 位、50-56 位、86-92 位、115-125 位、143-152 位、163-172 位、193-203 位、235-244 位、280-289 位、302-315 位、325-348 位、370-379 位、399-405 位、411-417 位、419-429 位、441-449 位、463-472 位、482-490 位、500-516 位、536-543 位、561-569 位、587-594 位、620-636 位、647-653 位、659-664 位、677-685 位、687-693 位、713-719 位、733-740 位、746-754 位、756-779 位、792-799 位、808-817 位、822-828 位、851-865 位、902-908 位、920-938 位、946-952 位、969-976 位、988-1005 位、1018-1027 位、1045-1057 位、1063-1069 位、1071-1078 位、1090-1099 位、1101-1109 位、1113-1127 位、1130-1137 位、1162-1174 位、1211-1221 位、1234-1242 位、1261-1268 位、1278-1284 位、1312-1317 位、1319-1326 位、1345-1353 位、1366-1378 位、1382-1394 位、1396-1413 位、1415-1424 位、1442-1457 位、1467-1474 位、1482-1490 位、1492-1530 位、1537-1549 位、1559-1576 位、1611-1616 位、1624-1641 和 1-414 位、443-614 位、997-1392 位 ;Seq ID No 170 的第 14-42 位、70-75 位、90-100 位、

158-181 和 1-164 位 ;Seq ID No 171 的第 4-21 位、30-36 位、54-82 位、89-97 位、105-118 位、138-147 和 126-207 位 ;Seq ID No 172 的第 4-21 位、31-66 位、96-104 位、106-113 位、131-142 和 180-204 位 ;Seq ID No 173 的第 5-23 位、31-36 位、38-55 位、65-74 位、79-88 位、101-129 位、131-154 位、156-165 位、183-194 位、225-237 位、245-261 位、264-271 位、279-284 位、287-297 位、313-319 位、327-336 位、343-363 位、380-386 和 11-197 位、204-219 位、258-372 位 ;Seq ID No 174 的第 4-20 位、34-41 位、71-86 位、100-110 位、113-124 位、133-143 位、150-158 位、160-166 位、175-182 位、191-197 位、213-223 位、233-239 位、259-278 位、298-322 和 195-289 位 ;Seq ID No 175 的第 4-10 位、21-35 位、44-52 位、54-62 位、67-73 位、87-103 位、106-135 位、161-174 位、177-192 位、200-209 位、216-223 位、249-298 位、304-312 位、315-329 和 12-130 位 ;Seq ID No 176 的第 10-27 位、33-38 位、48-55 位、70-76 位、96-107 位、119-133 位、141-147 位、151-165 位、183-190 位、197-210 位、228-236 位、245-250 位、266-272 位、289-295 位、297-306 位、308-315 位、323-352 位、357-371 位、381-390 位、394-401 位、404-415 位、417-425 位、427-462 位、466-483 位、485-496 位、502-507 位、520-529 位、531-541 位、553-570 位、577-588 位、591-596 位、600-610 位、619-632 位、642-665 位、671-692 位、694-707 和 434-444 位 ;Seq ID No 177 的第 6-14 位、16-25 位、36-46 位、52-70 位、83-111 位、129-138 位、140-149 位、153-166 位、169-181 位、188-206 位、212-220 位、223-259 位、261-269 位、274-282 位、286-293 位、297-306 位、313-319 位、329-341 位、343-359 位、377-390 位、409-415 位、425-430 和 360-375 位 ;Seq ID No 178 的第 4-26 位、28-48 位、54-62 位、88-121 位、147-162 位、164-201 位、203-237 位、245-251 和 254-260 位 ;Seq ID No 179 的第 12-21 位、26-32 位、66-72 位、87-93 位、98-112 位、125-149 位、179-203 位、209-226 位、233-242 位、249-261 位、266-271 位、273-289 位、293-318 位、346-354 位、360-371 位、391-400 和 369-382 位 ;Seq ID No 180 的第 11-38 位、44-65 位、70-87 位、129-135 位、140-163 位、171-177 位、225-232 位、238-249 位、258-266 位、271-280 位、284-291 位、295-300 位、329-337 位、344-352 位、405-412 位、416-424 位、426-434 位、436-455 位、462-475 位、478-487 和 270-312 位 ;Seq ID No 181 的第 5-17 位、34-45 位、59-69 位、82-88 位、117-129 位、137-142 位、158-165 位、180-195 位、201-206 位、219-226 位、241-260 位、269-279 位、292-305 位、312-321 位、341-347 位、362-381 位、396-410 位、413-432 位、434-445 位、447-453 位、482-487 位、492-499 位、507-516 位、546-552 位、556-565 位、587-604 和 486-598 位 ;Seq ID No 182 的第 4-15 位、17-32 位、40-47 位、67-78 位、90-98 位、101-107 位、111-136 位、161-171 位、184-198 位、208-214 位、234-245 位、247-254 位、272-279 位、288-298 位、303-310 位、315-320 位、327-333 位、338-349 位、364-374 和 378-396 位 ;Seq ID No 183 的第 5-27 位、33-49 位、51-57 位、74-81 位、95-107 位、130-137 位、148-157 位、173-184 和 75-235 位 ;Seq ID No 184 的第 6-23 位、47-53 位、57-63 位、75-82 位、97-105 位、113-122 位、124-134 位、142-153 位、159-164 位、169-179 位、181-187 位、192-208 位、215-243 位、247-257 位、285-290 位、303-310 和 30-51 位 ;Seq ID No 185 的第 17-29 位、44-52 位、59-73 位、77-83 位、86-92 位、97-110 位、118-153 位、156-166 位、173-179 位、192-209 位、225-231 位、234-240 位、245-251 位、260-268 位、274-279 位、297-306 位、328-340 位、353-360 位、369-382 位、384-397 位、414-423 位、431-436 位、452-465 位、492-498 位、500-508 位、

516-552 位、554-560 位、568-574 位、580-586 位、609-617 位、620-626 位、641-647 和 208-219 位 ;Seq ID No 186 的第 4-26 位、32-45 位、58-72 位、111-119 位、137-143 位、146-159 位、187-193 位、221-231 位、235-242 位、250-273 位、290-304 位、311-321 位、326-339 位、341-347 位、354-368 位、397-403 位、412-419 位、426-432 位、487-506 位、580-592 位、619-628 位、663-685 位、707-716 位、743-751 位、770-776 位、787-792 位、850-859 位、866-873 位、882-888 位、922-931 位、957-963 位、975-981 位、983-989 位、1000-1008 位、1023-1029 位、1058-1064 位、1089-1099 位、1107-1114 位、1139-1145 位、1147-1156 位、1217-1226 位、1276-1281 位、1329-1335 位、1355-1366 位、1382-1394 位、1410-1416 位、1418-1424 位、1443-1451 位、1461-1469 位、1483-1489 位、1491-1501 位、1515-1522 位、1538-1544 位、1549-1561 位、1587-1593 位、1603-1613 位、1625-1630 位、1636-1641 位、1684-1690 位、1706-1723 位、1765-1771 位、1787-1804 位、1850-1857 位、1863-1894 位、1897-1910 位、1926-1935 位、1937-1943 位、1960-1983 位、1991-2005 位、2008-2014 位、2018-2039 和 396-533 位、1342-1502 位、1672-1920 位 ;Seq ID No 187 的第 4-25 位、45-50 位、53-65 位、79-85 位、87-92 位、99-109 位、126-137 位、141-148 位、156-183 位、190-203 位、212-217 位、221-228 位、235-242 位、247-277 位、287-293 位、300-319 位、321-330 位、341-361 位、378-389 位、394-406 位、437-449 位、455-461 位、472-478 位、482-491 位、507-522 位、544-554 位、576-582 位、587-593 位、611-621 位、626-632 位、649-661 位、679-685 位、696-704 位、706-716 位、726-736 位、740-751 位、759-766 位、786-792 位、797-802 位、810-822 位、824-832 位、843-852 位、863-869 位、874-879 位、882-905 和 1-113 位、210-232 位、250-423 位、536-564 位 ;Seq ID No 188 的第 4-16 位、33-39 位、43-49 位、54-85 位、107-123 位、131-147 位、157-169 位、177-187 位、198-209 位、220-230 位、238-248 位、277-286 位、293-301 位、303-315 位、319-379 位、383-393 位、402-414 位、426-432 位、439-449 位、470-478 位、483-497 位、502-535 位、552-566 位、571-582 位、596-601 位、608-620 位、631-643 位、651-656 位、663-678 位、680-699 位、705-717 位、724-732 位、738-748 位、756-763 位、766-772 位、776-791 位、796-810 位、819-827 位、829-841 位、847-861 位、866-871 位、876-882 位、887-894 位、909-934 位、941-947 位、957-969 位、986-994 位、998-1028 位、1033-1070 位、1073-1080 位、1090-1096 位、1098-1132 位、1134-1159 位、1164-1172 位、1174-1201 和 617-635 位 ;Seq ID No 189 的第 7-25 位、30-40 位、42-64 位、70-77 位、85-118 位、120-166 位、169-199 位、202-213 位、222-244 和 190-203 位 ;SeqID No 190 的第 4-11 位、15-53 位、55-93 位、95-113 位、120-159 位、164-200 位、210-243 位、250-258 位、261-283 位、298-319 位、327-340 位、356-366 位、369-376 位、380-386 位、394-406 位、409-421 位、425-435 位、442-454 位、461-472 位、480-490 位、494-505 位、507-514 位、521-527 位、533-544 位、566-574 和 385-398 位 ;Seq ID No 191 的第 5-36 位、66-72 位、120-127 位、146-152 位、159-168 位、172-184 位、205-210 位、221-232 位、234-243 位、251-275 位、295-305 位、325-332 位、367-373 位、470-479 位、482-487 位、520-548 位、592-600 位、605-615 位、627-642 位、655-662 位、664-698 位、718-725 位、734-763 位、776-784 位、798-809 位、811-842 位、845-852 位、867-872 位、879-888 位、900-928 位、933-940 位、972-977 位、982-1003 和 12-190 位、276-283 位、666-806 位 ;Seq ID No 192 的第 4-38 位、63-68 位、100-114 位、160-173 位、183-192 位、

195-210 位、212-219 位、221-238 位、240-256 位、258-266 位、274-290 位、301-311 位、313-319 位、332-341 位、357-363 位、395-401 位、405-410 位、420-426 位、435-450 位、453-461 位、468-475 位、491-498 位、510-518 位、529-537 位、545-552 位、585-592 位、602-611 位、634-639 位、650-664 和 30-80 位、89-105 位、111-151 位 ;Seq ID No 193 的第 7-29 位、31-39 位、47-54 位、63-74 位、81-94 位、97-117 位、122-127 位、146-157 位、168-192 位、195-204 位、216-240 位、251-259 和 195-203 位 ;SeqID No 194 的第 5-16 位、28-34 位、46-65 位、79-94 位、98-105 位、107-113 位、120-134 位、147-158 位、163-172 位、180-186 位、226-233 位、237-251 位、253-259 位、275-285 位、287-294 位、302-308 位、315-321 位、334-344 位、360-371 位、399-412 位、420-426 和 32-50 位 ;Seq ID No 195 的第 8-20 位、30-36 位、71-79 位、90-96 位、106-117 位、125-138 位、141-147 位、166-174 和 75-90 位 ;Seq ID No 196 的第 4-13 位、15-33 位、43-52 位、63-85 位、98-114 位、131-139 位、146-174 位、186-192 位、198-206 位、227-233 和 69-88 位 ;Seq ID No197 的第 4-22 位、29-35 位、59-68 位、153-170 位、213-219 位、224-238 位、240-246 位、263-270 位、285-292 位、301-321 位、327-346 位、356-371 位、389-405 位、411-418 位、421-427 位、430-437 位、450-467 位、472-477 位、482-487 位、513-518 位、531-538 位、569-576 位、606-614 位、637-657 位、662-667 位、673-690 位、743-753 位、760-767 位、770-777 位、786-802 和 96-230 位、361-491 位、572-585 位 ;SeqID No 198 的第 4-12 位、21-36 位、48-55 位、74-82 位、121-127 位、195-203 位、207-228 位、247-262 位、269-278 位、280-289 和 102-210 位 ;Seq ID No 199 的第 13-20 位、23-31 位、38-44 位、78-107 位、110-118 位、122-144 位、151-164 位、176-182 位、190-198 位、209-216 位、219-243 位、251-256 位、289-304 位、306-313 和 240-248 位 ;Seq ID No200 的第 5-26 位、34-48 位、57-77 位、84-102 位、116-132 位、139-145 位、150-162 位、165-173 位、176-187 位、192-205 位、216-221 位、234-248 位、250-260 和 182-198 位 ;Seq ID No 201 的第 10-19 位、26-44 位、53-62 位、69-87 位、90-96 位、121-127 位、141-146 位、148-158 位、175-193 位、204-259 位、307-313 位、334-348 位、360-365 位、370-401 位、411-439 位、441-450 位、455-462 位、467-472 位、488-504 和 41-56 位 ;Seq ID No 202 的第 5-21 位、36-42 位、96-116 位、123-130 位、138-144 位、146-157 位、184-201 位、213-228 位、252-259 位、277-297 位、308-313 位、318-323 位、327-333 和 202-217 位 ;Seq ID No203 的第 6-26 位、33-51 位、72-90 位、97-131 位、147-154 位、164-171 位、187-216 位、231-236 位、260-269 位、275-283 和 1-127 位 ;Seq IDNo 204 的第 4-22 位、24-38 位、44-58 位、72-88 位、99-108 位、110-117 位、123-129 位、131-137 位、142-147 位、167-178 位、181-190 位、206-214 位、217-223 位、271-282 位、290-305 位、320-327 位、329-336 位、343-352 位、354-364 位、396-402 位、425-434 位、451-456 位、471-477 位、485-491 位、515-541 位、544-583 位、595-609 位、611-626 位、644-656 位、660-681 位、683-691 位、695-718 和 297-458 位 ;SeqID No 205 的第 5-43 位、92-102 位、107-116 位、120-130 位、137-144 位、155-163 位、169-174 位、193-213 和 24-135 位 ;Seq ID No 206 的第 4-25 位、61-69 位、73-85 位、88-95 位、97-109 位、111-130 位、135-147 位、150-157 位、159-179 位、182-201 位、206-212 位、224-248 位、253-260 位、287-295 位、314-331 位、338-344 位、365-376 位、396-405 位、413-422 位、424-430 位、432-449 位、478-485 位、487-494 位、503-517 位、522-536 位、544-560 位、564-578 位、585-590 位、597-613 位、615-623 位、629-636 位、

640-649 位、662-671 位、713-721 和 176-330 位 ;Seq ID No 207 的第 31-37 位、41-52 位、58-79 位、82-105 位、133-179 位、184-193 位、199-205 位、209-226 位、256-277 位、281-295 位、297-314 位、322-328 位、331-337 位、359-367 位、379-395 位、403-409 位、417-432 位、442-447 位、451-460 位、466-472 和 46-62 位、296-341 位 ;Seq ID No 208 的第 23-29 位、56-63 位、67-74 位、96-108 位、122-132 位、139-146 位、152-159 位、167-178 位、189-196 位、214-231 位、247-265 位、274-293 位、301-309 位、326-332 位、356-363 位、378-395 位、406-412 位、436-442 位、445-451 位、465-479 位、487-501 位、528-555 位、567-581 位、583-599 位、610-617 位、622-629 位、638-662 位、681-686 位、694-700 位、711-716 和 667-684 位 ;Seq ID No 209 的第 20-51 位、53-59 位、109-115 位、140-154 位、185-191 位、201-209 位、212-218 位、234-243 位、253-263 位、277-290 位、303-313 位、327-337 位、342-349 位、374-382 位、394-410 位、436-442 位、464-477 位、486-499 位、521-530 位、536-550 位、560-566 位、569-583 位、652-672 位、680-686 位、698-704 位、718-746 位、758-770 位、774-788 位、802-827 位、835-842 位、861-869 和 258-416 位 ;Seq ID No 210 的第 7-25 位、39-45 位、59-70 位、92-108 位、116-127 位、161-168 位、202-211 位、217-227 位、229-239 位、254-262 位、271-278 位、291-300 和 278-295 位 ;Seq ID No 211 的第 4-20 位、27-33 位、45-51 位、53-62 位、66-74 位、81-88 位、98-111 位、124-130 位、136-144 位、156-179 位、183-191 和 183-195 位 ;SeqID No 212 的第 12-24 位、27-33 位、43-49 位、55-71 位、77-85 位、122-131 位、168-177 位、179-203 位、209-214 位、226-241 和 63-238 位 ;Seq ID No 213 的第 4-19 位、37-50 位、120-126 位、131-137 位、139-162 位、177-195 位、200-209 位、211-218 位、233-256 位、260-268 位、271-283 位、288-308 和 1-141 位 ;Seq ID No 214 的第 11-17 位、40-47 位、57-63 位、96-124 位、141-162 位、170-207 位、223-235 位、241-265 位、271-277 位、281-300 位、312-318 位、327-333 位、373-379 和 231-368 位 ;Seq ID No 215 的第 9-33 位、41-48 位、57-79 位、97-103 位、113-138 位、146-157 位、165-186 位、195-201 位、209-215 位、223-229 位、237-247 位、277-286 位、290-297 位、328-342 和 247-260 位 ;Seq ID No 216 的第 7-15 位、39-45 位、58-64 位、79-84 位、97-127 位、130-141 位、163-176 位、195-203 位、216-225 位、235-247 位、254-264 位、271-279 和 64-72 位 ;Seq ID No 217 的第 4-12 位、26-42 位、46-65 位、73-80 位、82-94 位、116-125 位、135-146 位、167-173 位、183-190 位、232-271 位、274-282 位、300-306 位、320-343 位、351-362 位、373-383 位、385-391 位、402-409 位、414-426 位、434-455 位、460-466 位、473-481 位、485-503 位、519-525 位、533-542 位、554-565 位、599-624 位、645-651 位、675-693 位、717-725 位、751-758 位、767-785 位、792-797 位、801-809 位、819-825 位、831-836 位、859-869 位、890-897 和 222-362 位、756-896 位 ;Seq ID No 218 的第 11-17 位、22-28 位、52-69 位、73-83 位、86-97 位、123-148 位、150-164 位、166-177 位、179-186 位、188-199 位、219-225 位、229-243 位、250-255 和 153-170 位 ;Seq ID No 219 的第 4-61 位、71-80 位、83-90 位、92-128 位、133-153 位、167-182 位、184-192 位、198-212 和 56-73 位 ;Seq ID No 220 的第 4-19 位、26-37 位、45-52 位、58-66 位、71-77 位、84-92 位、94-101 位、107-118 位、120-133 位、156-168 位、170-179 位、208-216 位、228-238 位、253-273 位、280-296 位、303-317 位、326-334 和 298-312 位 ;Seq ID No 221 的第 7-13 位、27-35 位、38-56 位、85-108 位、113-121 位、123-160 位、163-169 位、172-183 位、188-200 位、206-211 位、219-238 位、

247-254 和 141-157 位 ;Seq ID No222 的第 23-39 位、45-73 位、86-103 位、107-115 位、125-132 位、137-146 位、148-158 位、160-168 位、172-179 位、185-192 位、200-207 位、210-224 位、233-239 位、246-255 位、285-334 位、338-352 位、355-379 位、383-389 位、408-417 位、423-429 位、446-456 位、460-473 位、478-503 位、522-540 位、553-562 位、568-577 位、596-602 位、620-636 位、640-649 位、655-663 和 433-440 位、572-593 位 ;Seq ID No 223 的第 4-42 位、46-58 位、64-76 位、118-124 位、130-137 位、148-156 位、164-169 位、175-182 位、187-194 位、203-218 位、220-227 位、241-246 位、254-259 位、264-270 位、275-289 位、296-305 位、309-314 位、322-334 位、342-354 位、398-405 位、419-426 位、432-443 位、462-475 位、522-530 位、552-567 位、593-607 位、618-634 位、636-647 位、653-658 位、662-670 位、681-695 位、698-707 位、709-720 位、732-742 位、767-792 位、794-822 位、828-842 位、851-866 位、881-890 位、895-903 位、928-934 位、940-963 位、978-986 位、1003-1025 位、1027-1043 位、1058-1075 位、1080-1087 位、1095-1109 位、1116-1122 位、1133-1138 位、1168-1174 位、1179-1186 位、1207-1214 位、1248-1267 和 17-319 位、417-563 位 ;Seq ID No 224 的第 6-19 位、23-33 位、129-138 位、140-150 位、153-184 位、190-198 位、206-219 位、235-245 位、267-275 位、284-289 位、303-310 位、322-328 位、354-404 位、407-413 位、423-446 位、453-462 位、467-481 位、491-500 和 46-187 位 ;Seq ID No 225 的第 4-34 位、39-57 位、78-86 位、106-116 位、141-151 位、156-162 位、165-172 位、213-237 位、252-260 位、262-268 位、272-279 位、296-307 位、332-338 位、397-403 位、406-416 位、431-446 位、448-453 位、464-470 位、503-515 位、519-525 位、534-540 位、551-563 位、578-593 位、646-668 位、693-699 位、703-719 位、738-744 位、748-759 位、771-777 位、807-813 位、840-847 位、870-876 位、897-903 位、910-925 位、967-976 位、979-992 和 21-244 位、381-499 位、818-959 位 ;Seq ID No 226 的第 19-29 位、65-75 位、90-109 位、111-137 位、155-165 位、169-175 和 118-136 位 ;Seq ID No227 的第 15-20 位、30-36 位、55-63 位、73-79 位、90-117 位、120-127 位、136-149 位、166-188 位、195-203 位、211-223 位、242-255 位、264-269 位、281-287 位、325-330 位、334-341 位、348-366 位、395-408 位、423-429 位、436-444 位、452-465 和 147-155 位 ;Seq ID No 228 的第 11-18 位、21-53 位、77-83 位、91-98 位、109-119 位、142-163 位、173-181 位、193-208 位、216-227 位、238-255 位、261-268 位、274-286 位、290-297 位、308-315 位、326-332 位、352-359 位、377-395 位、399-406 位、418-426 位、428-438 位、442-448 位、458-465 位、473-482 位、488-499 位、514-524 位、543-553 位、564-600 位、623-632 位、647-654 位、660-669 位、672-678 位、710-723 位、739-749 位、787-793 位、820-828 位、838-860 位、889-895 位、901-907 位、924-939 位、956-962 位、969-976 位、991-999 位、1012-1018 位、1024-1029 位、1035-1072 位、1078-1091 位、1142-1161 和 74-438 位 ;Seq ID No229 的第 4-31 位、41-52 位、58-63 位、65-73 位、83-88 位、102-117 位、123-130 位、150-172 位、177-195 位、207-217 位、222-235 位、247-253 位、295-305 位、315-328 位、335-342 位、359-365 位、389-394 位、404-413 和 156-420 位 ;Seq ID No 230 的第 4-42 位、56-69 位、98-108 位、120-125 位、210-216 位、225-231 位、276-285 位、304-310 位、313-318 位、322-343 和 79-348 位 ;Seq ID No 231 的第 12-21 位、24-30 位、42-50 位、61-67 位、69-85 位、90-97 位、110-143 位、155-168 和 53-70 位 ;Seq ID No 232 的第 4-26 位、41-54 位、

71-78 位、88-96 位、116-127 位、140-149 位、151-158 位、161-175 位、190-196 位、201-208 位、220-226 位、240-247 位、266-281 位、298-305 位、308-318 位、321-329 位、344-353 位、370-378 位、384-405 位、418-426 位、429-442 位、457-463 位、494-505 位、514-522 和 183-341 位 ;Seq ID No233 的第 4-27 位、69-77 位、79-101 位、117-123 位、126-142 位、155-161 位、171-186 位、200-206 位、213-231 位、233-244 位、258-263 位、269-275 位、315-331 位、337-346 位、349-372 位、376-381 位、401-410 位、424-445 位、447-455 位、463-470 位、478-484 位、520-536 位、546-555 位、558-569 位、580-597 位、603-618 位、628-638 位、648-660 位、668-683 位、717-723 位、765-771 位、781-788 位、792-806 位、812-822 和 92-231 位、618-757 位 ;Seq ID No 234 的第 11-47 位、63-75 位、108-117 位、119-128 位、133-143 位、171-185 位、190-196 位、226-232 位、257-264 位、278-283 位、297-309 位、332-338 位、341-346 位、351-358 位、362-372 和 41-170 位 ;Seq ID No 235 的第 6-26 位、50-56 位、83-89 位、108-114 位、123-131 位、172-181 位、194-200 位、221-238 位、241-259 位、263-271 位、284-292 位、304-319 位、321-335 位、353-358 位、384-391 位、408-417 位、424-430 位、442-448 位、459-466 位、487-500 位、514-528 位、541-556 位、572-578 位、595-601 位、605-613 位、620-631 位、634-648 位、660-679 位、686-693 位、702-708 位、716-725 位、730-735 位、749-755 位、770-777 位、805-811 位、831-837 位、843-851 位、854-860 位、863-869 位、895-901 位、904-914 位、922-929 位、933-938 位、947-952 位、956-963 位、1000-1005 位、1008-1014 位、1021-1030 位、1131-1137 位、1154-1164 位、1166-1174 和 20-487 位、757-1153 位 ;Seq ID No 236 的第 10-34 位、67-78 位、131-146 位、160-175 位、189-194 位、201-214 位、239-250 位、265-271 位、296-305 和 26-74 位、91-100 位、105-303 位 ;Seq ID No 237 的第 9-15 位、19-32 位、109-122 位、143-150 位、171-180 位、186-191 位、209-217 位、223-229 位、260-273 位、302-315 位、340-346 位、353-359 位、377-383 位、389-406 位、420-426 位、460-480 和 10-223 位、231-251 位、264-297 位、312-336 位 ;Seq ID No 238 的第 5-28 位、76-81 位、180-195 位、203-209 位、211-219 位、227-234 位、242-252 位、271-282 位、317-325 位、350-356 位、358-364 位、394-400 位、405-413 位、417-424 位、430-436 位、443-449 位、462-482 位、488-498 位、503-509 位、525-537 和 22-344 位 ;Seq ID No 239 的第 5-28 位、42-54 位、77-83 位、86-93 位、98-104 位、120-127 位、145-159 位、166-176 位、181-187 位、189-197 位、213-218 位、230-237 位、263-271 位、285-291 位、299-305 位、326-346 位、368-375 位、390-395 和 1-151 位 ;Seq ID No 240 的第 6-34 位、48-55 位、58-64 位、84-101 位、121-127 位、143-149 位、153-159 位、163-170 位、173-181 位、216-225 位、227-240 位、248-254 位、275-290 位、349-364 位、375-410 位、412-418 位、432-438 位、445-451 位、465-475 位、488-496 位、505-515 位、558-564 位、571-579 位、585-595 位、604-613 位、626-643 位、652-659 位、677-686 位、688-696 位、702-709 位、731-747 位、777-795 位、820-828 位、836-842 位、845-856 位、863-868 位、874-882 位、900-909 位、926-943 位、961-976 位、980-986 位、992-998 位、1022-1034 位、1044-1074 位、1085-1096 位、1101-1112 位、1117-1123 位、1130-1147 位、1181-1187 位、1204-1211 位、1213-1223 位、1226-1239 位、1242-1249 位、1265-1271 位、1273-1293 位、1300-1308 位、1361-1367 位、1378-1384 位、1395-1406 位、1420-1428 位、1439-1446 位、1454-1460 位、1477-1487 位、1509-1520 位、1526-1536 位、

1557-1574 位、1585-1596 位、1605-1617 位、1621-1627 位、1631-1637 位、1648-1654 位、1675-1689 位、1692-1698 位、1700-1706 位、1712-1719 位、1743-1756 和 91-263 位 ;Seq ID No 241 的第 4-16 位、75-90 位、101-136 位、138-144 位、158-164 位、171-177 位、191-201 位、214-222 位、231-241 位、284-290 位、297-305 位、311-321 位、330-339 位、352-369 位、378-385 位、403-412 位、414-422 位、428-435 位、457-473 位、503-521 位、546-554 位、562-568 位、571-582 位、589-594 位、600-608 位、626-635 位、652-669 位、687-702 位、706-712 位、718-724 位、748-760 位、770-775 和 261-272 位 ;Seq ID No 242 的第 4-19 位、30-41 位、46-57 位、62-68 位、75-92 位、126-132 位、149-156 位、158-168 位、171-184 位、187-194 位、210-216 位、218-238 位、245-253 位、306-312 位、323-329 位、340-351 位、365-373 位、384-391 位、399-405 位、422-432 位、454-465 位、471-481 位、502-519 位、530-541 位、550-562 位、566-572 位、576-582 位、593-599 位、620-634 位、637-643 位、645-651 位、657-664 位、688-701 和 541-551 位 ;Seq ID No 243 的第 6-11 位、17-25 位、53-58 位、80-86 位、91-99 位、101-113 位、123-131 位、162-169 位、181-188 位、199-231 位、245-252 和 84-254 位 ;Seq ID No 244 的第 13-30 位、71-120 位、125-137 位、139-145 位、184-199 和 61-78 位 ;Seq ID No 245 的第 9-30 位、38-53 位、63-70 位、74-97 位、103-150 位、158-175 位、183-217 位、225-253 位、260-268 位、272-286 位、290-341 位、352-428 位、434-450 位、453-460 位、469-478 位、513-525 位、527-534 位、554-563 位、586-600 位、602-610 位、624-640 位、656-684 位、707-729 位、735-749 位、757-763 位、766-772 位、779-788 位、799-805 位、807-815 位、819-826 位、831-855 和 568-580 位 ;Seq ID No 246 的第 11-21 位、29-38 和 5-17 位 ;Seq ID No 247 的第 2-9 位 ;Seq ID No 248 的第 4-10 位、16-28 和 7-18 位、26-34 位 ;Seq ID No 249 的第 10-16 和 1-15 位 ;Seq ID No 250 的第 4-11 位 ;Seq ID No 251 的第 4-40 位、42-51 和 37-53 位 ;Seq ID No 252 的第 4-21 和 22-29 位 ;Seq ID No 253 的第 2-11 位 ;9-17 位、Seq ID No 254 的第 32-44 和 1-22 位 ;Seq ID No 255 的第 19-25 位、27-32 和 15-34 位 ;Seq ID No 256 的第 4-12 位、15-22 和 11-33 位 ;Seq ID No 257 的第 10-17 位、24-30 位、39-46 位、51-70 和 51-61 位 ;Seq ID No 258 的第 6-19 位 ;Seq ID No 259 的第 6-11 位、21-27 位、31-54 和 11-29 位 ;Seq ID No 260 的第 4-10 位、13-45 和 11-35 位 ;Seq ID No 261 的第 4-14 位、23-32 和 11-35 位 ;Seq ID No 262 的第 14-39 位、45-51 和 15-29 位 ;Seq ID No 263 的第 4-11 位、14-28 和 4-17 位 ;Seq ID No 264 的第 4-16 和 2-16 位 ;Seq ID No 265 的第 4-10 位、12-19 位、39-50 和 6-22 位 ;Seq ID No 266 的第 2-13 位 ;Seq ID No 267 的第 4-11 位、22-65 和 3-19 位 ;Seq ID No 268 的第 17-23 位、30-35 位、39-46 位、57-62 和 30-49 位 ;Seq ID No 269 的第 4-19 和 14-22 位 ;Seq ID No 270 的第 2-9 位 ;Seq ID No 271 的第 7-18 位、30-43 和 4-12 位 ;Seq ID No 272 的第 4-30 位、39-47 和 5-22 位 ;Seq ID No 273 的第 6-15 和 14-29 位 ;Seq ID No 274 的第 4-34 和 23-35 位 ;Seq ID No 275 的第 4-36 位、44-57 位、65-72 和 14-27 位 ;Seq ID No 276 的第 4-18 和 11-20 位 ;Seq ID No 277 的第 5-19 位 ;Seq ID No 278 的第 18-36 和 6-20 位 ;Seq ID No 279 的第 4-10 位、19-34 位、41-84 位、96-104 和 50-63 位 ;Seq ID No 280 的第 4-9 位、19-27 和 8-21 位 ;Seq ID No 281 的第 4-16 位、18-28 和 22-30 位 ;Seq ID No 282 的第 4-15 和 21-35 位 ;Seq ID No 283 的第 4-17 和 3-13 位 ;Seq ID No 284 的第 4-12 和 4-18 位 ;Seq ID No 285 的第 4-24 位、31-36 和 29-45 位 ;Seq ID No 286 的第 12-22 位、

34-49 和 21-32 位 ;Seq ID No 287 的第 4-17 和 22-32 位 ;Seq ID No 288 的第 4-16 位、25-42 和 7-28 位 ;Seq ID No 289 的第 4-10 和 7-20 位 ;Seq ID No 290 的第 4-11 位、16-36 位、39-54 和 28-44 位 ;Seq ID No 291 的第 5-20 位、29-54 和 14-29 位 ;Seq ID No 292 的第 24-33 和 10-22 位 ;Seq ID No 293 的第 10-51 位、54-61 和 43-64 位 ;Seq ID No 294 的第 7-13 和 2-17 位 ;Seq ID No 295 的第 11-20 和 6-20 位 ;Seq ID No 296 的第 4-30 位、34-41 和 19-28 位 ;Seq ID No 297 的第 11-21 位 ;Seq ID No 298 的第 4-16 位、21-26 和 9-38 位 ;Seq ID No 299 的第 4-12 位、15-27 位、30-42 位、66-72 和 10-24 位 ;Seq ID No 300 的第 8-17 和 11-20 位 ;以及 Seq ID No 246 的第 2-19 位 ;Seq ID No 247 的第 1-12 位 ;Seq ID No 248 的第 21-38 位 ;Seq ID No 254 的第 2-22 位 ;Seq ID No 255 的第 15-33 位 ;Seq ID No 256 的第 11-32 位 ;Seq ID No 259 的第 11-28 位 ;Seq ID No 260 的第 10-27 位 ;Seq ID No 261 的第 9-26 位 ;Seq ID No 263 的第 4-16 位 ;Seq ID No 266 的第 1-18 位 ;Seq ID No 273 的第 12-29 位 ;Seq ID No 276 的第 6-23 位 ;Seq ID No 277 的第 1-21 位 ;Seq ID No 279 的第 47-64 位 ;Seq ID No 285 的第 28-45 位 ;Seq ID No 287 的第 18-35 位 ;Seq ID No 291 的第 14-31 位 ;Seq ID No 292 的第 7-24 位 ;Seq ID No 299 的第 8-25 位 ;Seq ID No 300 的第 1-20 位 ;Seq ID No 151 的第 18-33 位 ;Seq ID No 151 的第 62-72 位 ;Seq ID No 152 的第 118-131 位 ;Seq ID No 154 的第 195-220 位 ;Seq ID No 154 的第 215-240 位 ;Seq ID No 154 的第 255-280 位 ;Seq ID No 155 的第 72-81 位 ;Seq ID No 156 的第 174-186 位 ;Seq ID No 157 的第 317-331 位 ;Seq ID No 158 的第 35-59 位 ;Seq ID No 158 的第 54-84 位 ;Seq ID No 158 的第 79-104 位 ;Seq ID No 159 的第 33-58 位 ;Seq ID No 159 的第 81-101 位 ;Seq ID No 159 的第 136-150 位 ;Seq ID No 159 的第 173-186 位 ;Seq ID No 159 的第 231-251 位 ;Seq ID No 161 的第 22-48 位 ;Seq ID No 162 的第 24-39 位 ;Seq ID No 163 的第 475-489 位 ;Seq ID No 164 的第 38-56 位 ;Seq ID No 164 的第 583-604 位 ;Seq ID No 165 的第 202-223 位 ;Seq ID No 165 的第 222-247 位 ;Seq ID No 165 的第 242-267 位 ;Seq ID No 165 的第 262-287 位 ;Seq ID No 165 的第 282-307 位 ;Seq ID No 165 的第 302-327 位 ;Seq ID No 166 的第 25-48 位 ;Seq ID No 167 的第 204-217 位 ;Seq ID No 168 的第 259-276 位 ;Seq ID No 169 的第 121-139 位 ;Seq ID No 169 的第 260-267 位 ;Seq ID No 169 的第 215-240 位 ;Seq ID No 170 的第 115-140 位 ;Seq ID No 172 的第 182-204 位 ;Seq ID No 173 的第 144-153 位 ;Seq ID No 173 的第 205-219 位 ;Seq ID No 174 的第 196-206 位 ;Seq ID No 174 的第 240-249 位 ;Seq ID No 174 的第 272-287 位 ;Seq ID No 174 的第 199-223 位 ;Seq ID No 174 的第 218-237 位 ;Seq ID No 175 的第 226-249 位 ;Seq ID No 175 的第 287-306 位 ;Seq ID No 176 的第 430-449 位 ;Seq ID No 177 的第 61-375 位 ;Seq ID No 178 的第 241-260 位 ;Seq ID No 181 的第 483-502 位 ;Seq ID No 182 的第 379-396 位 ;Seq ID No 184 的第 31-51 位 ;Seq ID No 186 的第 1436-1460 位 ;Seq ID No 186 的第 1455-1474 位 ;Seq ID No 186 的第 1469-1487 位 ;Seq ID No 187 的第 215-229 位 ;Seq ID No 187 的第 534-561 位 ;Seq ID No 187 的第 59-84 位 ;Seq ID No 187 的第 79-104 位 ;Seq ID No 188 的第 618-635 位 ;Seq ID No 189 的第 191-203 位 ;Seq ID No 190 的第 386-398 位 ;Seq ID No 191 的第 65-83 位 ;Seq ID No 192 的第 90-105 位 ;Seq ID No 192 的第 112-136 位 ;Seq ID No 193 的第 290-209 位 ;Seq ID No 194 的第 33-50 位 ;Seq ID No 195 的第 76-90 位 ;Seq ID No 196 的第 70-88 位 ;Seq

ID No 197 的第 418-442 位 ;Seq ID No 197 的第 574-585 位 ;Seq ID No 198 的第 87-104 位 ;Seq ID No 198 的第 124-148 位 ;Seq ID No 198 的第 141-152 位 ;Seq ID No 199 的第 241-248 位 ;Seq ID No 200 的第 183-198 位 ;Seq ID No 200 的第 40-57 位 ;Seq ID No 202 的第 202-217 位 ;Seq ID No 203 的第 50-74 位 ;Seq ID No 203 的第 69-93Seq ID No 203 的第 88-112 位 ;Seq ID No 的第 203107-127 位 ;Seq ID No 205 的第 74-92 位 ;Seq ID No 206 的第 207-232 位 ;Seq ID No 206 的第 227-252 位 ;SeqID No 206 的第 247-272 位 ;Seq ID No 207 的第 47-60 位 ;Seq ID No 207 的第 297-305 位 ;Seq ID No 207 的第 312-337 位 ;Seq ID No 208 的第 667-384 位 ;Seq ID No 210 的第 279-295 位 ;Seq ID No 211 的第 179-198 位 ;Seq ID No 213 的第 27-51 位 ;Seq ID No 213 的第 46-70 位 ;Seq ID No 213 的第 65-89 位 ;Seq ID No 213 的第 84-108 位 ;Seq ID No 213 的第 112-141 位 ;Seq ID No 215 的第 248-260 位 ;Seq ID No 216 的第 59-78 位 ;Seq ID No 218 的第 154-170 位 ;Seq ID No 219 的第 57-73 位 ;Seq ID No 220 的第 297-314 位 ;Seq ID No 221 的第 142-157 位 ;Seq ID No 222 的第 428-447 位 ;Seq ID No 222 的第 573-593 位 ;Seq ID No 223 的第 523-544 位 ;Seq ID No 223 的第 46-70 位 ;Seq ID No 223 的第 65-89 位 ;Seq ID No 223 的第 84-108 位 ;Seq ID No 223 的第 122-151 位 ;Seq ID No 224 的第 123-142 位 ;Seq ID No 225 的第 903-921 位 ;Seq ID No 226 的第 119-136 位 ;Seq ID No 227 的第 142-161 位 ;Seq ID No 228 的第 258-277 位 ;SeqID No 228 的第 272-300 位 ;Seq ID No 228 的第 295-322 位 ;Seq ID No 229 的第 311-343 位 ;Seq ID No 229 的第 278-304 位 ;Seq ID No 230 的第 131-150 位 ;Seq ID No 230 的第 195-218 位 ;Seq ID No 231 的第 53-70 位 ;Seq ID No 232 的第 184-208 位 ;Seq ID No 232 的第 222-246 位 ;Seq ID No 232 的第 241-265 位 ;Seq ID No 232 的第 260-284 位 ;Seq ID No 232 的第 279-303 位 ;Seq ID No 232 的第 317-341 位 ;Seq ID No 233 的第 678-696 位 ;Seq ID No 235 的第 88-114 位 ;Seq ID No 235 的第 464-481 位 ;Seq ID No 236 的第 153-172 位 ;Seq ID No 236 的第 137-155 位、166-184 位 ;Seq ID No 236 的第 215-228 位 ;Seq ID No 237 的第 37-51 位 ;Seq ID No 237 的第 53-75 位 ;Seq ID No 237 的第 232-251 位 ;Seq ID No 237 的第 318-336 位 ;Seq ID No 238 的第 305-315 位 ;Seq ID No 238 的第 131-156 位 ;SeqID No 241 的第 258-275 位 ;Seq ID No 243 的第 107-137 位 ;Seq ID No 243 的第 138-162 位 ;Seq ID No 243 的第 157-181 位 ;Seq ID No 243 的第 195-227 位 ;Seq ID No 244 的第 62-78 位 ;Seq ID No 245 的第 567-584 位。

[0043] 本发明还提供了产生按照本发明的化脓链球菌超免疫血清反应性抗原或其片段的方法,包括在适当的表达系统中表达按照本发明的一个或多个核酸分子。

[0044] 此外,本发明提供了用于产生表达按照本发明的化脓链球菌超免疫血清反应性抗原或其片段的细胞的方法,包括用按照本发明的载体转化或转染适当的宿主细胞。

[0045] 依照本发明提供了含有本发明中所定义的超免疫血清反应性抗原或其片段或本发明中所定义的核酸分子的药物组合物,尤其是疫苗。

[0046] 在优选的实施方案中,药物组合物进一步包含免疫刺激物质,优选选自包含下列各项的组:聚阳离子聚合体,尤其是聚阳离子肽;免疫刺激性脱氧核苷酸 (ODNs);含至少两个赖氨酸-亮氨酸-赖氨酸基序的肽,尤其是 KLKL5KLK;刺激神经组织的化合物,尤其是人生长激素;清蛋白,弗氏完全或不完全佐剂或它们的组合。

[0047] 在更优选的实施方案中,免疫刺激物质是聚阳离子聚合体与免疫刺激性脱氧核苷酸的组合或含至少两个赖氨酸-亮氨酸-赖氨酸基序的肽和免疫刺激性脱氧核苷酸的组合。

[0048] 在还更优选的实施方案中,聚阳离子聚合体是聚阳离子肽,尤其是多精氨酸。

[0049] 依照本发明提供了按照本发明的核酸分子或按照本发明的超免疫血清反应性抗原或其片段在制备药物制剂,尤其是制备抗化脓链球菌感染的疫苗中的应用。

[0050] 同此还提供了至少与按照本发明的超免疫血清反应性抗原或其片段的选择性部分结合的抗体或至少其有效部分。

[0051] 在优选的实施方案中,所述抗体是单克隆抗体。

[0052] 在另一优选的实施方案中,所述抗体的有效部分包含 Fab 片段。

[0053] 在另一优选的实施方案中,所述抗体是嵌合抗体。

[0054] 在更优选的实施方案中,所述抗体是人源化抗体。

[0055] 本发明还提供了生产按照本发明的抗体的杂交瘤细胞系。

[0056] 此外,本发明提供了生产按照本发明的抗体的方法,其特征在于以下步骤:

[0057] • 按照本发明中所定义的,通过对所述动物施用超免疫血清反应性抗原或其片段在非人动物体内引发免疫应答,

[0058] • 从所述动物中取出含抗体的体液,并

[0059] • 通过将所述含抗体的体液进行进一步的纯化步骤生产该抗体。

[0060] 此外,本发明还提供了生产按照本发明的抗体的方法,其特征在于以下步骤:

[0061] • 按照本发明中所定义的,通过对所述动物施用超免疫血清反应性抗原或其片段在非人动物体内引发免疫应答,

[0062] • 从所述动物体内取出脾脏或脾细胞,

[0063] • 产生所述脾脏或脾细胞的杂交瘤细胞,

[0064] • 选择并克隆对所述超免疫血清反应性抗原或其片段特异的杂交瘤细胞,

[0065] • 通过培养所述被克隆的杂交瘤细胞和任选进一步的纯化步骤生产抗体。

[0066] 依照以上方法提供或生产的抗体可用于制备治疗或预防化脓链球菌感染的药物。

[0067] 依照另一方面,本发明提供了与按照本发明的超免疫血清反应性抗原或其片段结合的拮抗剂。

[0068] 这样的能与按照本发明的超免疫血清反应性抗原或其片段结合的拮抗剂可用包含以下步骤的方法进行鉴定:

[0069] a) 在能提供响应候选拮抗剂与所述超免疫血清反应性抗原或其片段的结合的可检测信号的组分的存在下,将按照本发明的分离的或固定的超免疫血清反应性抗原或其片段与候选的拮抗剂在允许所述候选拮抗剂结合所述超免疫血清反应性抗原或其片段的条件下结合,并

[0070] b) 检测响应拮抗剂与超免疫血清反应性抗原或其片段结合所产生的信号的存在或不存在。

[0071] 能降低或抑制按照本发明的超免疫血清反应性抗原或其片段与其相互作用配体之间的相互作用活性的拮抗剂可通过包含以下步骤的方法进行鉴定:

[0072] a) 提供按照本发明的超免疫血清反应性抗原或其超免疫片段,

[0073] b) 提供针对所述超免疫血清反应性抗原或其片段的相互作用配体,尤其是按照本发明的抗体,

[0074] c) 使所述的超免疫血清反应性抗原或其片段与所述的相互作用配体相互作用形成相互作用复合物,

[0075] d) 提供候选拮抗剂,

[0076] e) 允许候选拮抗剂和所述的相互作用复合物之间发生竞争性反应,

[0077] f) 确定候选拮抗剂是否抑制或降低了超免疫血清反应性抗原或其片段与相互作用配体之间的相互作用活性。

[0078] 按照本发明的超免疫血清反应性抗原或其片段可用于分离和/或纯化和/或鉴定所述超免疫血清反应性抗原或其片段的相互作用配体。

[0079] 本发明还提供了在体外诊断涉及按照本发明的超免疫血清反应性抗原或其片段的疾病的方法,包括测定编码所述的按照本发明的超免疫血清反应性抗原或片段的核酸序列的存在或按照本发明的超免疫血清反应性抗原或其片段的存在。

[0080] 本发明还提供了在体外诊断细菌感染,尤其是化脓链球菌感染的方法,包括分析编码所述的按照本发明的超免疫血清反应性抗原或片段的核酸序列的存在或按照本发明的超免疫血清反应性抗原或其片段的存在。

[0081] 此外,本发明提供了按照本发明的超免疫血清反应性抗原或其片段在生产结合所述超免疫血清反应性抗原或其片段的肽中的应用,其中的肽是抗促成素(anticaline)。

[0082] 本发明还提供了按照本发明的超免疫血清反应性抗原或其片段在制备功能性核酸中的应用,其中的功能性核酸选自包含适体(apramer)和 spiegelmer 的组。

[0083] 按照本发明的核酸分子也可用于制备功能性核糖核酸,其中的功能性核糖核酸选自核酶、反义核酸和 siRNA。

[0084] 本发明有利地提供了一套有效、相关且全面的分离的核酸分子以及它们编码的超免疫血清反应性抗原或其片段,它们可利用来自多个人血浆库的抗体制剂和来自化脓链球菌基因组的表面表达文库从化脓链球菌中鉴定而来。因此,本发明满足了对用于制备抗体和用于鉴定有效抗化脓链球菌感染的化合物的化脓链球菌抗原、疫苗、诊断剂和产品的广泛需求。

[0085] 有效的疫苗应由蛋白质或多肽组成,它们可由所有菌株表达并能诱导高亲和力的、丰富的抗化脓链球菌细胞表面组分的抗体。所述的抗体应是用于调理作用的 IgG1 和/或 IgG3,以及用于中和粘附和毒素作用的任何 IgG 亚型和 IgA。化学定义的疫苗必定确实优于全细胞疫苗(减毒的或灭活的),因为可除去与人组织交叉反应或抑制调理作用的化脓链球菌组分 {Whitnack, E. 等, 1985}, 而且可选择诱发保护性抗体和/或保护性免疫反应的个别蛋白质。

[0086] 已应用于本发明的方法是以 A 族链球菌蛋白质或肽与人血清中存在抗体之间的相互作用为基础的。由人免疫系统产生且存在于人血清中的抗化脓链球菌抗体是抗原性蛋白质体内表达以及它们的免疫原性的指征。此外,用预选择的血清库通过细菌表面展示表达文库鉴定的抗原性蛋白质在用单独选择或产生的血清进行的第二轮和第三轮筛选中被加工。由此本发明提供了一套有效的、相关的、全面的 A 族链球菌抗原作为药物组合物,尤其是预防化脓链球菌感染的疫苗。

[0087] 在用于鉴定按照本发明的一组全面的抗原的抗原鉴定程序中,用数个血清库或血浆组分或含其它集合抗体的体液(抗体库)筛选了至少两个不同的细菌表面表达文库。抗体库来自血清样品集合,已针对化脓链球菌的抗原性化合物对所述集合进行了检测,诸如全细胞提取物和培养物上清蛋白质等。优选地,使用两种不同的血清集合:1. 用非常稳定的抗体所有组分(repertoire):正常成人,临床健康人群,他们是非携带者且克服了先前的遭遇或目前正携带化脓链球菌但无急性疾病和症状的人群,2. 由于致病生物体的存在而具有急性诱发抗体的血清集合:具不同表现的急性疾病患者(如化脓链球菌咽炎、伤口感染和菌血症)。血清必须与多个A族链球菌特异抗原反应,以便被认为是超免疫和因此与用于本发明的筛选方法相关。由人免疫系统产生并存在于人血清中的抗链球菌抗体是抗原性蛋白质体外表达及其免疫原性的指征。

[0088] 表达文库用于本发明中时应可以表达所有潜在的抗原,例如,来自化脓链球菌所有表面蛋白质的抗原。细菌表面展示文库用于在细菌宿主膜上许多选定的外膜蛋白上展示一组(全部)A族链球菌所表达肽序列的细菌宿主重组文库代表{Georgiou,G.,1997;Etz,H.等,2001}。使用重组表达文库的优势之一在于确定的超免疫血清反应性抗原可通过表达被筛选和选择且表达超免疫血清反应性抗原的克隆的编码序列即刻产生,而无需另外的重组DNA技术或克隆步骤。

[0089] 用一轮或多轮另外的筛选进一步分析根据本发明所述程序鉴定的一组全面的抗原。因此利用了抗选定的具免疫原性的肽的单个抗体制剂或抗体。依照优选的实施方案,用于第二轮筛选的单独抗体制剂来自A族链球菌急性感染患者,尤其是显示出抗体效价高于一定最低水平的患者,例如抗体效价高于被测试人(患者或健康个体)血清的80个百分点,优选高于90个百分点,尤其是高于95个百分点。在第二轮筛选中应用所述的高效价单独抗体制剂可以高选择性的鉴定来自化脓链球菌的超免疫血清反应性抗原及其片段。

[0090] 在高通量的筛选步骤之后,如果不能表达于原核表达系统中则可表达为重组蛋白或体外翻译产物的选定抗原性蛋白或者已被鉴定的抗原性肽(合成产生)在第二轮筛选中用巨大的人血清集合(>100名未感染者,>50名患者血清)通过一系列ELISA和蛋白质印迹测定进行检测评估其免疫原性。

[0091] 重要的是单独的抗体制剂(也可以是选定的血清)可以从来自第一轮筛选的所有有希望候选者中选择性的鉴定超免疫血清反应性抗原。因此,在第二轮筛选中应优选至少10种单独的抗体制剂(即,来自感染选定病原体的至少10名不同个体的抗体制剂(如血清))用于鉴定这些抗原。当然,还有可能利用少于10种的单独制剂,不过,利用少数单独抗体制剂得到的步骤的选择性可能不是最佳的。另一方面,如果给定的超免疫血清反应性抗原(或其抗原性片段)被至少10种单独的抗体制剂所识别,优选至少30种,尤其是至少50种单独的抗体制剂识别,超免疫血清反应性抗原鉴定的选择性对于进行正确的鉴定而言也已足够。当然,超免疫血清反应性可用尽可能多的单独制剂进行检测(如,多于100或甚至多于1000)。

[0092] 因此,按照本发明方法的超免疫血清反应性抗体制剂的相关部分应优选至少10种,更优选至少30种,尤其是至少50种单独的抗体制剂。备选地(或联合的)超免疫血清反应性抗原还可用第二轮筛选中所用所有单独抗体制剂的至少20%、优选至少30%、尤其是至少40%进行鉴定。

[0093] 依照本发明某一优选的实施方案,从其中制备出用于第二轮筛选的单独抗体制剂的血清(或自身被用作抗体制剂)是根据其抗化脓链球菌的效价(如,抗此病原体的制剂,诸如裂解物、细胞壁组分和重组蛋白)进行选择。优选地,当整个生物体(总裂解物或全细胞)被用作 ELISA 中的抗原时,选择总 IgA 效价高于 4000U 的、尤其是高于 6000U 的以及/或 IgG 效价高于 10000U、尤其是高于 12000U 的那些生物体(U=单位,在给定稀释度下的 OD_{405nm} 读数计算而得)。

[0094] 人免疫系统所产生且存在于人血清中的抗链球菌抗体是抗原性蛋白质体内表达及其免疫原性的指征。抗体对线性表位的识别可能是以短到 4-5 个氨基酸长的序列为基础的。当然这并不意味着这些短肽能在体内产生所述的给定抗体。为此进一步在动物内(主要是在小鼠内)检测所限定表位、多肽和蛋白质在体内诱发抗选定蛋白质的抗体的能力。

[0095] 优选的抗原定位于细胞表面或被分泌且因而易于到达细胞外。抗细胞壁蛋白质抗体预计用于两个目的:抑制粘附并促进胞噬作用。抗被分泌蛋白质的抗体在对其毒素或毒性组分功能的中和作用中是有利的。还已知的是细菌通过分泌蛋白质相互交流。中和抗这些蛋白质的抗体可阻断链球菌种属之间或之内的生长促进交流对话。生物信息分析(信号序列、细胞壁定位信号、跨膜结构域)被证实对评估细胞表面定位或分泌作用非常有用。实验方法包括分离具相应表位的抗体和来自人血清的蛋白质,并在小鼠中产生针对通过细菌表面展示筛选所选定的(多)肽的免疫血清。然后将这些血清作为以下检测的试剂用于第三轮筛选中:在不同条件下生长的 A 族链球菌的细胞表面染色(FACS,显微镜观察),测定中和能力(毒素、粘附)以及促进调理作用和胞噬作用(体外胞噬作用检测)。

[0096] 为此目的,细菌中的大肠杆菌克隆被直接注射入小鼠中,采集免疫血清并在相应的体外试验中检测功能性的调理或中和抗体。或者,特异的抗体可以肽或蛋白质作为底物纯化自人或小鼠血清。

[0097] 宿主抗化脓链球菌的防卫性主要依赖于先天的免疫机制。通过接种疫苗诱发调理型和中和型的高亲和力抗体可帮助先天的免疫系统清除细菌和毒素。这使得按照本发明的方法成为鉴定 A 族链球菌抗原性蛋白质的最适工具。

[0098] 皮肤和黏膜是抗链球菌入侵的强大屏障。不过,一旦皮肤或黏膜破裂,非适应性的细胞防卫第一道防线通过补体和巨噬细胞、尤其是多形核白细胞(PMNs)开始联合作用。这些细胞可被认为是清除入侵细菌的基石。由于 A 族链球菌起初是胞外病原体,所以主要的抗链球菌适应性反应来自免疫系统的体液防卫装备且通过三种主要的机制介导:调理作用的促进、毒素中和作用和粘附的抑制。现认为调理作用尤为重要,因为它是有效的胞噬作用所需要的。为了进行有效的调理,微生物的表面必须用抗体和补体因子包被以便 PMNs 通过受体识别 IgG 分子的 Fc 片段或被活化的 C3b。经调理作用后,链球菌被胞噬并处死。结合细菌细胞表面特异抗原的抗体作为配体用于附着 PMNs 并促进胞噬作用。与粘附素和其它细胞表面蛋白结合的非常相似的抗体被预期可中和粘附作用并预防定植。本发明所提供的抗原的选择因此非常适于鉴定在动物模型或人中导致抗感染防卫作用的那些抗原。

[0099] 依照本文所用的抗原鉴定方法,如下所述,特别是,本发明可令人惊讶的提供化脓链球菌的一套全面的新核酸和新的超免疫血清反应性抗原及其片段。根据某一方面,本发明特别涉及编码超免疫血清反应性抗原且具有序列 Seq ID No :1-150 中所列出序列的核

核苷酸序列以及序列表 SeqID No 151-300 所列出的代表超免疫血清反应性抗原的相应编码氨基酸序列。

[0100] 在本发明的优选实施方案中,提供的核酸分子其全长展示出与 Seq IDNo 1, 4-8,10-18,20,22,24-32,34-35,38-40,43-46,49-51,53-54,57-61,63,65-71,73,75-77, 81-82,88,91-94 和 96-150 所列核苷酸序列具有 70% 的同一性。最优选的核酸包含了全长与 Seq ID No 1,4-8,10-18,20,22,24-32,34-35,38-40,43-46,49-51,53-54,57-61,63, 65-71,73,75-77,81-82,88,91-94 和 96-150 所列的核酸分子具有至少 80% 或至少 85% 同一性的区域。在这点上,特别优选全长与上述序列具有至少 90%、91%、92%、93%、94%、95% 或 96% 同一性的核酸分子。而且,特别优选至少具有 97% 同一性的那些核酸分子,尤其优选至少 98% 和至少 99% 的同一性,更优选至少 99% 或 99.5%,特别优选 100% 的同一性。此外,在这方面优选的实施方案是编码超免疫血清反应性抗原或其片段(多肽)的核酸,所述的超免疫血清反应性抗原或其片段保持与由序列 Seq ID No 1,4-8,10-18,20,22, 24-32,34-35,38-40,43-46,49-51,53-54,57-61,63,65-71,73,75-77,81-82,88,91-94 和 96-150 所列的所述核酸编码的成熟多肽基本上具有相同的生物学功能或活性。

[0101] 正如本领域已知且此处所用的,同一性是两个或多个多肽序列或者两个或多个多核苷酸序列之间的关系,这是通过比较这些序列确定的。在本领域内,同一性也意味着多肽或多核苷酸序列之间的序列相关程度,情形可能是通过在所述序列行与行之间进行比对确定的。同一性可方便的计算。虽然存在许多方法可以检测两个多核苷酸或两个多肽序列之间的同一性,但此术语对技术熟练人员而言是众所周知的(如,分子生物学中的序列分析 Sequence Analysis in Molecular Biology, von Heinje, G., Academic Press, 1987)。优选的测定同一性的方法是设计成能给出待测序列之间最大的匹配程度。确定同一性的方法被编成了计算机程序。测定两个序列之间同一性的优选计算机程序法包括,但不局限于, GCG 软件包 {Devereux, J. 等, 1984}、BLASTP、BLASTN 和 FASTA {Altschul, S. 等, 1990}。

[0102] 依照本发明的另一方面,提供了展示出与 Seq ID No 64 所列出的核酸序列具有至少 96% 同一性的核酸分子。

[0103] 依照本发明的另一方面,提出了与 Seq ID No 3,36,47-48,55,62,72,80,84,95 所列出的核酸序列相同的核酸分子。

[0104] 按照本发明的核酸分子可作为与上文所述第一个备选核酸至少基本上互补的第二个备选的核酸分子。用于此处时,互补意味着核酸链通过 Watson-Crick 碱基配对原则与第二条核酸链的碱基配对。基本上互补用于此处时意味着碱基配对并非存在于各自链的所有碱基之间,而是剩余一定数目或百分比的不配对碱基或错配碱基。正确配对的碱基百分率优选至少 70%,更优选至少 80%,还更优选至少 90%,最优选高于 90% 的任一百分率。应提及的是 70% 比例的碱基匹配时被认为是同源的且具此程度匹配碱基的杂交被认为是严谨的。此类严谨杂交的杂交条件参阅分子生物学通用方法 Current Protocols in Molecular Biology (John Wiley and Sons, Inc., 1987)。更具体而言,杂交条件如下:

[0105] • 例如,在 5x SSPE、5x Denhardt 氏试剂、0.1% SDS、100g/mL 剪切 DNA 中于 68°C 进行杂交

[0106] • 在 0.2xSSC、0.1% SDS 中于 42°C 进行中度严谨洗涤

[0107] • 在 0.1xSSC、0.1% SDS 于 68°C 进行高严谨性洗涤。

[0108] GC 含量为 50% 的基因组 DNA 的 T_m 近似 96°C。相对于 1% 的错配而言, T_m 下降约 1°C。

[0109] 此外, 本文所述的任何别的杂交条件原则上也是适用的。

[0110] 当然, 编码与本发明所鉴定分子相同的多肽分子的所有核酸序列分子都可包含在给定编码序列的任何部分内, 因为遗传密码的简并性是直接可应用于明确测定的且编码给定多肽分子的所有可能性核酸分子, 即使所述的简并核酸分子的数目可能是高的也是如此。这还可适用于给定多肽的片段, 只要该片段编码多肽适用于疫苗接种即可, 如作为主动或被动疫苗。

[0111] 按照本发明的核酸分子可作为第三个备选的核酸分子, 它包含核酸分子的一段至少 15 个碱基与上述按照本发明的第一个和第二个备选核酸分子一致。优选的所述碱基形成邻近的碱基伸展段 (stretch)。不过, 由许多碱基分隔开的两个或多个部分组成的伸展段也在本发明的范围内。

[0112] 按照本发明的核酸分子也可用作第四个备选的核酸分子, 它在严谨杂交条件下与符合上述第一、第二和第三备选方案的任一本发明核酸分子退火。严谨杂交条件通常是本文所述的条件。

[0113] 最后, 按照本发明的核酸分子可用作第五个备选的核酸分子, 若非遗传密码的简并性, 可与符合上述第一、第二和第三备选方案的任一本发明核酸分子的任一核酸分子杂交。此类核酸分子是指事实, 即优选按照本发明的核酸编码按照本发明的超免疫血清反应性抗原或其片段。此类核酸分子可特别用于检测按照本发明的核酸分子并从而诊断诸如化脓链球菌等相应的微生物以及涉及此类微生物的任何疾病或发病状况。优选地, 杂交可在与上述第四个备选方案相关的所述严谨条件下发生或完成。

[0114] 本文所用的核酸分子通常指任何核糖核酸分子或脱氧核糖核酸分子, 它们可以是未修饰的 RNA 或 DNA 或修饰过的 RNA 或 DNA。因此, 例如, 尤其是, 本文所用的核酸分子指单链和双链 DNA、作为单链和双链 RNA 混合物的 DNA 以及作为单链和双链区域混合物的 RNA, 含有 DNA 或 RNA 的可能是单链或更通常是双链或三链的杂合分子或单链和双链区域的混合物。此外, 用于此处时核酸分子涉及包含 RNA 或 DNA 或兼有 RNA 和 DNA 二者的三链区域。在所述区域内的链可能来自同一分子或来自不同的分子。这些区域可包含所有的一个或多个分子, 但更通常是只涉及某些分子的一个区域。三螺旋区域的分子之一常常是寡核苷酸。用于此处时, 术语核酸分子包括含有一个或多个修饰碱基的上述 DNA 或 RNA。因此, 由于稳定性或其它原因而其骨干被修饰的 DNA 或 RNA 是本文提及术语时所指的“核酸分子”。此外, 含诸如肌苷等稀有碱基或诸如三苯甲基化碱基等修饰碱基 (只是其中的两个例子) 的 DNA 或 RNA 是该术语用于此处时所指的核酸分子。还应理解的是为了达到许多本领域技术人员已知的有益的目的而对 DNA 和 RNA 进行了多种类型的修饰。术语核酸分子用于此处时包含化学、酶促或代谢修饰形式的核酸分子以及病毒和细胞的化学形式 DNA 和 RNA 特性, 其中包括简单和复杂细胞。术语核酸分子还包含常常被称为寡核苷酸的短核酸分子。“寡核苷酸”和“核酸”或“核酸分子”在此常常交换使用。

[0115] 本发明中所提供的核酸分子还包括许多单独的片段, 长于和短于化脓链球菌编码区序列列表中所列出的核酸分子序列都可, 它们可通过标准的克隆方法产生。为了达到唯一性, 片段必须有足够长度以将其与其它已知的核酸序列区分开来, 通过将任何选定的化脓

链球菌片段与诸如 GenBank 等计算机数据库中的核苷酸序列进行比较可最方便的确定。

[0116] 此外,可以对本发明所包含的核酸分子和多肽进行修饰。例如,可以进行不会影响核酸所编码多肽的核苷酸替代,因此编码超免疫血清反应性抗原或其片段的任何核酸分子都包含在本发明范围内。

[0117] 而且,可以利用诸如标准的克隆技术等标准技术将编码本发明所提供超免疫血清反应性抗原或其片段的任一核酸分子与预期的调控序列(无论是化脓链球菌调控序列或异源的调控序列都可)、异源前导序列、异源标记序列或异源编码序列进行功能型连接以产生融合蛋白。

[0118] 本发明的核酸分子可以通过克隆获得或通过化学合成技术产生或二者联合方法获得的诸如 mRNA 或 cRNA 等 RNA 形式或包括例如 cDNA 和基因组 DNA 在内的 DNA 形式。所述 DNA 可以是三链、双链或单链的。单链 DNA 可以是编码链,也已知为有义链,或者可以是非编码链,也称为反义链。

[0119] 本发明进一步涉及本文上述核酸分子的变体,它们编码具有序列表中所列出的推断化脓链球菌的氨基酸序列的超免疫血清反应性抗原和其片段的片段、类似物及衍生物。核酸分子变体可以是天然存在的变体,诸如天然存在的等位基因变体,或者可以是未知天然存在的变体。所述的核酸分子非天然存在变体可以通过诱变技术制备,包括应用于核酸分子、细胞或生物体的那些诱变技术。

[0120] 在变体中这方面的变体是通过核苷酸替代、缺失或添加而不同于上述核酸分子的变体。替代、缺失或添加可涉及一个或多个核苷酸。所述变体可以改变于编码区内或非编码区内或二者都有所改变。编码区的改造可能产生保守或非保守性氨基酸替代、缺失或添加。优选的是编码具有序列表中所提化脓链球菌序列的变体、类似物、衍生物或片段或者片段的变体、类似物或衍生物的核酸分子,其中数个、一些、5-10 个、1-5 个、1-3 个、2 个、1 个或无氨基酸以任一组合形式被替代、缺失或插入。其中尤其优选沉默替代、添加和缺失,它们不改变序列表中所列化脓链球菌多肽的特征和活性。在这点上还特别优选保守性替代。

[0121] 按照本发明的肽和片段还包括修饰的表位以及编码所述修饰表位的核酸序列,其中优选给定表位的一个或多个氨基酸根据上述例子中的原则被修饰或置换。

[0122] 同时还清楚的是,通过提高、保留或至少不显著阻碍所述表位的 T 细胞活化能力的氨基酸调换来自本发明表位的表位也在按照本发明的表位范围内。因此本发明的表位还覆盖了不包含化脓链球菌来源的原始序列但触发相同的或优选提高了的 T 细胞应答的表位。这些表位被称为“不规则变化的”,它们需要具有对 MHC/HLA 分子相似的或优选更大的亲合力,还需要具有以相似或优选更强的方式将 T 细胞受体(TCR)定向于原始表位的能力。

[0123] 不规则变化表位可通过合理的设计获得,即,如 {Rammensee, H. 等, 1999} 所述,考虑各个残基对结合 MHC/HLA 的贡献,联合与 TCR 潜在相互作用的残基的系统性交换并用被定向针对原始表位的 T 细胞检测所产生序列。这样的设计对于本领域技术人员而言是可能的而无需过多的实验。

[0124] 另一可能性包括用定向针对原始表位的 T 细胞筛选肽库。优选的方式是合成肽库的定位扫描。所述方法已详述于 {Hemmer, B. 等, 1999} 和本文所给定文献中。

[0125] 除了由本发明来源氨基酸序列表示的表位或不规则变化表位之外,还可应用模拟这些表位的物质,如“肽模拟物”或“反-逆-肽(retro-inverso-peptides)”。

[0126] 改良表位设计的另一方面是用增加其刺激 T 细胞能力的物质进行它们的配制或修饰。这些包括 T 辅助细胞表位、脂类或脂质体或 WO 01/78767 中所述的优选修饰。

[0127] 增加表位的 T 细胞刺激能力的另一方式将它们与免疫刺激物质配制在一起,例如细胞因子或趋化因子样白介素 -2、-7、-12、-18, I 类和 II 类干扰素 (IFN), 尤其是 γ -IFN、GM-CSF、 α -TNF、flt3 配体及其他。

[0128] 正如本文关于本发明核酸分子检测所另外论及的,例如,上述本发明核酸分子可用作针对 RNA、cDNA 和基因组 DNA 的杂交探针用于分离编码本发明多肽的全长 cDNAs 和基因组克隆,以及用于分离与本发明核酸分子具有高度序列相似性的其它基因的 cDNA 和基因组克隆。所述的探针通常包含至少 15 个碱基。优选地,所述探针具有至少 20 个、至少 25 个或至少 30 个碱基,还可能具有至少 50 个碱基。尤其优选的探针具有至少 30 个碱基,和具有 50 个或少于 50 个的碱基,诸如 30、35、40、45 或 50 个碱基。

[0129] 例如,本发明核酸分子的编码区可通过用已知的 DNA 序列合成寡核苷酸探针筛选有关的文库而被分离。具有与本发明基因序列互补的序列的已标记寡核苷酸然后被用于筛选 cDNA、基因组 DNA 或 mRNA 文库以确定探针与文库中的哪一个成员杂交。

[0130] 正如本文关于核酸分子试验所进一步讨论的,其中,本发明的核酸分子和多肽可用作试剂和材料用于研发疾病,尤其是人的疾病的治疗和诊断。

[0131] 寡核苷酸形式的本发明核酸分子可用于本文所述方法中,但优选用于 PCR,以确定此处鉴定的化脓链球菌基因是否全部或部分存在于和 / 或转录于诸如血液等受感染组织中。公认的,所述序列还可用于诊断病原体已达到的感染阶段和感染类型。为此目的和其它目的,可使用至少包含按照本发明的核酸之一的阵列。

[0132] 按照本发明的核酸分子可用于检测核酸分子和含这些核酸分子的生物体或样品。优选所述检测是用于诊断的,更优选用于诊断涉及或与化脓链球菌的存在或富集有关的疾病。

[0133] 可用各种技术在 DNA 水平检测感染了化脓链球菌的真核细胞(此处也称“个体”),特别是哺乳动物,尤其是人。用于将化脓链球菌从其它生物体中辨别出来的优选候选物是可以获得的。

[0134] 本发明提供了诊断化脓链球菌感染所引起疾病的方法,包括从分离或来自某一个体的样品中测定具有序列表中所列出的核酸分子序列的核酸分子表达水平的增加。核酸分子的表达可以用任何本领域已知的核酸分子定量方法进行检测,例如,PCR、RT-PCR、核糖核酸酶保护、RNA 印迹、其它的杂交方法以及本文所述的阵列。

[0135] 分离的用于此处是指“人为的”将其从自身的天然状态中隔离出来,即,如果它存在于自然界中,那么它已被改变或从其最初的环境中取走,或二者兼具。当该术语用于此处时,例如,在其天然状态下天然存在于活的生物体内的天然存在核酸分子或多肽不是“分离的”,但从其天然状态共存物质中隔离出来的同一核酸分子或多肽就是“分离的”。作为分离的一部分或分离后,所述核酸分子可与诸如 DNA 等其它核酸分子结合,例如,为了诱变形成融合蛋白,以及为了繁殖或表达于宿主中。分离的核酸分子,独自或与诸如载体等其它核酸分子结合,可引入培养物或完整生物体的宿主细胞内。该术语用于此处时,当引入培养物中或完整生物体的宿主细胞中时,所述 DNA 仍然是分离的,因为它们不会是处于其天然存在形式或环境中。同样的,核酸分子和多肽可存在于组合物中,诸如培养基制剂、将核酸分子

或多肽引入例如细胞中的溶液、化学或酶促反应的组合物或溶液，它们不是天然存在的组合物且其中保留了该术语用于此处时所具涵义的分离的核酸分子或多肽。

[0136] 按照本发明的核酸可化学合成。或者，所述核酸可用本领域技术人员已知的方法分离自化脓链球菌。

[0137] 依照本发明的另一方面，用此处所述的抗原鉴定方法提供了一套全面的、新的超免疫血清反应性抗原和其片段。在优选的本发明实施方案中，提供了含有此处所述任一核酸分子编码的氨基酸序列的超免疫血清反应性抗原及其片段。在本发明的另一优选实施方案中，提供了所含氨基酸序列选自 Seq ID No 151, 154-158, 160-168, 170, 172, 174-182, 184-185, 188-190, 193-196, 199-201, 203-204, 207-211, 213, 215-221, 223, 225-227, 231-232, 238, 241-244 和 246-300 所表示多肽序列的一组新的超免疫血清反应性抗原及其片段。在本发明进一步优选的实施方案中，提供了所含氨基酸序列选自 Seq ID No 214 中所示多肽序列的超免疫血清反应性抗原及其片段。在本发明更优选的实施方案中，提供了所含氨基酸序列选自 SeqID No 153, 186, 197-198, 205, 212, 222, 230, 234, 245 中所示多肽序列的超免疫血清反应性抗原及其片段。

[0138] 本发明中所提供的超免疫血清反应性抗原及其片段包括序列表中列出的任何多肽以及与序列表中提出多肽具有至少 70% 同一性的多肽，优选与序列表中提出多肽具有至少 80% 或 85% 的同一性、更优选与序列表中提出多肽具有至少 90% 的相似性（更优选至少 90% 的同一性）、还更优选与序列表中提出多肽具有至少 95%、96%、97%、98%、99% 或 99.5% 的相似性（更优选至少 95%、96%、97%、98%、99% 或 99.5% 的同一性）的多肽，还包括具有该多肽所述部分的所述多肽部分，该多肽的所述部分通常含至少 4 个氨基酸，更优选至少 8 个，还更优选至少 30 个，更优选至少 50 个氨基酸，诸如 4、8、10、20、30、35、40、45 或 50 个氨基酸。

[0139] 本发明还涉及这些超免疫血清反应性抗原及其片段的片段、类似物和衍生物。术语“片段”、“衍生物”和“类似物”在涉及其氨基酸序列列于序列表中的抗原时，指基本上保持与所述超免疫血清反应性抗原及其片段相同的生物学功能或活性的多肽。

[0140] 超免疫血清反应性抗原及其片段的片段、衍生物或类似物可能是 1) 其中一个或多个氨基酸残基被保守或不保守氨基酸残基替代（优选保守的氨基酸残基）且所述的替代的氨基酸残基可能是或可能不是由该遗传密码编码的，或 2) 其中一个或多个氨基酸残基包含一替代基团，或 3) 其中成熟的超免疫血清反应性抗原或其片段与另一化合物融合，诸如延长超免疫血清反应性抗原及其片段半寿期的化合物（例如，聚乙二醇），或 4) 其中附加的氨基酸残基与成熟的超免疫血清反应性抗原或其片段融合，诸如前导序列或分泌序列或用于纯化成熟的超免疫血清反应性抗原及其片段的序列或前蛋白质序列。所述片段、衍生物和类似物被认为对本领域技术人员而言在本发明的范围内。

[0141] 本发明还涉及不同化脓链球菌分离株的抗原。所述的同源物可以此处公开的核酸和氨基酸序列为基础方便的分离。迄今为止已分清的有 80 种以上的 M 蛋白血清型，且此分型是以 emm 基因 5' 端的可变区为基础进行的（参阅，例如 Vitali 等 2002）。因此任一抗原的存在都可针对每一种 M 血清型来确定。此外还可能按照针对 sic 基因所述确定各种 M 血清型中特殊抗原的可变性（Hoe 等，2001）。各种 M 血清型对其引起疾病种类的影响总结于最近的综述中（Cunningham, 2000）。具体而言，两组血清型可被区分：

[0142] 1) 引起咽炎和猩红热的血清型 (如, M 型 1, 3, 5, 6, 14, 18, 19, 24)

[0143] 2) 引起脓皮病和链球菌导致的皮肤感染的血清型 (如, M 型 2, 49, 57, 59, 60, 61)

[0144] 这可作为鉴定用作疫苗或通常用作靶向特定疾病的药物的适合性的基础。

[0145] 来自例如 CDC(<http://www.cdc.gov/ncidod/biotech/strep/emmtypes.htm>) 主页的信息给出了显示各种 M 血清型关系的系统树图。另外的相关文献有: Vitali 等, *Journal of Clinical Microbiology* 40:679-681. (2002) (分子 emm 分型方法), Enright 等, *Infection and Immunity* 69:2416-2427. (2001) (可供选择的分子分型方法 alternative molecular typing method(MLST)), Hoe 等, *The Journal of Infectious Diseases* 183:633-639. (2001) (在许多不同的血清型中某一抗原 (sic) 变化的例子 example for the variation of one antigen (sic) in many different serotypes) 以及 Cunningham, *CLINICAL MICROBIOLOGY REVIEWS* 13:470-511. (2000) (关于 GAS 发病机理的综述 review on GAS pathogenesis)。所有的 emm 型都完整列出并可下载自上述网址。

[0146] 系统树图是通过以下软件的顺序使用构建的: Wisconsin Package Version 10.1、Genetics Computer Group (GCG)、Madison programs Pileup, Distances 和 Growtree。基本上, 信号序列的 22 个残基加 83 个附加 N 末端残基被用于包括来自数据库的选定序列在内的序列比对。选定的序列包括编号 103-124 的新的 emm (如下表中所述) 以及它们的最接近的“传统的”M 蛋白匹配序列。尽管此分析局限于 C 末端是被任意截短的, 这在将不透明因子阳性株系 M 序列的分支与不透明因子阴性株系 M 序列分隔开的系统树图中仍是一种典型的结果。

[0147] emm 型 / 先前的命名 - GenBank 登录号 - 在哪一国家被分离 - 最接近的 N 末端 M 蛋白序列匹配程度 (% 同一性):

[0148] emm103/st2034 U74320 PNG, Bra, Egy, Mal, Nep, NZ, US M87 (66%)

[0149] emm104/st2034AF056300 PNG, Egy, Mal, Nep, NZ, US M66 (72%)

[0150] emm105/st4529AF060227 Mal, Nep, NZ, US M5 (45%)

[0151] emm106/st4532AF077666 Mal, Egy, Iran, Nep M27G (71%)

[0152] emm107/st4264AF163686 Mal, NZ M25 (52%)

[0153] emm108/st4547AF052426 Mal, Bra, Egy, Ira, NZ M70 (84%)

[0154] emm109/st3018 AF077667 Mal, Egy, NZ M28 (74%)

[0155] emm110/st4935U92492 Ind, Bul, NZ, Rus, US M13 (60%)

[0156] emm111/st4973AF128960 Ind, Bra, Nep, US M80 (40%)

[0157] emm112/stCmuk16AF091806 Thi, Bra, Rus, US M27L/77 (59%)

[0158] emm113/st2267AF078068 NZ, Thai, Chi M13 (50%)

[0159] emm114/st2967U50338 US, Can, Gam, NZ, PNG M73 (80%)

[0160] emm115/st2980AF028712 US, Bra, Rus M36 (64%)

[0161] emm116/st2370AF156180 US, Nep, NZ M52 (60%)

[0162] emm117/st436AF058801 US M13 (59%)

[0163] emm118/st448AF058802 US, Bra, Egy, Nep, NZ M49 (79%)

[0164] emm119/st3365AF083874 US, Br, Nep M52 (59%)

[0165] emm120/st1135AF296181 Egy M56 (78%)

[0166] emm121/st1161AF296182Egy M64(64%)

[0167] emm122/st1432AF222860Egy, Rus, Nep M18(40%)

[0168] emm123/st6949AF213451Arg, US, NZM80(68%)

[0169] st1160/emm124AF149048and AF018178Egy, Mal, NZM2(82%)

[0170] 缩写:Arg,阿根廷;Bra,巴西;Bul,保加利亚;Can,加拿大;Chi,智利;Egy,埃及;Gam,冈比亚;Ind,印度;Ira,伊朗;Mal,马来西亚;Nep,尼泊尔;NZ,新西兰;PNG,巴布亚新几内亚;Thi,泰国;Rus,俄罗斯;US,美国。%:最接近的成熟 M 蛋白序列与来自血清学表征 Lancefield 类型的预测 50 个成熟 N 末端残基相比。

[0171] emm 类型和序列类型:

[0172] 在许多情形中 emm 序列参照株系直接来自 Rebecca Lancefield 博士的 M 类型保藏中心。所述的株系被命名为 RCL。

[0173] 以“emm”开始的序列表明用此类型代表的分离株已经过除 CDC 链球菌实验室之外的数个对照实验室分析。emm94 至 emm124 中每一“新的”emm 型都是由回收自严重发病状态的多个独立分离株代表的,是用国际 GAS 对照实验室可获得的所有分型血清贮液都不可分型的 M 蛋白,并且通过在正常人血液中的繁殖证实了其体外抗噬菌细胞的特性。具有以“st”(序列类型)开始的 emm 序列的株系还未被所有参照实验室完全确认。

[0174] GAS 遗传学:

[0175] 长期以来已知针对血清不透明因子阳性(SOF+)株系的抗血清以株系特异的方式抑制 OF 活性。因此,从 GAS 不透明因子阳性株系中分析出代表至少 60 种单独 sof 基因的 500-2700 个碱基的 sof(血清不透明因子)基因可变区(有趣的是,通常在 OF 阴性 emm12 分离株和 emm/M 类型 12 对照株中找到类似物)。已发现 sof 基因序列也在不同 GAS 株系中显著可变,尽管它们在 emm 类型中通常是非常保守的。因此重要的株系包括 emm1, emm100, emm101, emm102, emm103, emm104, emm105, emm106, emm107, emm108, emm109, emm11, emm110, emm111, emm112, emm113, emm114, emm115, emm116, emm117, emm118, emm119, emm12, emm120, emm121, emm122, emm123, emm124, emm13L, emm14, emm15, emm17, emm18, emm19, emm2, emm22, emm23, emm24, emm25, emm26, emm27G, emm28, emm29, emm3, emm30, emm31, emm32, emm33, emm34, emm36, emm37, emm38, emm39, emm4, emm40, emm41, emm42, emm43, emm44, emm46, emm47, emm48, emm49, emm5, emm50, emm51, emm52, emm53, emm54, emm55, emm56, emm57, emm58, emm59, emm6, emm60, emm61, emm62, emm63, emm64, emm65, emm66, emm67, emm68, emm69, emm70, emm71, emm72, emm73, emm74, emm75, emm76, emm77, emm78, emm79, emm8, emm80, emm81, emm82, emm83, emm84, emm85, emm86, emm87, emm88, emm89, emm9, emm90, emm91, emm92, emm93, emm94, emm95, emm96, emm97, emm98, emm99, st1389, st1731, st1759, st1815, st1967, st1969, st1rp31, st11014, st2037, st204, st211, st213, st2147, st1207, st245, st2460, st2461, st2463, st2904, st2911, st2917, st2926, st2940, st369, st3757, st3765, st3850, st5282, st6735, st7700, st809, st833, st854, st980584, stck249, stck401, std432, std631, std633, stIL103, stIL62, stns292, stns554, sts104, stcl400, stcl741, stc36, stc3852, stc5344, stc5345, stc57, stc6979, stc74a, stc839, stgl10, stgl11, stgl1389, stgl166b, stgl1750, stg2078, stg3390, stg4222, stg4545, stg480, stg4831, stg485, stg4974, stg5063, stg6, stg62647, stg643, stg652, stg653, stg663, stg840, stg93464, stg97, stL1376, stL1929 和

stL2764。

[0176] 在本发明特别优选的实施方案中有关这方面的是序列表中列举的超免疫血清反应性抗原,其变体、类似物、衍生物和片段,以及片段的变体、类似物和衍生物。此外,含有所述超免疫血清反应性抗原、其变体、类似物、衍生物和片段以及片段的变体、类似物和衍生物的融合多肽也包括在本发明中。所述的融合多肽和蛋白质以及编码它们的核酸分子可用标准技术方便的制备,包括用于产生和表达编码融合蛋白的重组多核苷酸的标准重组技术。

[0177] 优选的变体是那些通过保守氨基酸替换从参照株系改变而来的变体。所述替代是用相似特性的另一氨基酸替代多肽中给定的氨基酸。通常被看做保守替代的是脂肪族氨基酸丙氨酸、缬氨酸、亮氨酸和异亮氨酸相互间的置换,羟基残基丝氨酸和苏氨酸之间的交换,酸性残基天冬氨酸和谷氨酸之间的调换,氨基残基天冬酰胺和谷氨酰胺之间的替代,碱性残基赖氨酸和精氨酸之间的交换和芳香族残基苯丙氨酸和酪氨酸之间的置换。

[0178] 在这方面更特别优选的是具有序列表中列举的任一多肽的氨基酸序列的变体、类似物、衍生物和片段以及片段的变体、类似物和衍生物,其中数个、一些、5-10 个、1-5 个、1-3 个、2 个、1 个或无氨基酸残基以任何组合方式被替代、缺失或插入。其中特别优选的不改变本发明多肽特征和活性的沉默替代、插入和缺失。在这点上还特别优选的是保守性替代。最优选的是具有序列表中所列举氨基酸序列而且无替代的多肽。特别适合的氨基酸替代是按照本申请书的序列表中所公开序列的同源物中所包含的那些替代。因此本文所公开的抗原或表位的适当序列衍生物包括存在于一个或多个化脓链球菌株系或血清型中的一种或多种变异(优选以所述同源变异为基础的 1、2、3、4、5、6、7、8、9 或 10 个氨基酸交换)。所述抗原包含序列可能是天然存在序列或新产生的人造序列。这些优选的抗原变体是以所述的天然存在序列变异为基础的,例如,形成针对按照本发明的多肽抗原区的“掌控序列”。所述同源变异或交换的适当例子参阅实施例部分的表 5。例如,某一给定的化脓链球菌序列可通过涉及所述的一种或多种变异被改进从而产生此给定(天然存在)抗原或表位序列的人造(即,非天然存在)的变体。

[0179] 本发明的超免疫血清反应性抗原及其片段优选以分离的形式提供,且优选纯化至均一。

[0180] 也属于本发明优选的实施方案的是含有具有序列表中所列举氨基酸序列的多肽片段以及序列表中所列举多肽的变体和衍生物片段的多肽。

[0181] 在这方面所述片段是具有与部分而非所有上述超免疫血清反应性抗原和它的片段以及变体或衍生物、类似物及其片段完全相同的氨基酸序列的多肽。所述片段可能是“独立的”,即,并非其它氨基酸或多肽的一部分或与它们相融合,或者它们可包含于较大的多肽中形成其中的一部分或区域。在本发明的这方面还优选的是通过本发明多肽的结构或功能特性表征的片段,即,含有本发明多肽的 α -螺旋和 α -螺旋形成区、 β -折叠和 β -折叠形成区、转角和转角形成区、卷曲和卷曲形成区、亲水区、疏水区、 α 两性分子区、 β 两性分子区、可变区、表面形成区、底物结合区和高抗原性指示区的片段以及所述片段的组合。优选的区域是介导本发明的超免疫血清反应性抗原及其片段活性的那些区域。在这点上最优选的是具有本发明超免疫血清反应性抗原及其片段的化学、生物学或其它活性的片段,包括那些具有相似的活性或改良的活性或不合需要的活性有所降低的片段。尤其优选的是含

酶受体或结构域的片段,所述酶具有化脓链球菌生存或在人体内引发疾病所必需的功能。更优选的多肽片段是包括或含有动物(尤其是人)中的抗原性或免疫原性决定簇的那些多肽片段。

[0182] 抗原性片段被定义为经鉴定本身可作为抗原或当作为半抗原时可制备抗原的抗原片段。因此,显示出一个或(对于较长片段而言)只有一些氨基酸交换的抗原或抗原片段也能用本发明激活,条件是具有氨基酸交换的所述片段的抗原能力并未因该交换而严重受损即可,即,适于在以此抗原进行免疫接种的个体中引起恰当的免疫反应并用来自个体血清的个体抗体制剂进行鉴定。

[0183] 所述超免疫血清反应性抗原片段的优选例子选自含有表 1 栏“预测免疫原性氨基酸”和“已鉴定的免疫原性区域的位置”的氨基酸序列、表 2 血清反应性表位的肽,尤其是含以下区段氨基酸序列的肽:Seq ID No 151 的第 4-44 位、57-65 位、67-98 位、101-107 位、109-125 位、131-144 位、146-159 位、168-173 位、181-186 位、191-200 位、206-213 位、229-245 位、261-269 位、288-301 位、304-317 位、323-328 位、350-361 位、374-384 位、388-407 位、416-425 和 1-114 位;Seq ID No 152 的第 5-17 位、49-64 位、77-82 位、87-98 位、118-125 位、127-140 位、142-150 位、153-159 位、191-207 位、212-218 位、226-270 位、274-287 位、297-306 位、325-331 位、340-347 位、352-369 位、377-382 位、390-395 和 29-226 位;Seq ID No 153 的第 4-16 位、20-26 位、32-74 位、76-87 位、93-108 位、116-141 位、148-162 位、165-180 位、206-219 位、221-228 位、230-236 位、239-245 位、257-268 位、313-328 位、330-335 位、353-359 位、367-375 位、394-403 位、414-434 位、437-444 位、446-453 位、456-464 位、478-487 位、526-535 位、541-552 位、568-575 位、577-584 位、589-598 位、610-618 位、624-643 位、653-665 位、667-681 位、697-718 位、730-748 位、755-761 位、773-794 位、806-821 位、823-831 位、837-845 位、862-877 位、879-889 位、896-919 位、924-930 位、935-940 位、947-955 位、959-964 位、969-986 位、991-1002 位、1012-1036 位、1047-1056 位、1067-1073 位、1079-1085 位、1088-1111 位、1130-1135 位、1148-1164 位、1166-1173 位、1185-1192 位、1244-1254 和 919-929 位;Seq ID No 154 的第 5-44 位、62-74 位、78-83 位、99-105 位、107-113 位、124-134 位、161-174 位、176-194 位、203-211 位、216-237 位、241-247 位、253-266 位、272-299 位、323-349 位、353-360 和 145-305 位;Seq ID No 155 的第 15-39 位、52-61 位、72-81 位、92-97 和 71-81 位;Seq ID No 156 的第 13-19 位、21-31 位、40-108 位、115-122 位、125-140 位、158-180 位、187-203 位、210-223 位、235-245 和 173-186 位;Seq ID No 157 的第 5-12 位、19-27 位、29-39 位、59-67 位、71-78 位、80-88 位、92-104 位、107-124 位、129-142 位、158-168 位、185-191 位、218-226 位、230-243 位、256-267 位、272-277 位、283-291 位、307-325 位、331-344 位、346-352 和 316-331 位;Seq ID No 158 的第 6-28 位、43-53 位、60-76 位、93-103 和 21-99 位;Seq ID No 159 的第 10-30 位、120-126 位、145-151 位、159-169 位、174-182 位、191-196 位、201-206 位、214-220 位、222-232 位、254-272 位、292-307 位、313-323 位、332-353 位、361-369 位、389-396 位、401-415 位、428-439 位、465-481 位、510-517 位、560-568 和 9-264 位;Seq ID No 160 的第 5-29 位、39-45 位、107-128 和 1-112 位;Seq ID No 161 的第 4-38 位、42-50 位、54-60 位、65-71 位、91-102 和 21-56 位;Seq ID No 162 的第 4-13 位、19-25 位、41-51 位、54-62 位、68-75 位、79-89 位、109-122 位、130-136 位、172-189 位、192-198 位、

217-224 位、262-268 位、270-276 位、281-298 位、315-324 位、333-342 位、353-370 位、376-391 和 23-39 位 ;Seq ID No 163 的第 6-41 位、49-58 位、62-103 位、117-124 位、147-166 位、173-194 位、204-211 位、221-229 位、255-261 位、269-284 位、288-310 位、319-325 位、348-380 位、383-389 位、402-410 位、424-443 位、467-479 位、496-517 位、535-553 位、555-565 位、574-581 位、583-591 和 474-489 位 ;SeqID No 164 的第 8-35 位、52-57 位、66-73 位、81-88 位、108-114 位、125-131 位、160-167 位、174-180 位、230-235 位、237-249 位、254-262 位、278-285 位、308-314 位、321-326 位、344-353 位、358-372 位、376-383 位、393-411 位、439-446 位、453-464 位、471-480 位、485-492 位、502-508 位、523-529 位、533-556 位、558-563 位、567-584 位、589-597 位、605-619 位、625-645 位、647-666 位、671-678 位、690-714 位、721-728 位、741-763 位、766-773 位、777-787 位、792-802 位、809-823 位、849-864 和 37-241 位、409-534 位、582-604 位、743-804 位 ;Seq ID No 165 的第 4-17 位、24-36 位、38-44 位、59-67 位、72-90 位、92-121 位、126-149 位、151-159 位、161-175 位、197-215 位、217-227 位、241-247 位、257-264 位、266-275 位、277-284 位、293-307 位、315-321 位、330-337 位、345-350 位、357-366 位、385-416 和 202-337 位 ;Seq ID No 166 的第 4-20 位、22-46 位、49-70 位、80-89 位、96-103 位、105-119 位、123-129 位、153-160 位、181-223 位、227-233 位、236-243 位、248-255 位、261-269 位、274-279 位、283-299 位、305-313 位、315-332 位、339-344 位、349-362 位、365-373 位、380-388 位、391-397 位、402-407 和 1-48 位 ;Seq ID No 167 的第 18-37 位、41-63 位、100-106 位、109-151 位、153-167 位、170-197 位、199-207 位、212-229 位、232-253 位、273-297 和 203-217 位 ;Seq ID No 168 的第 20-26 位、54-61 位、80-88 位、94-101 位、113-119 位、128-136 位、138-144 位、156-188 位、193-201 位、209-217 位、221-229 位、239-244 位、251-257 位、270-278 位、281-290 位、308-315 位、319-332 位、339-352 位、370-381 位、388-400 位、411-417 位、426-435 位、468-482 位、488-497 位、499-506 位、512-521 和 261-273 位 ;Seq ID No169 的第 6-12 位、16-36 位、50-56 位、86-92 位、115-125 位、143-152 位、163-172 位、193-203 位、235-244 位、280-289 位、302-315 位、325-348 位、370-379 位、399-405 位、411-417 位、419-429 位、441-449 位、463-472 位、482-490 位、500-516 位、536-543 位、561-569 位、587-594 位、620-636 位、647-653 位、659-664 位、677-685 位、687-693 位、713-719 位、733-740 位、746-754 位、756-779 位、792-799 位、808-817 位、822-828 位、851-865 位、902-908 位、920-938 位、946-952 位、969-976 位、988-1005 位、1018-1027 位、1045-1057 位、1063-1069 位、1071-1078 位、1090-1099 位、1101-1109 位、1113-1127 位、1130-1137 位、1162-1174 位、1211-1221 位、1234-1242 位、1261-1268 位、1278-1284 位、1312-1317 位、1319-1326 位、1345-1353 位、1366-1378 位、1382-1394 位、1396-1413 位、1415-1424 位、1442-1457 位、1467-1474 位、1482-1490 位、1492-1530 位、1537-1549 位、1559-1576 位、1611-1616 位、1624-1641 和 1-414 位、443-614 位、997-1392 位 ;Seq ID No 170 的第 14-42 位、70-75 位、90-100 位、158-181 和 1-164 位 ;Seq ID No 171 的第 4-21 位、30-36 位、54-82 位、89-97 位、105-118 位、138-147 和 126-207 位 ;Seq ID No 172 的第 4-21 位、31-66 位、96-104 位、106-113 位、131-142 和 180-204 位 ;Seq ID No 173 的第 5-23 位、31-36 位、38-55 位、65-74 位、79-88 位、101-129 位、131-154 位、156-165 位、183-194 位、225-237 位、245-261 位、264-271 位、279-284 位、287-297 位、313-319 位、

327-336 位、343-363 位、380-386 和 11-197 位、204-219 位、258-372 位 ;Seq ID No 174 的第 4-20 位、34-41 位、71-86 位、100-110 位、113-124 位、133-143 位、150-158 位、160-166 位、175-182 位、191-197 位、213-223 位、233-239 位、259-278 位、298-322 和 195-289 位 ;Seq ID No 175 的第 4-10 位、21-35 位、44-52 位、54-62 位、67-73 位、87-103 位、106-135 位、161-174 位、177-192 位、200-209 位、216-223 位、249-298 位、304-312 位、315-329 和 12-130 位 ;Seq ID No 176 的第 10-27 位、33-38 位、48-55 位、70-76 位、96-107 位、119-133 位、141-147 位、151-165 位、183-190 位、197-210 位、228-236 位、245-250 位、266-272 位、289-295 位、297-306 位、308-315 位、323-352 位、357-371 位、381-390 位、394-401 位、404-415 位、417-425 位、427-462 位、466-483 位、485-496 位、502-507 位、520-529 位、531-541 位、553-570 位、577-588 位、591-596 位、600-610 位、619-632 位、642-665 位、671-692 位、694-707 和 434-444 位 ;Seq ID No 177 的第 6-14 位、16-25 位、36-46 位、52-70 位、83-111 位、129-138 位、140-149 位、153-166 位、169-181 位、188-206 位、212-220 位、223-259 位、261-269 位、274-282 位、286-293 位、297-306 位、313-319 位、329-341 位、343-359 位、377-390 位、409-415 位、425-430 和 360-375 位 ;Seq ID No 178 的第 4-26 位、28-48 位、54-62 位、88-121 位、147-162 位、164-201 位、203-237 位、245-251 和 254-260 位 ;Seq ID No 179 的第 12-21 位、26-32 位、66-72 位、87-93 位、98-112 位、125-149 位、179-203 位、209-226 位、233-242 位、249-261 位、266-271 位、273-289 位、293-318 位、346-354 位、360-371 位、391-400 和 369-382 位 ;Seq ID No 180 的第 11-38 位、44-65 位、70-87 位、129-135 位、140-163 位、171-177 位、225-232 位、238-249 位、258-266 位、271-280 位、284-291 位、295-300 位、329-337 位、344-352 位、405-412 位、416-424 位、426-434 位、436-455 位、462-475 位、478-487 和 270-312 位 ;Seq ID No 181 的第 5-17 位、34-45 位、59-69 位、82-88 位、117-129 位、137-142 位、158-165 位、180-195 位、201-206 位、219-226 位、241-260 位、269-279 位、292-305 位、312-321 位、341-347 位、362-381 位、396-410 位、413-432 位、434-445 位、447-453 位、482-487 位、492-499 位、507-516 位、546-552 位、556-565 位、587-604 和 486-598 位 ;Seq ID No 182 的第 4-15 位、17-32 位、40-47 位、67-78 位、90-98 位、101-107 位、111-136 位、161-171 位、184-198 位、208-214 位、234-245 位、247-254 位、272-279 位、288-298 位、303-310 位、315-320 位、327-333 位、338-349 位、364-374 和 378-396 位 ;Seq ID No 183 的第 5-27 位、33-49 位、51-57 位、74-81 位、95-107 位、130-137 位、148-157 位、173-184 和 75-235 位 ;Seq ID No 184 的第 6-23 位、47-53 位、57-63 位、75-82 位、97-105 位、113-122 位、124-134 位、142-153 位、159-164 位、169-179 位、181-187 位、192-208 位、215-243 位、247-257 位、285-290 位、303-310 和 30-51 位 ;Seq ID No 185 的第 17-29 位、44-52 位、59-73 位、77-83 位、86-92 位、97-110 位、118-153 位、156-166 位、173-179 位、192-209 位、225-231 位、234-240 位、245-251 位、260-268 位、274-279 位、297-306 位、328-340 位、353-360 位、369-382 位、384-397 位、414-423 位、431-436 位、452-465 位、492-498 位、500-508 位、516-552 位、554-560 位、568-574 位、580-586 位、609-617 位、620-626 位、641-647 和 208-219 位 ;Seq ID No 186 的第 4-26 位、32-45 位、58-72 位、111-119 位、137-143 位、146-159 位、187-193 位、221-231 位、235-242 位、250-273 位、290-304 位、311-321 位、326-339 位、341-347 位、354-368 位、397-403 位、412-419 位、426-432 位、487-506 位、580-592 位、619-628 位、663-685 位、707-716 位、743-751 位、

770-776 位、787-792 位、850-859 位、866-873 位、882-888 位、922-931 位、957-963 位、975-981 位、983-989 位、1000-1008 位、1023-1029 位、1058-1064 位、1089-1099 位、1107-1114 位、1139-1145 位、1147-1156 位、1217-1226 位、1276-1281 位、1329-1335 位、1355-1366 位、1382-1394 位、1410-1416 位、1418-1424 位、1443-1451 位、1461-1469 位、1483-1489 位、1491-1501 位、1515-1522 位、1538-1544 位、1549-1561 位、1587-1593 位、1603-1613 位、1625-1630 位、1636-1641 位、1684-1690 位、1706-1723 位、1765-1771 位、1787-1804 位、1850-1857 位、1863-1894 位、1897-1910 位、1926-1935 位、1937-1943 位、1960-1983 位、1991-2005 位、2008-2014 位、2018-2039 和 396-533 位、1342-1502 位、1672-1920 位；Seq ID No 187 的第 4-25 位、45-50 位、53-65 位、79-85 位、87-92 位、99-109 位、126-137 位、141-148 位、156-183 位、190-203 位、212-217 位、221-228 位、235-242 位、247-277 位、287-293 位、300-319 位、321-330 位、341-361 位、378-389 位、394-406 位、437-449 位、455-461 位、472-478 位、482-491 位、507-522 位、544-554 位、576-582 位、587-593 位、611-621 位、626-632 位、649-661 位、679-685 位、696-704 位、706-716 位、726-736 位、740-751 位、759-766 位、786-792 位、797-802 位、810-822 位、824-832 位、843-852 位、863-869 位、874-879 位、882-905 和 1-113 位、210-232 位、250-423 位、536-564 位；Seq ID No 188 的第 4-16 位、33-39 位、43-49 位、54-85 位、107-123 位、131-147 位、157-169 位、177-187 位、198-209 位、220-230 位、238-248 位、277-286 位、293-301 位、303-315 位、319-379 位、383-393 位、402-414 位、426-432 位、439-449 位、470-478 位、483-497 位、502-535 位、552-566 位、571-582 位、596-601 位、608-620 位、631-643 位、651-656 位、663-678 位、680-699 位、705-717 位、724-732 位、738-748 位、756-763 位、766-772 位、776-791 位、796-810 位、819-827 位、829-841 位、847-861 位、866-871 位、876-882 位、887-894 位、909-934 位、941-947 位、957-969 位、986-994 位、998-1028 位、1033-1070 位、1073-1080 位、1090-1096 位、1098-1132 位、1134-1159 位、1164-1172 位、1174-1201 和 617-635 位；Seq ID No 189 的第 7-25 位、30-40 位、42-64 位、70-77 位、85-118 位、120-166 位、169-199 位、202-213 位、222-244 和 190-203 位；Seq ID No 190 的第 4-11 位、15-53 位、55-93 位、95-113 位、120-159 位、164-200 位、210-243 位、250-258 位、261-283 位、298-319 位、327-340 位、356-366 位、369-376 位、380-386 位、394-406 位、409-421 位、425-435 位、442-454 位、461-472 位、480-490 位、494-505 位、507-514 位、521-527 位、533-544 位、566-574 和 385-398 位；Seq ID No 191 的第 5-36 位、66-72 位、120-127 位、146-152 位、159-168 位、172-184 位、205-210 位、221-232 位、234-243 位、251-275 位、295-305 位、325-332 位、367-373 位、470-479 位、482-487 位、520-548 位、592-600 位、605-615 位、627-642 位、655-662 位、664-698 位、718-725 位、734-763 位、776-784 位、798-809 位、811-842 位、845-852 位、867-872 位、879-888 位、900-928 位、933-940 位、972-977 位、982-1003 和 12-190 位、276-283 位、666-806 位；Seq ID No 192 的第 4-38 位、63-68 位、100-114 位、160-173 位、183-192 位、195-210 位、212-219 位、221-238 位、240-256 位、258-266 位、274-290 位、301-311 位、313-319 位、332-341 位、357-363 位、395-401 位、405-410 位、420-426 位、435-450 位、453-461 位、468-475 位、491-498 位、510-518 位、529-537 位、545-552 位、585-592 位、602-611 位、634-639 位、650-664 和 30-80 位、89-105 位、111-151 位；Seq ID No 193 的第 7-29 位、31-39 位、47-54 位、63-74 位、81-94

位、97-117 位、122-127 位、146-157 位、168-192 位、195-204 位、216-240 位、251-259 和 195-203 位 ;Seq ID No 194 的第 5-16 位、28-34 位、46-65 位、79-94 位、98-105 位、107-113 位、120-134 位、147-158 位、163-172 位、180-186 位、226-233 位、237-251 位、253-259 位、275-285 位、287-294 位、302-308 位、315-321 位、334-344 位、360-371 位、399-412 位、420-426 和 32-50 位 ;Seq ID No 195 的第 8-20 位、30-36 位、71-79 位、90-96 位、106-117 位、125-138 位、141-147 位、166-174 和 75-90 位 ;Seq ID No 196 的第 4-13 位、15-33 位、43-52 位、63-85 位、98-114 位、131-139 位、146-174 位、186-192 位、198-206 位、227-233 和 69-88 位 ;Seq ID No 197 的第 4-22 位、29-35 位、59-68 位、153-170 位、213-219 位、224-238 位、240-246 位、263-270 位、285-292 位、301-321 位、327-346 位、356-371 位、389-405 位、411-418 位、421-427 位、430-437 位、450-467 位、472-477 位、482-487 位、513-518 位、531-538 位、569-576 位、606-614 位、637-657 位、662-667 位、673-690 位、743-753 位、760-767 位、770-777 位、786-802 和 96-230 位、361-491 位、572-585 位 ;Seq ID No 198 的第 4-12 位、21-36 位、48-55 位、74-82 位、121-127 位、195-203 位、207-228 位、247-262 位、269-278 位、280-289 和 102-210 位 ;Seq ID No 199 的第 13-20 位、23-31 位、38-44 位、78-107 位、110-118 位、122-144 位、151-164 位、176-182 位、190-198 位、209-216 位、219-243 位、251-256 位、289-304 位、306-313 和 240-248 位 ;Seq ID No 200 的第 5-26 位、34-48 位、57-77 位、84-102 位、116-132 位、139-145 位、150-162 位、165-173 位、176-187 位、192-205 位、216-221 位、234-248 位、250-260 和 182-198 位 ;Seq ID No 201 的第 10-19 位、26-44 位、53-62 位、69-87 位、90-96 位、121-127 位、141-146 位、148-158 位、175-193 位、204-259 位、307-313 位、334-348 位、360-365 位、370-401 位、411-439 位、441-450 位、455-462 位、467-472 位、488-504 和 41-56 位 ;Seq ID No 202 的第 5-21 位、36-42 位、96-116 位、123-130 位、138-144 位、146-157 位、184-201 位、213-228 位、252-259 位、277-297 位、308-313 位、318-323 位、327-333 和 202-217 位 ;Seq ID No 203 的第 6-26 位、33-51 位、72-90 位、97-131 位、147-154 位、164-171 位、187-216 位、231-236 位、260-269 位、275-283 和 1-127 位 ;Seq ID No 204 的第 4-22 位、24-38 位、44-58 位、72-88 位、99-108 位、110-117 位、123-129 位、131-137 位、142-147 位、167-178 位、181-190 位、206-214 位、217-223 位、271-282 位、290-305 位、320-327 位、329-336 位、343-352 位、354-364 位、396-402 位、425-434 位、451-456 位、471-477 位、485-491 位、515-541 位、544-583 位、595-609 位、611-626 位、644-656 位、660-681 位、683-691 位、695-718 和 297-458 位 ;Seq ID No 205 的第 5-43 位、92-102 位、107-116 位、120-130 位、137-144 位、155-163 位、169-174 位、193-213 和 24-135 位 ;Seq ID No 206 的第 4-25 位、61-69 位、73-85 位、88-95 位、97-109 位、111-130 位、135-147 位、150-157 位、159-179 位、182-201 位、206-212 位、224-248 位、253-260 位、287-295 位、314-331 位、338-344 位、365-376 位、396-405 位、413-422 位、424-430 位、432-449 位、478-485 位、487-494 位、503-517 位、522-536 位、544-560 位、564-578 位、585-590 位、597-613 位、615-623 位、629-636 位、640-649 位、662-671 位、713-721 和 176-330 位 ;Seq ID No 207 的第 31-37 位、41-52 位、58-79 位、82-105 位、133-179 位、184-193 位、199-205 位、209-226 位、256-277 位、281-295 位、297-314 位、322-328 位、331-337 位、359-367 位、379-395 位、403-409 位、417-432 位、442-447 位、451-460 位、466-472 和 46-62 位、296-341 位 ;SeqID No 208 的第 23-29 位、56-63 位、67-74

位、96-108 位、122-132 位、139-146 位、152-159 位、167-178 位、189-196 位、214-231 位、247-265 位、274-293 位、301-309 位、326-332 位、356-363 位、378-395 位、406-412 位、436-442 位、445-451 位、465-479 位、487-501 位、528-555 位、567-581 位、583-599 位、610-617 位、622-629 位、638-662 位、681-686 位、694-700 位、711-716 和 667-684 位 ;Seq ID No 209 的第 20-51 位、53-59 位、109-115 位、140-154 位、185-191 位、201-209 位、212-218 位、234-243 位、253-263 位、277-290 位、303-313 位、327-337 位、342-349 位、374-382 位、394-410 位、436-442 位、464-477 位、486-499 位、521-530 位、536-550 位、560-566 位、569-583 位、652-672 位、680-686 位、698-704 位、718-746 位、758-770 位、774-788 位、802-827 位、835-842 位、861-869 和 258-416 位 ;Seq ID No 210 的第 7-25 位、39-45 位、59-70 位、92-108 位、116-127 位、161-168 位、202-211 位、217-227 位、229-239 位、254-262 位、271-278 位、291-300 和 278-295 位 ;Seq ID No 211 的第 4-20 位、27-33 位、45-51 位、53-62 位、66-74 位、81-88 位、98-111 位、124-130 位、136-144 位、156-179 位、183-191 和 183-195 位 ;Seq ID No 212 的第 12-24 位、27-33 位、43-49 位、55-71 位、77-85 位、122-131 位、168-177 位、179-203 位、209-214 位、226-241 和 63-238 位 ;Seq ID No 213 的第 4-19 位、37-50 位、120-126 位、131-137 位、139-162 位、177-195 位、200-209 位、211-218 位、233-256 位、260-268 位、271-283 位、288-308 和 1-141 位 ;Seq ID No 214 的第 11-17 位、40-47 位、57-63 位、96-124 位、141-162 位、170-207 位、223-235 位、241-265 位、271-277 位、281-300 位、312-318 位、327-333 位、373-379 和 231-368 位 ;Seq ID No 215 的第 9-33 位、41-48 位、57-79 位、97-103 位、113-138 位、146-157 位、165-186 位、195-201 位、209-215 位、223-229 位、237-247 位、277-286 位、290-297 位、328-342 和 247-260 位 ;Seq ID No 216 的第 7-15 位、39-45 位、58-64 位、79-84 位、97-127 位、130-141 位、163-176 位、195-203 位、216-225 位、235-247 位、254-264 位、271-279 和 64-72 位 ;Seq ID No 217 的第 4-12 位、26-42 位、46-65 位、73-80 位、82-94 位、116-125 位、135-146 位、167-173 位、183-190 位、232-271 位、274-282 位、300-306 位、320-343 位、351-362 位、373-383 位、385-391 位、402-409 位、414-426 位、434-455 位、460-466 位、473-481 位、485-503 位、519-525 位、533-542 位、554-565 位、599-624 位、645-651 位、675-693 位、717-725 位、751-758 位、767-785 位、792-797 位、801-809 位、819-825 位、831-836 位、859-869 位、890-897 和 222-362 位、756-896 位 ;Seq ID No 218 的第 11-17 位、22-28 位、52-69 位、73-83 位、86-97 位、123-148 位、150-164 位、166-177 位、179-186 位、188-199 位、219-225 位、229-243 位、250-255 和 153-170 位 ;SeqID No 219 的第 4-61 位、71-80 位、83-90 位、92-128 位、133-153 位、167-182 位、184-192 位、198-212 和 56-73 位 ;Seq ID No 220 的第 4-19 位、26-37 位、45-52 位、58-66 位、71-77 位、84-92 位、94-101 位、107-118 位、120-133 位、156-168 位、170-179 位、208-216 位、228-238 位、253-273 位、280-296 位、303-317 位、326-334 和 298-312 位 ;SeqID No 221 的第 7-13 位、27-35 位、38-56 位、85-108 位、113-121 位、123-160 位、163-169 位、172-183 位、188-200 位、206-211 位、219-238 位、247-254 和 141-157 位 ;Seq ID No 222 的第 23-39 位、45-73 位、86-103 位、107-115 位、125-132 位、137-146 位、148-158 位、160-168 位、172-179 位、185-192 位、200-207 位、210-224 位、233-239 位、246-255 位、285-334 位、338-352 位、355-379 位、383-389 位、408-417 位、423-429 位、446-456 位、460-473 位、478-503 位、522-540 位、553-562 位、568-577 位、

596-602 位、620-636 位、640-649 位、655-663 和 433-440 位、572-593 位 ;Seq ID No 223 的第 4-42 位、46-58 位、64-76 位、118-124 位、130-137 位、148-156 位、164-169 位、175-182 位、187-194 位、203-218 位、220-227 位、241-246 位、254-259 位、264-270 位、275-289 位、296-305 位、309-314 位、322-334 位、342-354 位、398-405 位、419-426 位、432-443 位、462-475 位、522-530 位、552-567 位、593-607 位、618-634 位、636-647 位、653-658 位、662-670 位、681-695 位、698-707 位、709-720 位、732-742 位、767-792 位、794-822 位、828-842 位、851-866 位、881-890 位、895-903 位、928-934 位、940-963 位、978-986 位、1003-1025 位、1027-1043 位、1058-1075 位、1080-1087 位、1095-1109 位、1116-1122 位、1133-1138 位、1168-1174 位、1179-1186 位、1207-1214 位、1248-1267 和 17-319 位、417-563 位 ;Seq ID No 224 的第 6-19 位、23-33 位、129-138 位、140-150 位、153-184 位、190-198 位、206-219 位、235-245 位、267-275 位、284-289 位、303-310 位、322-328 位、354-404 位、407-413 位、423-446 位、453-462 位、467-481 位、491-500 和 46-187 位 ;Seq ID No 225 的第 4-34 位、39-57 位、78-86 位、106-116 位、141-151 位、156-162 位、165-172 位、213-237 位、252-260 位、262-268 位、272-279 位、296-307 位、332-338 位、397-403 位、406-416 位、431-446 位、448-453 位、464-470 位、503-515 位、519-525 位、534-540 位、551-563 位、578-593 位、646-668 位、693-699 位、703-719 位、738-744 位、748-759 位、771-777 位、807-813 位、840-847 位、870-876 位、897-903 位、910-925 位、967-976 位、979-992 和 21-244 位、381-499 位、818-959 位 ;Seq ID No 226 的第 19-29 位、65-75 位、90-109 位、111-137 位、155-165 位、169-175 和 118-136 位 ;Seq ID No 227 的第 15-20 位、30-36 位、55-63 位、73-79 位、90-117 位、120-127 位、136-149 位、166-188 位、195-203 位、211-223 位、242-255 位、264-269 位、281-287 位、325-330 位、334-341 位、348-366 位、395-408 位、423-429 位、436-444 位、452-465 和 147-155 位 ;Seq ID No 228 的第 11-18 位、21-53 位、77-83 位、91-98 位、109-119 位、142-163 位、173-181 位、193-208 位、216-227 位、238-255 位、261-268 位、274-286 位、290-297 位、308-315 位、326-332 位、352-359 位、377-395 位、399-406 位、418-426 位、428-438 位、442-448 位、458-465 位、473-482 位、488-499 位、514-524 位、543-553 位、564-600 位、623-632 位、647-654 位、660-669 位、672-678 位、710-723 位、739-749 位、787-793 位、820-828 位、838-860 位、889-895 位、901-907 位、924-939 位、956-962 位、969-976 位、991-999 位、1012-1018 位、1024-1029 位、1035-1072 位、1078-1091 位、1142-1161 和 74-438 位 ;Seq ID No 229 的第 4-31 位、41-52 位、58-63 位、65-73 位、83-88 位、102-117 位、123-130 位、150-172 位、177-195 位、207-217 位、222-235 位、247-253 位、295-305 位、315-328 位、335-342 位、359-365 位、389-394 位、404-413 和 156-420 位 ;Seq ID No 230 的第 4-42 位、56-69 位、98-108 位、120-125 位、210-216 位、225-231 位、276-285 位、304-310 位、313-318 位、322-343 和 79-348 位 ;Seq ID No 231 的第 12-21 位、24-30 位、42-50 位、61-67 位、69-85 位、90-97 位、110-143 位、155-168 和 53-70 位 ;Seq ID No 232 的第 4-26 位、41-54 位、71-78 位、88-96 位、116-127 位、140-149 位、151-158 位、161-175 位、190-196 位、201-208 位、220-226 位、240-247 位、266-281 位、298-305 位、308-318 位、321-329 位、344-353 位、370-378 位、384-405 位、418-426 位、429-442 位、457-463 位、494-505 位、514-522 和 183-341 位 ;Seq ID No 233 的第 4-27 位、69-77 位、79-101 位、117-123 位、126-142 位、155-161 位、171-186 位、200-206 位、213-231 位、233-244 位、

258-263 位、269-275 位、315-331 位、337-346 位、349-372 位、376-381 位、401-410 位、424-445 位、447-455 位、463-470 位、478-484 位、520-536 位、546-555 位、558-569 位、580-597 位、603-618 位、628-638 位、648-660 位、668-683 位、717-723 位、765-771 位、781-788 位、792-806 位、812-822 和 92-231 位、618-757 位 ;Seq ID No 234 的第 11-47 位、63-75 位、108-117 位、119-128 位、133-143 位、171-185 位、190-196 位、226-232 位、257-264 位、278-283 位、297-309 位、332-338 位、341-346 位、351-358 位、362-372 和 41-170 位 ;Seq ID No 235 的第 6-26 位、50-56 位、83-89 位、108-114 位、123-131 位、172-181 位、194-200 位、221-238 位、241-259 位、263-271 位、284-292 位、304-319 位、321-335 位、353-358 位、384-391 位、408-417 位、424-430 位、442-448 位、459-466 位、487-500 位、514-528 位、541-556 位、572-578 位、595-601 位、605-613 位、620-631 位、634-648 位、660-679 位、686-693 位、702-708 位、716-725 位、730-735 位、749-755 位、770-777 位、805-811 位、831-837 位、843-851 位、854-860 位、863-869 位、895-901 位、904-914 位、922-929 位、933-938 位、947-952 位、956-963 位、1000-1005 位、1008-1014 位、1021-1030 位、1131-1137 位、1154-1164 位、1166-1174 和 20-487 位、757-1153 位 ;Seq ID No 236 的第 10-34 位、67-78 位、131-146 位、160-175 位、189-194 位、201-214 位、239-250 位、265-271 位、296-305 和 26-74 位、91-100 位、105-303 位 ;Seq ID No 237 的第 9-15 位、19-32 位、109-122 位、143-150 位、171-180 位、186-191 位、209-217 位、223-229 位、260-273 位、302-315 位、340-346 位、353-359 位、377-383 位、389-406 位、420-426 位、460-480 和 10-223 位、231-251 位、264-297 位、312-336 位 ;Seq ID No 238 的第 5-28 位、76-81 位、180-195 位、203-209 位、211-219 位、227-234 位、242-252 位、271-282 位、317-325 位、350-356 位、358-364 位、394-400 位、405-413 位、417-424 位、430-436 位、443-449 位、462-482 位、488-498 位、503-509 位、525-537 和 22-344 位 ;Seq ID No 239 的第 5-28 位、42-54 位、77-83 位、86-93 位、98-104 位、120-127 位、145-159 位、166-176 位、181-187 位、189-197 位、213-218 位、230-237 位、263-271 位、285-291 位、299-305 位、326-346 位、368-375 位、390-395 和 1-151 位 ;Seq ID No 240 的第 6-34 位、48-55 位、58-64 位、84-101 位、121-127 位、143-149 位、153-159 位、163-170 位、173-181 位、216-225 位、227-240 位、248-254 位、275-290 位、349-364 位、375-410 位、412-418 位、432-438 位、445-451 位、465-475 位、488-496 位、505-515 位、558-564 位、571-579 位、585-595 位、604-613 位、626-643 位、652-659 位、677-686 位、688-696 位、702-709 位、731-747 位、777-795 位、820-828 位、836-842 位、845-856 位、863-868 位、874-882 位、900-909 位、926-943 位、961-976 位、980-986 位、992-998 位、1022-1034 位、1044-1074 位、1085-1096 位、1101-1112 位、1117-1123 位、1130-1147 位、1181-1187 位、1204-1211 位、1213-1223 位、1226-1239 位、1242-1249 位、1265-1271 位、1273-1293 位、1300-1308 位、1361-1367 位、1378-1384 位、1395-1406 位、1420-1428 位、1439-1446 位、1454-1460 位、1477-1487 位、1509-1520 位、1526-1536 位、1557-1574 位、1585-1596 位、1605-1617 位、1621-1627 位、1631-1637 位、1648-1654 位、1675-1689 位、1692-1698 位、1700-1706 位、1712-1719 位、1743-1756 和 91-263 位 ;Seq ID No 241 的第 4-16 位、75-90 位、101-136 位、138-144 位、158-164 位、171-177 位、191-201 位、214-222 位、231-241 位、284-290 位、297-305 位、311-321 位、330-339 位、352-369 位、378-385 位、403-412 位、414-422 位、428-435 位、457-473 位、503-521 位、546-554 位、

562-568 位、571-582 位、589-594 位、600-608 位、626-635 位、652-669 位、687-702 位、706-712 位、718-724 位、748-760 位、770-775 和 261-272 位 ;Seq ID No 242 的第 4-19 位、30-41 位、46-57 位、62-68 位、75-92 位、126-132 位、149-156 位、158-168 位、171-184 位、187-194 位、210-216 位、218-238 位、245-253 位、306-312 位、323-329 位、340-351 位、365-373 位、384-391 位、399-405 位、422-432 位、454-465 位、471-481 位、502-519 位、530-541 位、550-562 位、566-572 位、576-582 位、593-599 位、620-634 位、637-643 位、645-651 位、657-664 位、688-701 和 541-551 位 ;Seq ID No 243 的第 6-11 位、17-25 位、53-58 位、80-86 位、91-99 位、101-113 位、123-131 位、162-169 位、181-188 位、199-231 位、245-252 和 84-254 位 ;Seq ID No 244 的第 13-30 位、71-120 位、125-137 位、139-145 位、184-199 和 61-78 位 ;Seq ID No245 的第 9-30 位、38-53 位、63-70 位、74-97 位、103-150 位、158-175 位、183-217 位、225-253 位、260-268 位、272-286 位、290-341 位、352-428 位、434-450 位、453-460 位、469-478 位、513-525 位、527-534 位、554-563 位、586-600 位、602-610 位、624-640 位、656-684 位、707-729 位、735-749 位、757-763 位、766-772 位、779-788 位、799-805 位、807-815 位、819-826 位、831-855 和 568-580 位 ;Seq ID No 246 的第 11-21 位、29-38 和 5-17 位 ;Seq ID No 247 的第 2-9 位 ;Seq ID No248 的第 4-10 位、16-28 和 7-18 位、26-34 位 ;Seq ID No 249 的第 10-16 和 1-15 位 ;Seq ID No 250 的第 4-11 位 ;Seq ID No 251 的第 4-40 位、42-51 和 37-53 位 ;Seq ID No 252 的第 4-21 和 22-29 位 ;Seq ID No 253 的第 2-11 位 ;9-17 位、Seq ID No 254 的第 32-44 和 1-22 位 ;Seq ID No 255 的第 19-25 位、27-32 和 15-34 位 ;Seq ID No 256 的第 4-12 位、15-22 和 11-33 位 ;Seq ID No 257 的第 10-17 位、24-30 位、39-46 位、51-70 和 51-61 位 ;Seq ID No 258 的第 6-19 位 ;Seq ID No 259 的第 6-11 位、21-27 位、31-54 和 11-29 位 ;Seq ID No 260 的第 4-10 位、13-45 和 11-35 位 ;Seq ID No 261 的第 4-14 位、23-32 和 11-35 位 ;Seq ID No 262 的第 14-39 位、45-51 和 15-29 位 ;Seq ID No 263 的第 4-11 位、14-28 和 4-17 位 ;Seq ID No 264 的第 4-16 和 2-16 位 ;Seq ID No 265 的第 4-10 位、12-19 位、39-50 和 6-22 位 ;Seq ID No 266 的第 2-13 位 ;Seq ID No 267 的第 4-11 位、22-65 和 3-19 位 ;Seq ID No 268 的第 17-23 位、30-35 位、39-46 位、57-62 和 30-49 位 ;Seq ID No 269 的第 4-19 和 14-22 位 ;Seq ID No270 的第 2-9 位 ;Seq ID No 271 的第 7-18 位、30-43 和 4-12 位 ;Seq IDNo 272 的第 4-30 位、39-47 和 5-22 位 ;Seq ID No 273 的第 6-15 和 14-29 位 ;Seq ID No 274 的第 4-34 和 23-35 位 ;Seq ID No 275 的第 4-36 位、44-57 位、65-72 和 14-27 位 ;Seq ID No 276 的第 4-18 和 11-20 位 ;Seq IDNo 277 的第 5-19 位 ;Seq ID No 278 的第 18-36 和 6-20 位 ;Seq ID No 279 的第 4-10 位、19-34 位、41-84 位、96-104 和 50-63 位 ;Seq ID No 280 的第 4-9 位、19-27 和 8-21 位 ;Seq ID No 281 的第 4-16 位、18-28 和 22-30 位 ;Seq ID No 282 的第 4-15 和 21-35 位 ;Seq ID No 283 的第 4-17 和 3-13 位 ;Seq ID No 284 的第 4-12 和 4-18 位 ;Seq ID No 285 的第 4-24 位、31-36 和 29-45 位 ;Seq ID No 286 的第 12-22 位、34-49 和 21-32 位 ;Seq ID No 287 的第 4-17 和 22-32 位 ;Seq ID No 288 的第 4-16 位、25-42 和 7-28 位 ;Seq ID No 289 的第 4-10 和 7-20 位 ;Seq ID No 290 的第 4-11 位、16-36 位、39-54 和 28-44 位 ;Seq ID No 291 的第 5-20 位、29-54 和 14-29 位 ;Seq ID No 292 的第 24-33 和 10-22o 位 ;Seq ID No 293 的第 10-51 位、54-61 和 43-64 位 ;Seq ID No 294 的第 7-13 和 2-17 位 ;Seq IDNo 295 的第 11-20 和 6-20 位 ;Seq ID No 296 的第 4-30 位、

34-41 和 19-28 位 ;Seq ID No 297 的第 11-21 位 ;Seq ID No 298 的第 4-16 位、21-26 和 9-38 位 ;Seq ID No 299 的第 4-12 位、15-27 位、30-42 位、66-72 和 10-24 位 ;Seq ID No 300 的第 8-17 和 11-20 位 ;以及 Seq ID No 246 的第 2-19 位 ;Seq ID No 247 的第 1-12 位 ;Seq ID No 248 的第 21-38 位 ;Seq ID No 254 的第 2-22 位 ;Seq ID No 255 的第 15-33 位 ;Seq ID No 256 的第 11-32 位 ;Seq ID No 259 的第 11-28 位 ;Seq ID No 260 的第 10-27 位 ;Seq ID No 261 的第 9-26 位 ;Seq ID No 263 的第 4-16 位 ;Seq ID No 266 的第 1-18 位 ;Seq ID No 273 的第 12-29 位 ;Seq ID No 276 的第 6-23 位 ;Seq ID No 277 的第 1-21 位 ;Seq ID No 279 的第 47-64 位 ;Seq ID No 285 的第 28-45 位 ;Seq ID No 287 的第 18-35 位 ;Seq ID No 291 的第 14-31 位 ;Seq ID No 292 的第 7-24 位 ;Seq ID No 299 的第 8-25 位 ;Seq ID No 300 的第 1-20 位 ;Seq ID No 151 的第 18-33 位 ;Seq ID No 151 的第 62-72 位 ;Seq ID No 152 的第 118-131 位 ;Seq ID No 154 的第 195-220 位 ;Seq ID No 154 的第 215-240 位 ;Seq ID No 154 的第 255-280 位 ;Seq ID No 155 的第 72-81 位 ;Seq ID No 156 的第 174-186 位 ;Seq ID No 157 的第 317-331 位 ;Seq ID No 158 的第 35-59 位 ;Seq ID No 158 的第 54-84 位 ;Seq ID No 158 的第 79-104 位 ;Seq ID No 159 的第 33-58 位 ;Seq ID No 159 的第 81-101 位 ;Seq ID No 159 的第 136-150 位 ;Seq ID No 159 的第 173-186 位 ;Seq ID No 159 的第 231-251 位 ;Seq ID No 161 的第 22-48 位 ;Seq ID No 162 的第 24-39 位 ;Seq ID No 163 的第 475-489 位 ;Seq ID No 164 的第 38-56 位 ;Seq ID No 164 的第 583-604 位 ;Seq ID No 165 的第 202-223 位 ;Seq ID No 165 的第 222-247 位 ;Seq ID No 165 的第 242-267 位 ;Seq ID No 165 的第 262-287 位 ;Seq ID No 165 的第 282-307 位 ;Seq ID No 165 的第 302-327 位 ;Seq ID No 166 的第 25-48 位 ;Seq ID No 167 的第 204-217 位 ;Seq ID No 168 的第 259-276 位 ;Seq ID No 169 的第 121-139 位 ;Seq ID No 169 的第 260-267 位 ;Seq ID No 169 的第 215-240 位 ;Seq ID No 170 的第 115-140 位 ;Seq ID No 172 的第 182-204 位 ;Seq ID No 173 的第 144-153 位 ;Seq ID No 173 的第 205-219 位 ;Seq ID No 174 的第 196-206 位 ;Seq ID No 174 的第 240-249 位 ;Seq ID No 174 的第 272-287 位 ;Seq ID No 174 的第 199-223 位 ;Seq ID No 174 的第 218-237 位 ;Seq ID No 175 的第 226-249 位 ;Seq ID No 175 的第 287-306 位 ;Seq ID No 176 的第 430-449 位 ;Seq ID No 177 的第 61-375 位 ;Seq ID No 178 的第 241-260 位 ;Seq ID No 181 的第 483-502 位 ;Seq ID No 182 的第 379-396 位 ;Seq ID No 184 的第 31-51 位 ;Seq ID No 186 的第 1436-1460 位 ;Seq ID No 186 的第 1455-1474 位 ;Seq ID No 186 的第 1469-1487 位 ;Seq ID No 187 的第 215-229 位 ;Seq ID No 187 的第 534-561 位 ;Seq ID No 187 的第 59-84 位 ;Seq ID No 187 的第 79-104 位 ;Seq ID No 188 的第 618-635 位 ;Seq ID No 189 的第 191-203 位 ;Seq ID No 190 的第 386-398 位 ;Seq ID No 191 的第 65-83 位 ;Seq ID No 192 的第 90-105 位 ;Seq ID No 192 的第 112-136 位 ;Seq ID No 193 的第 290-209 位 ;Seq ID No 194 的第 33-50 位 ;Seq ID No 195 的第 76-90 位 ;Seq ID No 196 的第 70-88 位 ;Seq ID No 197 的第 418-442 位 ;Seq ID No 197 的第 574-585 位 ;Seq ID No 198 的第 87-104 位 ;Seq ID No 198 的第 124-148 位 ;Seq ID No 198 的第 141-152 位 ;Seq ID No 199 的第 241-248 位 ;Seq ID No 200 的第 183-198 位 ;Seq ID No 200 的第 40-57 位 ;Seq ID No 202 的第 202-217 位 ;Seq ID No 203 的第 50-74 位 ;Seq ID No 203 的第 69-93 位 ;Seq ID No 203 的第 88-112 位 ;Seq ID No 203 的第 107-127 位 ;Seq ID No 205 的第 74-92 位 ;Seq ID No 206 的第

207-232 位 ;Seq ID No 206 的第 227-252 位 ;Seq ID No 206 的第 247-272 位 ;SeqID No 207 的第 47-60 位 ;Seq ID No 207 的第 297-305 位 ;Seq ID No 207 的第 312-337 位 ;Seq ID No 208 的第 667-384 位 ;Seq ID No 210 的第 279-295 位 ;Seq ID No 211 的第 179-198 位 ;Seq ID No 213 的第 27-51 位 ;Seq ID No 213 的第 46-70 位 ;Seq ID No 213 的第 65-89 位 ;Seq ID No 213 的第 84-108 位 ;Seq ID No 213 的第 112-141 位 ;Seq ID No 215 的第 248-260 位 ;Seq ID No 216 的第 59-78 位 ;Seq ID No 218 的第 154-170 位 ;Seq ID No 219 的第 57-73 位 ;Seq ID No 220 的第 297-314 位 ;Seq ID No 221 的第 142-157 位 ;Seq ID No 222 的第 428-447 位 ;Seq ID No 222 的第 573-593 位 ;Seq ID No 223 的第 523-544 位 ;Seq ID No 223 的第 46-70 位 ;Seq ID No 223 的第 65-89 位 ;Seq ID No 223 的第 84-108 位 ;Seq ID No 223 的第 122-151 位 ;Seq ID No 224 的第 123-142 位 ;Seq ID No 225 的第 903-921 位 ;Seq ID No 226 的第 119-136 位 ;Seq ID No 227 的第 142-161 位 ;Seq ID No 228 的第 258-277 位 ;Seq ID No 228 的第 272-300 位 ;SeqID No 228 的第 295-322 位 ;Seq ID No 229 的第 311-343 位 ;Seq ID No 229 的第 278-304 位 ;Seq ID No 230 的第 131-150 位 ;Seq ID No 230 的第 195-218 位 ;Seq ID No 231 的第 53-70 位 ;Seq ID No 232 的第 184-208 位 ;Seq ID No 232 的第 222-246 位 ;Seq ID No 232 的第 241-265 位 ;Seq ID No 232 的第 260-284 位 ;Seq ID No 232 的第 279-303 位 ;Seq ID No 232 的第 317-341 位 ;Seq ID No 233 的第 678-696 位 ;Seq ID No 235 的第 88-114 位 ;Seq ID No 235 的第 464-481 位 ;Seq ID No 236 的第 153-172 位 ;SeqID No 236 的第 137-155 位、166-184 位 ;Seq ID No 236 的第 215-228 位 ;Seq ID No 237 的第 37-51 位 ;Seq ID No 237 的第 53-75 位 ;Seq ID No 237 的第 232-251 位 ;Seq ID No 237 的第 318-336 位 ;Seq ID No 238 的第 305-315 位 ;Seq ID No 238 的第 131-156 位 ;Seq ID No 241 的第 258-275 位 ;Seq ID No 243 的第 107-137 位 ;Seq ID No 243 的第 138-162 位 ;SeqID No 243 的第 157-181 位 ;Seq ID No 243 的第 195-227 位 ;Seq ID No 244 的第 62-78 位 ;Seq ID No 245 的第 567-584 位,以及含所述序列的至少 6 个、优选 8 个以上、尤其优选 10 个以上氨基酸的片段。所有这些片段个别的且各自独立的形成本发明的优选方面。

[0184] 用一组具有至少 10 个氨基酸长且有 1 个氨基酸重叠的肽通过分析蛋白质抗原的完整序列可鉴定某一具体抗原的所有线性超免疫血清反应性片段。随后,用已表达的全长蛋白质或其结构域多肽通过分析与超免疫血清在一起的蛋白质抗原可鉴定非线性表位。假如某一蛋白质的单独结构域足以形成不依赖于其天然蛋白质的 3D 结构,分析与超免疫血清在一起的相应重组或合成产生的结构域多肽可鉴定多结构域蛋白质个别结构域内的构象表位。对于那些其中的结构域具有线性表位以及构象表位的抗原而言,利用对应于线性表位的肽进行的竞争性实验可用于确认构象表位的存在。

[0185] 其中,应理解本发明还尤其涉及编码上述片段的核酸分子、与编码所述片段的核酸分子杂交的核酸分子、尤其是那些在严谨条件下杂交的核酸分子以及用于扩增编码所述片段的核酸分子的核酸分子,诸如 PCR 引物。在这些方面,如上所述,优选的核酸分子是与优选片段相应的那些核酸分子。

[0186] 本发明还涉及含有本发明的一种或多种核酸分子的载体、含本发明载体的遗传改造过的宿主细胞和用重组技术生产超免疫血清反应性抗原及其片段。

[0187] 多种表达载体可用于表达按照本发明的超免疫血清反应性抗原或其片段。通常,

在宿主中适于维持、繁殖或表达核酸从而表达多肽的任何载体都可用于这方面的表达。依照本方面的这一方面,载体可以是,例如,质粒载体、单链或双链噬菌体载体、单链或双链 RNA 或 DNA 病毒载体。本文所公开的起始质粒可购买获得、公开获得或可通过众所周知的已发表方法的常规应用构建自有用的质粒。在某些方面,载体中优选的是用于表达本发明的核酸分子和超免疫血清反应性抗原或其片段的那些载体。宿主细胞中的核酸构建体可以常规方式用于生产由所述重组序列编码的基因产物。或者,本发明的超免疫血清反应性抗原及其片段可通过传统的肽合成仪合成产生。成熟蛋白质可在适当启动子的调控下表达于哺乳动物细胞、酵母、细菌或其它细胞中。无细胞翻译体系还可通过来自本发明 DNA 构建体的 RNA 应用于生产所述蛋白质。

[0188] 宿主细胞可遗传改造以引入核酸分子并表达本发明的核酸分子。适当宿主的代表性例子包括细菌细胞,诸如链球菌、葡萄球菌、大肠杆菌、链霉菌和枯草芽孢杆菌细胞;真菌细胞,诸如酵母细胞和曲霉菌细胞;昆虫细胞,诸如果蝇 S2 和 *Spodoptera Sf9* 细胞;动物细胞,诸如 CHO、COS、HeLa、C127、3T3、BHK、293 和 Bowes 黑素瘤细胞;和植物细胞。

[0189] 本发明还提供了生产化脓链球菌超免疫血清反应性抗原及其片段的方法,包括从宿主细胞表达由本发明所提供核酸分子编码的超免疫血清反应性抗原或其片段。本发明另外提供了生产能表达化脓链球菌超免疫血清反应性抗原或其片段的细胞,包括用按照本发明的载体转化或转染合适的宿主细胞从而使被转化或转染的细胞表达由该载体中所含核酸编码的多肽。

[0190] 多肽可以修饰过的形式表达,诸如融合蛋白,而且不仅可包含分泌信号,还可包含附加的异源功能区。这样,例如,某一区段的附加氨基酸,尤其是带电荷氨基酸,可加入多肽的 N- 或 C- 末端以便在纯化期间或随后的操作和储存期间提高其在宿主细胞中的稳定性和持久性。此外,某些区段可加入多肽中以方便纯化。所述的区域可在该多肽最终制备步骤前去除。其中,在多肽中添加某部分肽产生分泌物或排泄物以提高稳定性或方便纯化是本领域中熟知和常规的技术。优选的融合蛋白包含来自免疫球蛋白且对稳定或纯化多肽有效的异源区段。例如,EP-A-0 464 533(加拿大副本 2045869)公开了含免疫球蛋白分子恒定区多个部分以及另一蛋白质或其部分的融合蛋白。在药物开发中,例如,为了高通量筛选检测的目的将蛋白质与抗体 Fc 部分融合以鉴定拮抗剂。参阅,例如,{Bennett, D. 等,1995} 和 {Johanson, K. 等,1995}。

[0191] 化脓链球菌超免疫血清反应性抗原或其片段可通过众所周知的方法回收和纯化自重组细胞培养物,包括硫酸铵或乙醇沉淀、酸提取、阴离子或阳离子交换层析、磷酸纤维素层析、疏水相互作用层析、羟磷灰石层析和外源凝集素层析。

[0192] 按照本发明的超免疫血清反应性抗原及其片段可通过化学合成以及生物技术手段产生。后者包括用含有按照本发明的核酸的载体转染或转化宿主细胞并在本领域技术人员已知的条件下培养已被转染或已被转化的宿主细胞。生产方法还可包括纯化步骤以纯化或分离待制备的多肽。在优选实施方案中的载体是按照本发明的载体。

[0193] 按照本发明的超免疫血清反应性抗原及其片段可用于检测样品中的某种或几种生物体,所述样品含这些生物体或其来源多肽。优选这样的检测是用于诊断的,更优选用于疾病的诊断,最优选用于诊断与革兰氏阳性细菌的存在或富集相关或有联系的疾病,尤其是包括链球菌、葡萄球菌和乳酸菌在内的细菌类型。更优选地,微生物可选自无乳链球菌

(*Streptococcusagalactiae*)、肺炎链球菌和变异链球菌 (*Streptococcus mutans*) 类型,尤其是所述微生物为化脓链球菌。

[0194] 本发明还涉及诸如定量测定等诊断检测法和用于检测细胞和组织中本发明超免疫血清反应性抗原及其片段水平的诊断检测中,包括正常和异常水平的测定。因而,例如,依照本发明用于检测较正常对照组织样品而言过表达的多肽的诊断检测法可用于检测感染的存在,并识别正被感染的生物体。可用于确定宿主来源样品中多肽水平的检测技术是本领域技术人员众所周知的。这样的检测方法包括放射免疫检测法、竞争性结合检测法、蛋白质印迹分析和 ELISA 检测法。其中,ELISA 是常常优选的。ELISA 检测最初包括制备特异于所述多肽的抗体,优选单克隆抗体。此外,通常制备与所述单克隆抗体结合的报道抗体。该报道抗体与可检测试剂附着,诸如放射性试剂、荧光试剂或酶促试剂,诸如辣根过氧化物酶。

[0195] 按照本发明的超免疫血清反应性抗原及其片段还可用于阵列的目的或与连同阵列使用。更具体而言,可将至少一个按照本发明的超免疫血清反应性抗原及其片段固定于支持物上。所述支持物通常包含各种超免疫血清反应性抗原及其片段,凭此通过使用一种或数种按照本发明的超免疫血清反应性抗原及其片段和 / 或不同的超免疫血清反应性抗原及其片段可建立多样性。所述阵列的特征性质,也是一般而言任何阵列的特征性质,是单独的多肽固定于所述支持物或其表面的单独的或预定区域或位置。因此,阵列的单独位置或区域的任一活性可能都与某一特殊的多肽相关。固定于支持物上的不同超免疫血清反应性抗原及其片段的数目可以从少到 10 个多到数千个不同的超免疫血清反应性抗原及其片段之间变化。在优选的实施方案中每 cm^2 的超免疫血清反应性抗原及其片段的密度少到每 cm^2 10 个肽 / 多肽至每 cm^2 至少 400 个不同的肽 / 多肽且更单独的是每 cm^2 至少 1000 个不同的超免疫血清反应性抗原及其片段。

[0196] 所述阵列的制备是本领域技术人员已知的,例如,参阅美国专利 5,744,309。阵列优选包含一个平坦、多孔或非多孔且具有至少第一个表面的固体支持物。如本文所公开的超免疫血清反应性抗原及其片段被固定于所述的表面上。其中,优选的支持物材料是玻璃或纤维素。用于本文所述任一诊断应用中的阵列也在本发明的范围内。除了按照本发明的超免疫血清反应性抗原及其片段之外,按照本发明的核酸分子也可用于上述阵列的产生中。本申请也适用于抗体构成的阵列,其中抗体优选此处所述的单克隆抗体。

[0197] 另一方面,本发明涉及针对按照本发明的任一超免疫血清反应性抗原及其片段、衍生物或其片段的抗体。本发明包括,例如,单克隆和多克隆抗体,嵌合的、单链的和人源化的抗体,以及 Fab 片段或 Fab 表达文库的产物。也属于本发明范围内的是抗体可能是嵌合的,即,其不同部分来自不同物种或至少各自的序列取自不同的物种。

[0198] 针对相应于本发明序列的超免疫血清反应性抗原及其片段产生的抗体可通过将超免疫血清反应性抗原及其片段直接注射入动物中或通过将超免疫血清反应性抗原及其片段施用于动物,优选非人动物而获得。由此获得的抗体然后与超免疫血清反应性抗原及其片段自身结合。如此,即使只编码超免疫血清反应性抗原及其片段的片段的序列也可用于产生结合完整天然超免疫血清反应性抗原及其片段的抗体。所述的抗体然后可用于从表达那些超免疫血清反应性抗原及其片段的组织中分离超免疫血清反应性抗原及其片段。

[0199] 为了制备单克隆抗体,本领域已知提供连续细胞系培养所产生抗体的任何技术都

可使用。（如，最初在 {Kohler, G. 等, 1975} 中所述的。）

[0200] 用于单链抗体产生所述的技术（美国专利文件 4, 946, 778）可适用于生产针对按照本发明超免疫血清反应性抗原及其片段的单链抗体。此外，转基因小鼠，或诸如其它哺乳动物等其它生物体可用于表达针对按照本发明的超免疫血清反应性抗原及其片段的人源化抗体。

[0201] 或者，噬菌体展示技术或核糖体展示可应用于选择对超免疫血清反应性抗原及其片段具结合活性的抗体基因，它们可来自为具有相应靶抗原而筛选的人淋巴细胞 PCR 扩增 v- 基因的全部组分或来自天然的文库 {McCafferty, J. 等, 1990} ; {Marks, J. 等, 1992}。这些抗体的亲合力也可通过链的改组得到提高 {Clackson, T. 等, 1991}。

[0202] 如果存在两个抗原结合结构域，各结构域可直接针对不同表位 - 被称为“双特异性的”抗体。

[0203] 上述抗体可应用于分离或鉴定表达超免疫血清反应性抗原及其片段的克隆或通过将抗体附着于固体支持物上进行亲和层析分离和 / 或纯化而纯化本发明的超免疫血清反应性抗原及其片段。

[0204] 因此，特别地，针对本发明超免疫血清反应性抗原及其片段的抗体可应用于抑制和 / 或治疗感染，尤其是细菌感染，更尤其是由化脓链球菌引起的感染。

[0205] 超免疫血清反应性抗原及其片段包括形成本领域特殊的一个方面的抗原性、表位性或免疫学等价衍生物。术语“抗原等价衍生物”用于此处时包含被确定抗体特异识别的超免疫血清反应性抗原及其片段或其等同物，在用按照本发明的蛋白质或超免疫血清反应性抗原及其片段接种免疫后，所述抗体干扰了病原体 and 哺乳动物宿主之间的相互作用。术语“免疫学等价衍生物”用于此处时包含肽或其等价物，当将它们用于适当制剂中以在脊椎动物中产生抗体时，所述抗体发挥了干扰病原体和哺乳动物宿主之间相互作用的功能。

[0206] 超免疫血清反应性抗原及其片段，诸如抗原性或免疫学等价衍生物或它们的融合蛋白可用作抗原来免疫小鼠或诸如大鼠或小鸡等其它动物。融合蛋白可使超免疫血清反应性抗原及其片段具有稳定性。抗原可以通过例如偶联与例如牛血清白蛋白 (BSA) 或匙孔血蓝蛋白 (KLH) 等免疫原性载体蛋白联合。或者，包含多拷贝所述蛋白质或超免疫血清反应性抗原及其片段或抗原性或免疫学等价的超免疫血清反应性抗原及其片段的抗原性肽对于提高免疫原性而言可能已具有足够的抗原性，从而避免了载体的使用。

[0207] 优选的对抗体或其衍生物进行修饰使其在个体中的免疫原性降低。例如，如果所述个体是人，则抗体可能最好被人源化，其中杂交瘤衍生抗体的互补性决定区被移植入人的单克隆抗体内，参阅 {Jones, P. 等, 1986} 或 {Tempest, P. 等, 1991}。

[0208] 本发明多核苷酸在遗传免疫中的应用优选采用合适的运送方法，诸如将质粒 DNA 直接注射入肌肉中、运送具特异蛋白质载体的 DNA 复合物、DNA 与磷酸钙共沉淀、将 DNA 封装入各种形式的脂质体中、微粒轰击 {Tang, D. 等, 1992}，{Eisenbraun, M. 等, 1993} 以及用克隆的逆转录病毒载体进行体内感染 {Seeger, C. 等, 1984}。

[0209] 另一方面，本发明涉及与按照本发明的任一超免疫血清反应性抗原及其片段结合的肽，以及制备所述肽的方法，由此该方法的特征在于按照本发明的超免疫血清反应性抗原及其片段的应用，而基本的步骤则是本领域技术人员所已知的。

[0210] 所述肽可依照本领域不同情形下利用不同的方法产生，诸如噬菌体展示或核糖体

展示。假使用噬菌体展示,基本上肽库是以噬菌体的形式产生的,且

[0211] 将此类文库与靶分子(在本情形中按照本发明的超免疫血清反应性抗原及其片段)接触。随后从各自的反应液中洗下与靶分子结合的肽,优选作为与靶分子在一起的复合物被洗下。结合特征,至少一定程度上,取决于具体实行的实验安排,诸如盐浓度等等,这是本领域技术人员所已知的。在将以较高的亲和力或较大的势能与靶分子结合的那些肽从文库中未结合的成员中分隔出来后,以及任选的同样在将靶分子从靶分子和肽的复合物中取出后,可随后对相应的肽进行定性。4 在定性前可任选的进行扩增,诸如,通过培养肽编码噬菌体。所述定性优选包括对靶结合肽进行测序。基本上,所述的肽不局限于其长度,不过,优选以相应的方法获得长度为大约 8-20 个氨基酸的肽。文库的大小可能在大约 10^2 - 10^{18} 种不同的肽,优选 10^8 - 10^{15} 种不同的肽,不过,并不局限于此。

[0212] 结合超免疫血清反应性抗原及其片段的具体形式靶分子是德国专利申请 DE 197 42 706 中所述的所谓“抗促成素”。

[0213] 另一方面,本发明涉及与按照本发明的任一超免疫血清反应性抗原及其片段相互作用的功能性核酸,以及制备所述功能性核酸的方法,因此该方法的特征在于按照本发明的超免疫血清反应性抗原及其片段的应用,而基本的步骤则是本领域技术人员所已知的。功能性核酸优选为适体和 spiegelmers。

[0214] 适体是单链或双链且特异与靶分子相互作用的 D- 核酸。适体的制备和选择参阅,例如欧洲专利 EP 0 533 838。基本上采用了以下步骤。首先,提供核酸的混合物,即,可能的适体,因此各核酸通常包含一区段数个、优选至少 8 个连续的随机核苷酸。此混合物随后与靶分子接触从而使所述核酸与靶分子结合,这是基于与候选混合物相比,该混合物对靶分子的亲和力有所增加或对其具有更大的势能。结合的核酸(一个或多个分子)随后从留下的混合物中分离出来。任选的,用例如聚合酶链式反应扩增由此获得的核酸。这些步骤可重复数次得到具有特异结合靶分子的较高比率核酸的混合物,然后任选的从中选择最终结合的核酸。这些特异结合的核酸被称为适体。显而易见在产生或鉴定单个核酸混合物的适体样品的方法的任一阶段都可用标准技术测定其序列。可通过例如引入本领域技术人员已知产生适体的预定化学基团使适体稳定,这也在本发明的范围内。这样的修饰可能,例如,位于核苷酸的糖基部分 2' 位置处引入一氨基。适体目前被用作治疗剂。不过,由此选择或产生的适体可用于靶分子的确认和 / 或用于先导物质被开发成药物,优选以小分子为基础的药物开发,这也在本发明的范围内。这事实上是通过竞争性试验完成的,其中靶分子和适体之间的相互作用被候选药物通过从靶分子和适体的复合物中置换出适体而被抑制,假设各个候选药物可特异的抑制靶分子和适体之间的相互作用,且如果相互作用是特异的,所述的候选药物就,至少在原则上,适于封闭靶分子并因此降低其在含所述靶分子的相应体系中的生物利用度或活性。由此获得的小分子可然后被进一步的衍生和修饰以使它物理、化学、生物学和 / 或医学特性,诸如毒性、特异性、生物降解性和生物利用度最优化。

[0215] Spiegelmer 及其产生或制备是基于相似的原理。spiegelmer 的制备参阅国际专利申请 WO 98/08856。Spiegelmer 是 L 型核酸,这意味着它们是由 L 型核苷酸而非象适体那样由 D 型核苷酸组成的。Spiegelmer 实际上的特征在于与适体相比,它们在生物学体系中具有非常高的稳定性,特异地与它们所针对的靶分子相互作用。在 spiegelmer 产生过程中,建立了一组异源 D- 核酸,将所述的那组异源 D 核酸与靶分子的旋光对映体接触,在本情

形中例如按照本发明的超免疫血清反应性抗原及其片段的天然存在 L 型对映体的 D 对映体。随后,分隔开不与靶分子的旋光对映体相互作用的那些 D 核酸。而那些与靶分子旋光对映体相互作用的那些 D 核酸被分离、任选的测定和 / 或测序并随后基于获自 D 核酸的核酸信息合成相应的 L 核酸。这些 L 核酸在序列形式上与跟靶分子的旋光对映体相互作用的上述 D 核酸相同,将特异的与天然存在的靶分子而非其旋光对映体相互作用。与适体产生方法类似,也可能数次重复多个步骤并因此富集与靶分子的旋光对映体特异相互作用的那些核酸。

[0216] 另一方面,本发明涉及与按照本发明的核酸分子相互作用的功能性核酸,以及用于制备所述功能性核酸的方法,其中该方法的特征在于使用了按照本发明的核酸分子及其相应序列且基本步骤是本领域技术人员所已知的。功能性核酸优选是核酶、反义寡核苷酸和 siRNA。

[0217] 核酶是优选由 RNA 组成且基本上包含两部分的催化活性核酸。第一部分显示出催化活性,而第二部分负责与靶核酸的特异相互作用,在本情形中靶核酸是编码按照本发明的超免疫血清反应性抗原及其片段的核酸。基于靶核酸与核酶第二部分之间的相互作用,通常是通过杂交和在两条杂交链上基本上互补的碱基伸展段的 Watson-Crick 碱基配对,催化活性部分可被活化,即,如果所述核酶的催化活性是磷酸二酯酶活性,那么它在分子内或分子间催化靶核酸。随后,靶核酸可能进一步降解,这最后导致了靶核酸以及来源与所述靶核酸的蛋白质的降解。核酶、其用途和设计原则是本领域技术人员所已知的,可参阅 {Doherty, E. 等, 2001} 和 {Lewin, A. 等, 2001}。

[0218] 用于药物制备及作为诊断剂的反义寡核苷酸的活性和设计分别基于相似的作用模式。基本上,反义寡核苷酸基于碱基互补与靶 RNA 杂交,优选与 mRNA 杂交,从而激活 RNase H。RNase H 被磷酸二酯和硫代磷酸酯二者偶联的 DNA 激活。不过,磷酸二酯偶联的 DNA 被细胞核酸酶快速降解,而硫代磷酸酯偶联的 DNA 则不会。这些抗性、非天然存在 DNA 衍生物通过与 RNA 杂交而不抑制 RNase H。换言之,反义多核苷酸只作为 DNA RNA 杂交复合物有效。其中,这类反义寡核苷酸的例子参阅美国专利 5,849,902 和 US 5,989,912。换言之,基于靶分子的核酸序列,即,在本案例中是针对按照本发明的超免疫血清反应性抗原及其片段的核酸分子,或来自相应核酸序列原则上可从中推导出来的靶蛋白,或通过同样的了解所述核酸序列,尤其是 mRNA,可基于碱基互补的原则设计适当的反义寡核苷酸。

[0219] 尤其优选的是具有硫代磷酸酯的短伸展段的反义寡核苷酸 (3-6 个碱基)。最短 3 个 DNA 碱基是细菌 RNase H 活化所需要的,而最短 5 个碱基是哺乳动物 RNase H 活化所需要的。在这些嵌合寡核苷酸中存在形成 RNase H 底物的中心区,两侧是由不会形成 RNase H 底物的修饰过的核苷酸组成的杂交“臂”。嵌合寡核苷酸可被诸如 2' -O- 甲基或 2' - 氟代等修饰。备选的方案是在所述的臂中使用磷酸甲酯或氨基磷酸酯键。在本发明实践中有效的反义寡核苷酸的别的实施方案是 P- 甲氧基寡核苷酸、部分的 P- 甲氧基寡脱氧核糖核苷酸或 P- 甲氧基寡核苷酸。

[0220] 对于本发明而言特别相关和有效的是如以上提及的两个美国专利中所更具体描述的那些反义寡核苷酸。这些寡核苷酸包含非天然存在的 5' → 3' - 连接核苷酸。更确切的说,寡核苷酸具有两种类型的核苷酸: 2' - 脱氧硫代磷酸酯,它激活 RNase H, 以及 2' - 修饰的核苷酸,它不激活 RNaseH。2' - 已修饰的核苷酸之间的键可以是磷酸二酯、硫

代磷酸酯或 P- 乙氧基磷酸二酯。RNase H 的活化是通过含 3-5 个 2' 脱氧硫代磷酸酯核苷酸的邻近 RNaseH- 活化区激活细菌的 RNase H 以及含 5-10 个 2' - 脱氧硫代磷酸酯核苷酸的邻近 RNase H- 活化区激活真核的、尤其是哺乳动物的 RNase H 而完成的。防止降解的保护作用是通过形成 5' 和 3' 末端碱基高度核酸酶抗性以及任选的通过安置 3' 末端封闭基团而完成。

[0221] 更具体而言,所述的反义寡核苷酸包含 -5' 末端和 -3' 末端;和 11-59 个独立选择自 2'- 修饰磷酸二酯核苷酸和 2'- 修饰 P- 烷氧基磷酸三酯核苷酸的 5' → 3' - 连接核苷酸;且其中 5' - 末端核苷附着于 3-10 个毗连硫代磷酸酯连接脱氧核糖核苷酸的 RNase H- 活化区,且其中所述寡核苷酸的 3' - 末端选自:反向的脱氧核糖核苷酸、1-3 个硫代磷酸酯 2' - 修饰核糖核苷酸的毗连伸展段、生物素基团和 P- 烷氧基磷酸三酯核苷酸。

[0222] 还可使用其中不是 5' 末端核苷附着于 RNase H- 活化区而是 3' 末端核苷如上所述附着于 RNase H- 活化区的反义寡核苷酸。此外,5' 末端选自不同于所述寡核苷酸 3' 末端的特殊基团。

[0223] 按照本发明的核酸以及超免疫血清反应性抗原及其片段可用作或用于制备药物组合物,尤其是疫苗。优选所述的药物组合物、优选疫苗是用于预防或治疗化脓链球菌引起、涉及或相关的疾病。在此范围内,本发明的另一方面涉及在个体、尤其是哺乳动物内引起免疫应答的方法,它包括用按照本发明的超免疫血清反应性抗原及其片段或其片段或变体接种个体,足以产生抗体以保护所述个体不受感染,尤其是链球菌感染,最具体而言是化脓链球菌感染。

[0224] 本发明的另一方面涉及在个体中诱发免疫应答的方法,它包括,通过基因治疗或其它方法运送功能性编码超免疫血清反应性抗原及其片段或其片段或变体的核酸,以便在体内表达超免疫血清反应性抗原及其片段或其片段或变体,从而诱导免疫应答产生抗体或细胞介导的 T 细胞应答(产生细胞因子的 T 细胞或细胞毒性 T 细胞),以保护所述的个体免受疾病的侵袭,无论该疾病是已存在于个体内的或未存在其中的。施加基因的一个方式是将其包被在颗粒或其它物质上加速进入预定细胞中。

[0225] 本发明的另一方面涉及免疫学组合物,所述组合物在引入能在其中诱发免疫应答的宿主时在该宿主中诱发免疫应答,其中的组合物包含编码和表达按照本发明的超免疫血清反应性抗原及其片段的抗原的重组 DNA。免疫学应答可用于治疗或预防,且可采用抗体免疫性或细胞免疫性的形式,诸如由 CTL 或 CD4+T 细胞产生的免疫性。

[0226] 本发明的超免疫血清反应性抗原及其片段或其片段可与辅助蛋白质融合,所述辅助蛋白质可能自身不产生抗体,但能稳定上述第一种蛋白质并产生具有免疫原性和保护特性的融合蛋白。这种融合的重组蛋白质优选进一步包含抗原性辅助蛋白,诸如谷胱甘肽-S- 转移酶 (GST) 或 β - 半乳糖苷酶,稳定蛋白质并方便其产生和纯化的相对大的辅助蛋白。此外,辅助蛋白在提供免疫系统全身性刺激的意义上可用作佐剂。辅助蛋白可附着于第一种蛋白质的氨基或羧基末端。

[0227] 此外,本发明提供了在化脓链球菌感染的动物模型的所述遗传免疫接种实验中使用所述的核酸分子或其特殊片段的方法。所述片段对于鉴定能引起预防或治疗免疫应答的蛋白质表位是特别有效的。此方法可随后从成功抵抗或清除感染所必需的器官内产生特殊意义的单克隆抗体,用于哺乳动物、尤其是人体内化脓链球菌预防剂或治疗方法的研发。

[0228] 超免疫血清反应性抗原及其片段可作为抗原用于宿主的疫苗接种从而产生特异抗体保护宿主抵抗细菌入侵,例如,通过阻断细菌在受损伤组织的粘附。组织损伤的例子包括由例如机械、化学或热损伤或留置装置的植入造成的皮肤或结缔组织伤口,或诸如口腔、乳腺、尿道或阴道等黏膜中的伤口。

[0229] 本发明还包括含与适当载体在一起的免疫原性重组蛋白质的疫苗制剂。由于蛋白质在胃中可分解,因此优选非肠道施用,包括,例如,皮下、肌肉内、静脉内或皮内施用。适于非肠道施用的制剂包括以下各项的水性和非水性无菌注射液:抗氧化剂、缓冲液、抑菌剂和使制剂与个体的体液(优选血液)等渗的溶质,以及可包含悬浮剂或增稠剂在内的水性和非水性无菌悬浮液。制剂可以单位剂量或多剂量包装出现,例如,封装在安瓿和小瓶中,且可以冻干状态储存,只需在临用前加入无菌液态载体即可。疫苗制品还可包括用于增强制品免疫原性的佐剂体系,诸如水包油体系和本领域已知的其它体系。剂量取决于疫苗的特异性且可用常规实验方便的测定。

[0230] 另一方面,本发明涉及包含了本发明所提供针对化脓链球菌而言的超免疫血清反应性抗原或其片段的药物组合物。所述药物组合物可包含抗化脓链球菌的一个或多个超免疫血清反应性抗原或其片段。任选的,所述的化脓链球菌超免疫血清反应性抗原或其片段还可与抗其它病原体的抗原组合于联合药物组合物中。优选地,所述的药物组合物是用于预防或治疗化脓链球菌和/或其它病原体引起的感染的疫苗,针对所述病原体的抗原已包含于该疫苗中。

[0231] 依照另一方面,本发明涉及包含如上所鉴定针对化脓链球菌的编码超免疫血清反应性抗原或其片段的核酸分子的药物组合物。所述的药物组合物可包含编码针对化脓链球菌的超免疫血清反应性抗原或其片段的一个或多个核酸分子。任选的,所述编码超免疫血清反应性抗原或其片段的化脓链球菌核酸分子还可与编码针对其它病原体的抗原的核酸分子组合于联合型药物组合物中。优选地,所述的药物组合物是用于预防或治疗化脓链球菌和/或其它病原体引起的感染的疫苗,针对所述病原体的抗原已包含于该疫苗中。

[0232] 药物组合物可包含任何适当的辅助物质,诸如缓冲物质、稳定剂或更进一步的活性成分,尤其是已知与药物组合物和/或疫苗产生有联系的成分。

[0233] 对于按照本发明的超免疫血清反应性抗原、其片段或其编码核酸分子而言优选的载体和/或赋形剂是用于进一步刺激免疫应答产生超免疫血清反应性抗原、其片段或其编码核酸分子的免疫刺激化合物。优选的在按照本发明的药用制剂中的免疫刺激化合物选自聚阳离子物质,尤其是聚阳离子肽、免疫刺激核酸分子,优选免疫刺激脱氧核苷酸、明矾、弗氏完全佐剂、弗氏不完全佐剂、神经刺激化合物,尤其是人生长激素,或它们的组合。

[0234] 也在本发明范围内的是,药物组合物、尤其是疫苗除了按照本发明的超免疫血清反应性抗原、其片段和/或其编码核酸分子之外还包含生物学或制药学活性的其它化合物。优选地,所述组合物包含至少一种聚阳离子肽。依照本发明使用的聚阳离子化合物可能是显示出符合 WO 97/30721 的特征性作用的任何聚阳离子化合物。优选的聚阳离子化合物选自碱性多肽、有机聚阳离子、碱性聚氨基酸或它们的混合物。这些聚氨基酸应具有至少 4 个氨基酸残基的链长(WO 97/30721)。尤其优选的是类似聚赖氨酸、聚精氨酸和多肽的物质或它们的混合物,其中在 8 个以上氨基酸的范围内、尤其是 20 个以上氨基酸残基的范围内含 20%以上、尤其是 50%以上碱性氨基酸。其它优选的聚阳离子和它们的药物组合物参

阅 WO 97/30721 (例如, 聚乙烯亚胺) 和 WO 99/38528。优选这些多肽包含 20-500 个之间的氨基酸残基, 尤其是在 30-200 个残基之间。

[0235] 这些聚阳离子化合物可化学或重组产生或可来自天然来源。

[0236] 阳离子(多)肽还可以是具有如综述 {Ganz, T., 1999} 中所述的抗微生物特性。这些(多)肽可能是原核的或动物的或植物来源的或可化学或重组产生(WO 02/13857)。肽还可以属于防卫素类型(WO 02/13857)。所述肽的序列可发现于, 例如, 以下网址的抗微生物序列数据库:

[0237] <http://www.bbcm.univ.trieste.it/~tossi/pag2.html>

[0238] 所述宿主防御肽或防御工具也是优选形式的按照本发明的聚阳离子聚合体。通常, 允许作为优选由 APC (包括树状细胞) 介导的适应性免疫系统活化(或下调)终产物的化合物被用作聚阳离子聚合体。

[0239] 在本发明中尤其优选用作聚阳离子物质的是 cathelicidin 来源抗微生物肽或其衍生物(在此收编作为参考的国际专利申请 WO 02/13857), 尤其是来源于哺乳类 cathelicidin、优选来自人、牛或小鼠的抗微生物肽。

[0240] 天然来源的聚阳离子化合物包括 HIV-REV 或 HIV-TAT (来自阳离子肽、触角(antennapedia)肽、壳聚糖或几丁质的其它衍生物)或通过生化或重组产生的来源于这些肽或蛋白质的其它肽。其它优选的聚阳离子化合物是 cathelin 或 cathelin 的相关或衍生物。例如, 小鼠的 cathelin 是具有氨基酸序列 $\text{NH}_2\text{-RLAGLLRKGGEKIGELKKIGOKIKNFFQKLVPQPE-COOH}$ 的肽。相关或衍生 cathelin 物质包含 cathelin 的全部或具有至少 15-20 个氨基酸残基的部分。衍生可包括用不在 20 个标准氨基酸之内的氨基酸进行的天然氨基酸替代或修饰。此外, 更多的阳离子残基可引入所述 cathelin 分子中。这些 cathelin 分子优选与抗原组合。这些 cathelin 分子令人惊讶的显示出也可有效的作为抗原的佐剂而无需添加另外的佐剂。因此可能利用所述 cathelin 分子作为疫苗制品中的有效佐剂, 可以有或没有另外的免疫活化物质。

[0241] 可依照本发明使用的另一优选的多阳离子物质是含有由 3-7 个疏水氨基酸组成的接头隔开的至少 2 个 KKK-基序的合成肽(国际专利申请 W002/32451, 结合于此作为参考)。

[0242] 本发明的药物组合物可进一步包含免疫刺激核酸。免疫刺激核酸是, 例如, 含有核酸的中性或人造 CpG, 所述核酸是来自非脊椎动物的核酸短伸展段或者是在某一碱基段含非甲基化的胞嘧啶-鸟嘌呤双核苷(CpG)的短寡核苷酸形式(ODNs) (例如, 参阅 WO 96/02555)。或者, 如 W001/93903 中所述基于肌苷和胞苷的核酸或含脱氧肌苷和/或脱氧尿苷残基的脱氧核酸(参阅 WO 01/93905 和 PCT/EP 02/05448, 结合于此作为参考)也可优选用作本发明的免疫刺激核酸。优选地, 不同免疫刺激核酸的混合物可依照本发明使用。

[0243] 也属于本发明范围的是任一上述聚阳离子化合物与任一上述免疫刺激核酸组合。优选地, 所述的组合是与结合于此作为参考的以下文献之一所述的一致的: WO 01/93905、WO 02/32451、WO 01/54720、WO 01/93903、WO 02/13857 和 PCT/EP 02/05448 以及奥地利专利申请 A 1924/2001。

[0244] 另外或备选的所述疫苗组合物除了包含按照本发明的超免疫血清反应性抗原及其片段以及它们的编码核酸之外, 还可包含刺激神经组织的化合物。优选地, 刺激神经组织

的化合物是人生长因子,如 WO 01/24822 中所述。同样优选的是,刺激神经组织的化合物与上述任一聚阳离子化合物和 / 或免疫刺激核酸组合。

[0245] 本发明的另一方面涉及药物组合物。所述的药物组合物是,例如,本文所述的疫苗。药物组合物还是包含以下任一种化合物或其组合的药物组合物:按照本发明的核酸分子、按照本发明的超免疫血清反应性抗原及其片段、按照本发明的载体、按照本发明的细胞、按照本发明的抗体、按照本发明的功能性核酸以及诸如抗促成素等按照本发明的结合肽、本文所述的任何激动剂和拮抗剂。与之有关的是任一这些化合物都可与针对细胞、组织或器官的未消毒或消毒载体(一种或多种)组合使用,诸如适于施用于某受试者的药用载体。所述组合物包括,例如,培养基添加剂或治疗有效量的本发明超免疫血清反应性抗原及其片段和药用可接受载体或赋形剂。所述载体可包括,但不局限于:盐、缓冲盐、葡萄糖、水、甘油、乙醇及它们的组合。该制剂应与施用模式相配。

[0246] 药物组合物可以任一有效的、方便的形式施用,其中包括,例如,局部、口、肛门、阴道、静脉内、腹膜内、肌肉内、皮下、鼻内或皮内途径施用。

[0247] 在治疗中或作为预防药,活性试剂可作为可注射组合物施用于个体,例如,作为无菌水分散液,优选等渗的。

[0248] 或者组合物可制备成局部应用的,例如以药膏、霜、洗液、眼膏、滴眼液、滴耳液、漱口、浸渍敷料和缝合线和气雾剂形式,且可包含合适的常规添加剂,包括,例如,防腐剂、帮助药物渗透的溶剂以及在药膏和霜中的润肤剂。这样的局部制剂还可包含兼容的常规载体,例如霜或软膏基质、以及用于洗液的乙醇或油醇。所述的载体组成可占制剂重量的约 1% -98%,更通常的是组成高达制剂重量的约 80%。

[0249] 除了上述治疗之外,本发明组合物通常还可用作伤口处理剂以防止细菌粘附于伤口组织中暴露的基质蛋白质上,且可作为抗生素预防疗法的另一备选方案或与其结合使用于牙科治疗的预防性应用中。

[0250] 疫苗组合物便利的是以可注射形式。常规的佐剂可用于增强免疫应答。疫苗接种的适当单位剂量是 0.05-5 μ g/kg 抗原,这样的剂量优选施用 1-3 次且间隔 1-3 周。

[0251] 采用指定的剂量范围,应不会观察到本发明化合物的毒理学副作用,如果有就将停止它们施用于适当个体。

[0252] 在另一实施方案中,本发明涉及包含了填充有一种或多种本发明上述组合物成分的一个或多个容器的诊断和药物包裹和试剂盒。所述成分可以有效量、剂量、制剂或组合存在。与所述容器相关的是管理制药的政府机构规定形式的说明书、药品或生物产品的用途或销售,反映出制药机构同意该产品的应用或销售是针对人进行施用的。

[0253] 与本发明有关且涉及本文所述用途(例如,药物组合物或疫苗的应用)的任何疾病具体而言是由链球菌、更优选化脓链球菌引起、连接或相关的疾病或疾病状况。与此有关应注意的是化脓链球菌包含本文所述那些菌株在内的数个菌株。涉及依照本发明待预防和 / 或治疗的细菌感染、尤其引起或与其相关的疾病除了包括细菌性咽炎之外,还包括人的猩红热、脓疱病、风湿热、坏死性筋膜炎和脓毒病。

[0254] 在另一实施方案中,本发明涉及利用按照本发明的任一超免疫血清反应性抗原或核酸的筛选方法。筛选方法对于本领域技术人员而言同样是已知的,且可设计成筛选激动剂或拮抗剂的。优选在抑制或阻碍按照本发明的任一超免疫血清反应性抗原及其片段与相

互作用配体的结合的情况下筛选拮抗剂。所述的相互作用配体可以是天然存在的相互作用配体或非天然存在的相互作用配体。

[0255] 本发明还提供了筛选方法用于鉴定化合物是否增强（激动剂）或阻断（拮抗剂）本发明的超免疫血清反应性抗原及其片段或核酸分子的功能，诸如它们与结合分子的相互作用。筛选方法可涉及高通量。

[0256] 例如，为了筛选激动剂或拮抗剂，分别按照本发明的核酸分子和核酸的相互作用配体可能是合成的反应混合物、细胞区室（诸如膜、细胞被膜或细胞壁）或它们中任一种的制剂，可从表达结合本发明的超免疫血清反应性抗原及其片段的分子的细胞中制备。在可能是激动剂或拮抗剂的候选分子缺乏或存在的条件下将所述制剂与标记好的超免疫血清反应性抗原及其片段一起温育。候选分子与结合分子结合的能力反映在被标记配体的结合有所下降。无偿结合的分子，即，不诱发超免疫血清反应性抗原及其片段的功能性作用的分子，最有可能是好的拮抗剂。可很好结合且引起与超免疫血清反应性抗原及其片段相同或非常接近的功能作用的分子是良好的激动剂。

[0257] 潜在激动剂和拮抗剂功能作用可通过，例如，在候选分子与细胞或适当细胞制剂相互作用后测定报道体系的活性来检测，且将此效用与本发明超免疫血清反应性抗原及其片段或所引起作用与超免疫血清反应性抗原及其片段相同的分子的效用进行比较。这方面可用的报道体系包括但不局限于转换成产物的标记底物的比色分析、对超免疫血清反应性抗原及其片段功能活性中的变化作出响应的报道基因以及本领域已知的结合试验。

[0258] 检测拮抗剂的另一例子是在适于进行竞争性抑制试验的条件下将本发明的超免疫血清反应性抗原及其片段和具有膜结合分子、重组结合分子、天然底物或配体或底物或配体模拟物的可能拮抗剂结合。超免疫血清反应性抗原及其片段可被诸如放射性或比色分析化合物标记，从而可准确测定与结合分子结合或转换成产物的超免疫血清反应性抗原及其片段的分子数目以评估可能性拮抗剂的效力。

[0259] 可能性的拮抗剂包括与本发明的超免疫血清反应性抗原及其片段结合并从而抑制或消除其活性的小有机分子、肽、多肽和抗体。可能的拮抗剂还可以是结合于结合分子的同一点且不会诱导产生本发明的超免疫血清反应性抗原及其片段的功能活性的小有机分子、肽、诸如紧密相连的蛋白质或抗体等多肽。

[0260] 可能的拮抗剂包括结合并占领超免疫血清反应性抗原及其片段的结合位点因此阻碍其与细胞结合分子结合从而使正常生物学活性被抑制的小分子。小分子的例子包括但不局限于小的有机分子、肽或肽类似分子。其它的可能性拮抗剂包括反义分子。

[0261] 其它的可能性拮抗剂包括反义分子（参阅 {Okano, H. 等, 1991} ; 作为基因表达反义抑制物的寡脱氧核苷酸 OLIGODEOXYNUCLEOTIDES AS ANTISENSE INHIBITORS OF GENE EXPRESSION ; CRC 出版社, Boca Raton, FL (1988), 以描述这些分子）。

[0262] 优选的可能性拮抗剂包括本发明的超免疫血清反应性抗原及其片段的衍生物。

[0263] 用于此处时，按照本发明的超免疫血清反应性抗原及其片段的活性是它们与其任一相互作用配体结合的性能或与自身的或任何相互作用配体结合的所述性能的程度。

[0264] 在某一具体方面，本发明提供了本发明的超免疫血清反应性抗原及其片段、核酸分子或抑制剂在干预病原体和被造成感染后遗症的哺乳动物宿主之间最初的物理学相互作用中的应用。具体而言本发明的分子可用于：i) 防止化脓链球菌对留置装置上的哺乳

动物胞外基质蛋白质或伤口处的胞外基质蛋白质的粘附 ;ii) 封闭通过例如启动哺乳动物酪氨酸激酶的磷酸化作用而介导哺乳动物细胞侵入的蛋白质 {Rosenshine, I. 等,1992} ; iii) 阻断哺乳动物胞外基质蛋白质和介导组织损伤的细菌蛋白质之间的细菌粘附 ;iv) 阻断了通过留置装置植入或通过其它外科技术之外的方式引起的感染中发病的进程。

[0265] 本文所提供的各个 DNA 编码序列可用于抗菌化合物的探索和开发中。表达后的编码蛋白质可用作抗菌药物筛选的靶目标。此外,编码所述编码蛋白质氨基末端区域的 DNA 序列或相应 mRNA 的 Shine-Delgarno 或其它方便翻译的序列可用于构建反义序列以控制目的编码序列的表达。

[0266] 拮抗剂和激动剂可用于,例如,抑制链球菌、尤其是化脓链球菌引起的疾病,诸如脓毒病。

[0267] 另一方面,本发明涉及亲和力装置,所述的亲和力装置包括至少一种支持材料和附着在支持材料上的任一按照本发明的超免疫血清反应性抗原及其片段。由于按照本发明的超免疫血清反应性抗原及其片段对其靶细胞或靶分子或其相互作用配体的特异性,如果符合结合条件,超免疫血清反应性抗原及其片段可从加到支持物的任一种样品中选择性的除去它们的相互作用配体。样品可以是生物学或医学样品,包括但不限于,发酵液、细胞碎片、细胞制剂、器官制剂、血液、尿、淋巴液、流体等。

[0268] 超免疫血清反应性抗原及其片段可以共价或非共价方式附着于基质上。适当的支持材料是本领域技术人员所已知的且可选自纤维素、硅、玻璃、铝、顺磁珠、淀粉和右旋糖苷。

[0269] 通过以下图、实施例和序列列表对本发明进行了进一步的描述,其中可见更多的特征、实施方案和优势。应理解这些实施例只是以例证方式给出而不是对内容的局限。

[0270] 与本发明有关:

[0271] 图 1 显示了化脓链球菌特异人血清的特性。

[0272] 图 2 显示了来自化脓链球菌 SF370/M1 的小片段基因组文库 LSPy-70 的特性。

[0273] 图 3 显示了用生物素化的人 IgGs 通过 MACS 选择细菌细胞。

[0274] 图 4 显示了用已确定抗原进行基因分布研究的例子。

[0275] 图 5 显示了用流式细胞计着色的细胞表面。

[0276] 图 6 显示了已确定的重组化脓链球菌抗原的预防用途。

[0277] 表 1 显示了用基因组化脓链球菌文库和人血清进行的所有筛选的概况。

[0278] 表 2 显示了有关人血清的表位血清学。

[0279] 表 3 显示了针对 50 个化脓链球菌菌株中已确定抗原的基因分布分析概况。

[0280] 表 4 总结了用于免疫接种实验中的抗原蛋白的有关信息。

[0281] 表 5 显示了在 6 个不同化脓链球菌菌株中抗原蛋白的可变性。

[0282] 本说明书中可提及的图详述如下。

[0283] 图 1 显示了用 ELISA 检测的人血清对化脓链球菌的鉴定。

[0284] 图 2 显示了化脓链球菌 SF370/M1 小片段基因组文库 LSPy-70 的片段大小分布。测序后 576 个随机选择克隆序列被修整以消除载体残基,并对克隆数目和多种基因组片段大小作图。(B) 化脓链球菌染色体上 LSPy-70 的同一套随机测序克隆的分布图解。蓝圈表示在 +/+ 方向上与所注解 ORF 匹配的序列。红色矩形代表在 +/+ 方向上与非编码染色体序列

完全匹配的克隆。绿色菱形标出了具互补性或嵌合序列的所有克隆的位置。碱基对中的数字间隔标示于各个环形基因组上用于定位。在图的底部给出了文库中多个克隆组分配的数目和百分比。

[0285] 图 3A 显示了用生物素化人 IgG 进行的 MACS 选择。在第一轮中用 $10\ \mu\text{g}$ 、第二轮中用 $1\ \mu\text{g}$ 生物素化的人血清 (P4-IgG) 筛选 pMAL9.1 中的 LSPy-70 文库。作为阴性对照,在用于筛选的文库细胞中不添加血清。针对各轮选择均显示了在第一和第二轮洗脱后选择的细胞数。图 3B 显示了通过细菌表面展示根据用人血清 (P4-IgG) 进行的蛋白质印迹分析而分离的特异克隆 (1-52) 的反应性,所述人血清以 1 : 3,000 的稀释度供 MACS 选择。作为负荷对照,用 1 : 5,000 稀释度且直接针对平台蛋白 LamB 的抗体分析相同的印迹。LB,来自表达 LamB 且无外源肽插入的克隆的提取物。

[0286] 图 4A 显示了针对基因分布研究分析的化脓链球菌的 emm 类型。图 4B 显示了用相应寡核苷酸对基因 Spy0269 的基因分布进行的 PCR 分析。预计的 PCR 片段的大小是 1,000bp。1-50,在 A 下列举的化脓链球菌菌株 ;N,未加入基因组 DNA ;P,来自化脓链球菌 SF310 的基因组 DNA,用作文库构建的模板。

[0287] 图 5 用流式细胞计进行的结合于 A 族链球菌细胞表面的特异抗体的检测。在图 5A 中将免疫前小鼠血清和针对化脓链球菌裂解物产生的多克隆血清与化脓链球菌菌株 SF370/M1 一起温育并用流式细胞计进行分析。对照代表二抗对化脓链球菌细胞表面的非特异结合水平。图 5B 和 5C 的柱状图显示分别针对两个平台蛋白 LamB 和 FhuA 而言与对照血清相比,由于抗 -Spy0012 (B) 或抗 -Spy1315 和抗 -Spy1798 (C) 抗体的特异结合而使荧光增强。

[0288] 图 6 用 3 个连续剂量的重组蛋白质 ($50\ \mu\text{g}/\text{剂}$) 间隔两周分别于 0、14 和 28 天免疫 NMRI 小鼠。作为阴性对照,在佐剂的存在下用 PBS 免疫小鼠。M1 蛋白 (Spy2018) 被用作阳性对照用于攻击实验。用 5×10^7 个化脓链球菌 AP1 细胞静脉注射进行细菌的攻击并对小鼠的生存状况分别观察 A) 18 天, B) 21 天和 C) 19 天。

[0289] 表 1 :用细菌表面展示鉴定的免疫原性蛋白质。

[0290] A, 用 IC3-IgG (1588) 鉴定的 lamB 中的 LSPy-7 文库, B, 用 IC3-IgA (1539) 鉴定的 lamB 中的 LSPy-7 文库, C, 用 IC6-IgG (1173) 鉴定的 lamB 中的 LSPy-7 文库, D, 用 P4-IgG (1138) 鉴定的 lamB 中的 LSPy-7 文库, E, 用 P4-IgA (981) 鉴定的 lamB 中的 LSPy-7 文库, F, 用 IC3-IgG (991) 鉴定的 btuB 中的 LSPy-150 文库, G, 用 IC6-IgG (1036) 鉴定的 btuB 中的 LSPy-150 文库, H, 用 P4-IgG (681) 鉴定的 btuB 中的 LSPy-150 文库, I, 用 IC3-IgG (559) 鉴定的 fhuA 中的 LSPy-400 文库, K, 用 IC6-IgG (543) 鉴定的 fhuA 中的 LSPy-400 文库, L, 用 P4-IgG (20) 鉴定的 fhuA 中的 LSPy-400 文库, *, 长于 5 个氨基酸的抗原序列的预测用软件 ANTIGENIC {Kolaskar, A. 等, 1990} 进行。

[0291] 表 2 :关于人血清的表位血清学。

[0292] 显示了代表用个别人血清选定表位的个别合成肽的免疫反应性。反应性范围是图案 / 灰色编码 ; 白色 : - ($< 50\text{U}$), 灰色 : + ($50-119\text{U}$), 斜线 : ++ ($120-199\text{U}$), 对角交叉线 : +++ ($200-1000\text{U}$) 以及垂直交叉 : ++++ ($> 1000\text{U}$)。在背景校正后从 $\text{OD}_{545\text{nm}}$ 的读数和血清稀释度计算 ELISA 单位 (U)。得分, 所有反应性的总和 (所有 + 数目的相加) ; P1 至 P10 血清来自急性咽炎患者而 N1 至 N10 血清来自健康成人。P 和 N 用作内在对照。

[0293] 肽名称:SP00012,注释为 ORF Spy0012;SPA0450, Spy0450 选择性读码框架中的潜在新 ORF;SPC0406, Spy0406 互补链上的潜在新 ORF;SPN0001,非编码区中的潜在新 ORF。

[0294] 表 3:化脓链球菌菌株中的基因分布。

[0295] 用特异于相应抗原编码基因的寡核苷酸通过 PCR 检测图 4A 中所示的 50 个化脓链球菌菌株。对某一选定 PCR 片段的 PCR 片段进行测序以证实正确 DNA 片段的扩增。*,与化脓链球菌 SF370 (M1) 相比菌株 M89 中的氨基酸替代的数目。#,因为基因不存在于 M89 中而用于测序的备选菌株。

[0296] 表 4:用于 NMRI 小鼠中免疫接种实验的重组蛋白质。

[0297] 如实验步骤中所述用重组抗原进行免疫接种以及用致病的化脓链球菌 AP1 攻击。A,给出了与全长蛋白质相比的用于动物免疫接种实验的重组蛋白质中所含的相应抗原的氨基酸。B,生存率的百分比表示为保护状态且圆括号中插入语描述了阴性对照 (PBS 免疫) 的保护状态的百分比,随后是阳性对照 (Spy2018) 保护状态的百分比。C,由于尽管在观察结束时所有小鼠都死了,但小鼠仍显示出较好的生存状态这一事实而选择 Spy0269。这通过检测如下的平均生存时间可以反映出来:14.6 (Spy0269)、11.6 (PBS) 和 19.3 天 (Spy2018)。

[0298] 表 5:来自化脓链球菌的抗原蛋白的序列变异。

[0299] 针对在实验步骤下所列举的 6 个不同化脓链球菌中的氨基酸交换分析抗原蛋白质。残基编号指示该氨基酸在全长蛋白质中的位置。以 Spy1666 为例,列举了相对于肺炎链球菌 TIGR4 (SP0334) 中同源基因的变化,因为基因在化脓链球菌以及肺炎链球菌中是高度保守的。A,来自化脓链球菌 SF370 的蛋白质中的氨基酸残基;B,如果不同于化脓链球菌 SF370,可能存在于来自另 5 个化脓链球菌菌株的任一被分析基因中的氨基酸残基;C,催化活性中所涉及的 Spy0416 残基。预期这些残基的变化以使所述酶失活并因此在实验上用丙氨酸、丝氨酸、苏氨酸交换甘氨酸以产生酶失活的重组蛋白质。

[0300] 实施例

[0301] 实施例 1:从人血清制备抗体

[0302] 人免疫体系产生且存在人血清中的抗 A 族链球菌抗体是抗原蛋白质体内表达及其免疫原性的的指征。这些分子是在本发明所述方法中个别抗原的鉴定所必需的,它是以特异的抗链球菌抗体及其相应的化脓链球菌肽或蛋白质之间的相互作用为基础的。为了得到相应的抗体全部组分,从以下来源收集人血清:

[0303] I. 急性化脓链球菌感染患者,诸如咽炎、伤口感染和菌血症(通过医学微生物学检测显示出化脓链球菌是病原体),

[0304] II. 未感染的健康成人,由于 A 族链球菌感染是常见的,抗体作为天然免疫的结果存在于先前接触了链球菌的个体中。

[0305] 用一系列 ELISA 和免疫印迹检测针对抗化脓链球菌抗体鉴定血清。数种链球菌抗原已用于显示所检测的效价不是交叉反应抗体的总和结果。为此目的使用了两种不同的抗原制剂:从过夜培养(稳定期)于 THB(Todd-Hewitt 肉汤)生长培养基中的化脓链球菌 SF370/M1 中制备全细胞提取物或培养物上清蛋白质。IgG 和 IgA 两种抗体水平均被测定。以针对两种抗原制剂的总抗体滴度为基础通过免疫印迹选择血清以供进一步的分析。

[0306] 在反应为线性的给定稀释度比较所述滴度(图 1)。将血清按针对多种链球菌组分的反应性进行排列,选用那些反应性最高的血清进行进一步的免疫印迹分析。此广泛的抗

体表征方法已使抗链球菌超免疫血清得以明确的鉴定。

[0307] 最近,据报道不仅是 IgG,还有 IgA 血清抗体可被 PMN 的 FcRIII 受体识别并促进调理作用 {Phillips-Quagliata, J. 等,2000 ;Shibuya, A. 等,2000}。IgA 抗体的首要任务是起中和作用,主要在黏膜表面。血清 IgA 的水平反映了二聚体分泌 IgA 的质量、数量和特异性。为此原因,不只分析血清集合的抗链球菌 IgG,还可分析其 IgA 水平。在 ELISA 检测法中使用了高度特异的二级试剂来检测高亲和力类型的抗体,诸如 IgG 和 IgA,而消除了 IgM。IgM 抗体的产生出现于初级适应性体液应答期间并产生了低亲和力抗体,而 IgG 和 IgA 抗体已经过亲和成熟且在抗击或预防疾病中更有价值。

[0308] 实验步骤

[0309] 肽合成

[0310] 用 SyroII 合成仪 (Multisyn tech, Witten, Germany) 在 Rink 酰胺树脂上 (PepChem, Tübingen, Germany) 通过标准 F-moc 化学法小量合成肽 (4mg 树脂,平行最多合成 288mg)。序列装配好后,用 Fmoc- ϵ -氨基己酸 (作为接头) 和生物素 (Sigma, St. Louis, MO; 象正常氨基酸一样被激活) 延伸肽。利用 93% TFA、5% 三乙基硅烷和 2% 水作用 1 小时将肽从树脂上分离下来。将肽从乙腈/水 (1 : 1) 中真空干燥并冻干三次。用质谱技术在 Reflex IIIMALDI-TOF (Bruker, Bremen Germany) 上证实正确质量的存在。所述的肽直接使用而无需进一步纯化。

[0311] 酶联免疫检测法 (ELISA)。

[0312] 对于血清鉴定而言:用稀释于包被缓冲液 (0.1M 碳酸钠 pH 9.2) 的 5-10 μ g/ml 总蛋白包被 ELISA 平板 (Maxisorb, Millipore)。在 PBS-BSA 中制备了三种稀释度的血清 (2,000X、10,000X、50,000X)。

[0313] 对于肽血清学而言:依照制造商的说明书以 10 μ g/ml 的浓度将生物素标记的肽包被于链霉亲和素 ELISA 平板 (EXICON) 上。在两个稀释度检测血清,200X 和 1,000X。

[0314] 依照制造商的建议 (稀释度:1,000x) 使用高特异性的辣根过氧化物酶 (HRP)-偶联的抗人 IgG 或抗人 IgA 二抗 (Southern Biotech)。基于自动 ELISA 读数仪 (TECAN SUNRISE) 中的 OD_{405nm} 读数通过检测底物 (ABTS) 向显色产物的转换从而对抗原-抗体复合物进行定量。在人工包被后,用内置读数仪 (GENIOS, TECAN) 通过 Gemini 160ELISA 自动仪 (TECAN) 对肽平板进行处理和分析。

[0315] 免疫印迹法

[0316] 从体外培养的化脓链球菌 SF370/M1 中制备总细菌裂解产物和培养物上清样品。用 BioRad Mini-Protean 3 Cell 电泳体系通过 SDS-PAGE 分离 10-25 μ g 总蛋白/泳道并将蛋白质转到硝酸纤维素膜 (ECL, AmershamPharmacia) 上。在 5% 的牛奶中过夜封闭后,加入 2,000x 倍稀释的抗血清,并将 HRP 标记的抗小鼠 IgG 用于检测。

[0317] 细菌抗原提取物的制备

[0318] 总细菌裂解产物:重复冻融裂解细菌:温育于干冰/乙醇混合物中直至冻结 (1 分钟),然后融解于 37°C (5 分钟);重复 3 次。随后超声波降解并离心收集上清 (3,500rpm, 15 分钟,4°C)。

[0319] 培养物上清液:除去细菌后,用冰冷的乙醇 (100%) 沉淀过夜培养的细菌培养物上清液:1 份上清液/3 份乙醇于 -20°C 放置过夜。离心收集沉淀 (2,600g, 15 分钟) 并

干燥。将干的沉淀或者溶解于 PBS 以备 ELISA, 或者溶解于尿素和 SDS 样品缓冲液中以备 SDS-PAGE 和免疫印迹。用 Bradford 检测法检测样品的蛋白质浓度。

[0320] 纯化抗体以备基因组筛选。基于筛选步骤中所用血清集合的总抗链球菌滴度选择来自患者以及未感染组的 5 份血清。通过将热灭活血清与所有样品中的大肠杆菌细胞 (DH5 α , 用 pHIE11 转化, 在与细菌表面展示同样的条件下培养) 一起保温去除抗大肠杆菌蛋白质的抗体。依照制造商的说明书 (UltraLink Immobilized Protein G, Pierce) 通过蛋白 G 亲和层析从收集的缺乏血清中产生高度富集的 IgG 制品。IgA 抗体也是通过亲和层析利用固定于链霉亲和素-琼脂糖 (GIBCO BRL) 上的生物素标记抗人 IgA (SouthernBiotech) 进行纯化。通过 SDS-PAGE、蛋白质印迹、ELISA 和蛋白质浓度检测检查缺失和纯化的效率。

[0321] 实施例 2: 化脓链球菌的高度随机且框架选定的小片段基因组 DNA 文库的产生

[0322] 实验步骤

[0323] 链球菌基因组 DNA 的制备。将化脓链球菌 SF370/M1 从冻存穿刺斜面接种于 50ml Todd-Hewitt Broth 培养基中并在 37°C 通风振荡培养 18 小时。然后收获培养基, 1,600xg 离心 15 分钟去上清。细菌沉淀用 PBS 洗 3 次并小心重悬于 0.5ml 溶菌酶溶液 (100mg/ml) 中。加入 0.1ml 的 10mg/ml 热处理 RNase A 和 20U 的 RNase T1, 小心混和并将此溶液在 37°C 保温 1 小时。加入 0.2ml 的 20% SDS 溶液和 0.1ml 的蛋白酶 K (10mg/ml) 后将此试管在 55°C 保温过夜。然后加入 1/3 体积的饱和 NaCl 并将该溶液在 4°C 保温 20 分钟。在微量离心管中沉淀提取物 (13,000rpm) 并将上清转移至一新的试管中。用 PhOH/CHCl₃/IAA (25 : 24 : 1) 和 CHCl₃/IAA (24 : 1) 抽提该溶液。加入 0.6 倍体积的异丙醇室温沉淀 DNA, 用消毒的巴斯德移液管从所述溶液中将 DNA 缠绕在上面并转移入含 80% 冰冻乙醇的试管中。10-12,000xg 离心沉淀回收 DNA, 然后空气干燥并溶于 ddH₂O 中。

[0324] 小基因组 DNA 片段的制备。将基因组 DNA 片段用杯状钻头状超声波仪 (装备了 BB5 杯状钻头的 Bandelin Sonoplus UV 2200 超声波仪, 10 秒脉冲, 100% 功率输出) 机械切割成大小 150-300bp 之间的片段或通过温和的 DNase I 处理 (Novagen) 切割成大小 50-70bp 范围内的片段。可观察到当将 DNA 断裂成 150-300bp 大小范围内的片段时, 超声波降解法产生了密集得多的片段大小分布。不过, 尽管 DNA 广泛暴露于超声波诱导的流体力学剪切力中, 并不能同时有效和重复的达到减小片段大小的目的。因此, 50-70bp 大小的片段通过温和的 DNase I 处理利用 Novagen 的鸟枪裂解试剂盒获得。制备试剂盒所提供的 DNase I 的 1 : 20 稀释液并在 MnCl₂ 存在的条件下于 60 μ l 的体积中在 20°C 消化 5 分钟以保证双链都被酶所切割。用 2 μ l 的 0.5M EDTA 终止反应且在 2% TAE-琼脂糖凝胶上评估片段化效率。此处理导致基因组 DNA 全部断裂成近似 50-70bp 的片段。然后在各种 dNTP 各 100 μ M 存在的条件下用 T4 DNA 聚合酶将片段末端补平以保证末端的有效对齐。片段可立即用于连接反应或冻存于 -20°C 备用。

[0325] 载体的描述。在 pASK-IBA 主链 {Skerra, A., 1994} 上构建载体 pMAL4.31, 用 β -内酰胺酶 (bla) 基因调换卡那霉素抗性基因。此外将 bla 基因克隆入多克隆位点。编码成熟 β -内酰胺酶的序列之前有 ompA 的前导序列以使其能有效分泌穿过细胞质膜。而且编码成熟 β -内酰胺酶的头 12 个氨基酸的序列 (间隔序列) 接着 ompA 前导肽序列之后以避免序列直接在前导肽酶切割位点之后融合, 因为, 例如, 此区域的正电荷氨基酸簇会减少或完全破坏穿越细胞质膜的迁移 {Kajava, A. 等, 2000}。SmaI 限制性位点是用于文库插入的。用

于回收选定片段的上游 FseI 位点和下游 NotI 位点在 SmaI 位点两侧。所述的三个限制性位点在编码所述的 12 个氨基酸间隔序列的序列之后插入,插入方式是 bla 基因转录于 -1 阅读框架,导致在 NotI 位点之后 15bp 产生了一个终止密码子。+1bp 插入恢复了 bla ORF 从而产生了 β 内酰胺酶蛋白并随之获得了氨苄青霉素抗性。

[0326] 通过将 lamB 基因克隆入 pEH1 的多克隆位点构建载体 pMAL9.1 {Hashemzadeh-Bonehi, L. 等,1998}。随后,在第 154 位氨基酸之后将含限制位点 FseI、SmaI 和 NotI 的序列插入 lamB 中。此插入的阅读框架以此方式构建,即,通过用 FseI 和 NotI 消化从质粒 pMAL4.31 切下框架选定的 DNA 片段进行转移产生了 lamB 的连续阅读框架以及各自的插入片段。

[0327] 通过将 btuB 基因克隆入 pEH1 的多克隆位点构建载体 pMAL10.1。随后,在第 236 位氨基酸之后将含限制位点 FseI、XbaI 和 NotI 的序列插入 btuB 中。此插入的阅读框架以此方式构建,即,通过用 FseI 和 NotI 消化从质粒 pMAL4.31 切下框架选定的 DNA 片段进行转移产生了 btuB 的连续阅读框架以及各自的插入片段。

[0328] 通过将 fhuA 基因克隆入 pEH1 的多克隆位点构建载体 pHIE11。随后,在第 405 位氨基酸之后将含限制位点 FseI、XbaI 和 NotI 的序列插入 fhuA 中。此插入的阅读框架以此方式选择,即,通过用 FseI 和 NotI 消化从质粒 pMAL4.31 切下框架选定的 DNA 片段进行转移产生了 fhuA 的连续阅读框架以及各自的插入片段。

[0329] 有关框架选择的文库克隆和评估。将化脓链球菌基因组 DNA 片段在 SmaI 位点连接入载体 pMAL4.31 中。使重组 DNA 通过电穿孔进入 DH10B 电感受态大肠杆菌细胞 (GIBCO BRL) 并将转化子铺于补充了卡那霉素 (50 μ g/ml) 和氨苄青霉素 (50 μ g/ml) 的 LB 琼脂上。平板在 37°C 保温过夜并收集菌落供大量提取 DNA。贮藏并保留一个有代表性的平板以便收集菌落供菌落 PCR 分析和大规模测序。最初用简单菌落 PCR 检测确定大致的片段大小分布以及插入效率。结合插入位点的完整性以及框架选择的精确性 (3n+1 准则) 从测序资料评估片段的大小。

[0330] 有关细菌表面展示的文库克隆和评估。用限制性酶 FseI 和 NotI 从 pMAL4.31 载体上切下含化脓链球菌文库的基因组 DNA 片段。然后将全部片段群转移入已用 FseI 和 NotI 消化的质粒 pMAL9.1 (LamB)、pMAL10.1 (BtuB) 或 pHIE11 (FhuA)。利用这两种识别 8bp 富 GC 序列的限制性酶,在各个平台载体中保持 pMAL4.31 载体中选定的阅读框架。然后通过电穿孔将所述质粒文库转化入大肠杆菌 DH5 α 细胞中。细胞铺于补充了 50 μ g/ml 卡那霉素的大 LB 琼脂平板上并以产生清晰可见单菌落的密度在 37°C 培养过夜。然后将细胞从这些平板表面刮下,用新鲜 LB 培养基漂洗并以等份试样贮存于 -80°C 以备文库筛选。

[0331] 结果

[0332] 用于框架选择的文库。在 pMAL4.31 载体中产生了大小分别约为 70、150 和 300bp 的三个文库 (LSPy70、LSPy150 和 LSPy300)。对于每个文库而言,大约 1 μ g pMAL4.31 质粒 DNA 和 50ng 片段化的基因组化脓链球菌 DNA 的连接和随后的转化在框架选择后产生了 4×10^5 – 2×10^6 个克隆。为了评估文库的随机性,对大约 600 个随机选择的 LSPy70 克隆进行测序。生物信息分析显示这些克隆中只有极少的克隆出现了一次以上。此外,还显示出 90% 的克隆大小落在大小 16–61bp 之间的范围内,平均大小为 34bp (图 2)。所有序列遵从 3n+1 原则,显示出所有克隆均为正确的框架选择。

[0333] 细菌表面展示文库。大肠杆菌表面的肽展示要求将来自 LSPy 文库的插入片段从框架选择载体 pMAL4.31 转移至展示质粒 pMAL9.1 (LamB)、pMAL10.1 (BtuB) 或 pHIE11 (FhuA)。用 FseI 和 NotI 限制酶切割基因组 DNA 片段并将 5ng 插入片段与 0.1 μ g 质粒 DNA 连接并随后转化入 DH5a 细胞,产生 $2-5 \times 10^6$ 个克隆。将克隆从 LB 平板上刮下并冻存而无需进一步扩增。

[0334] 实施例 3:用细菌表面展示基因文库和人血清鉴定来自化脓链球菌的高免疫原性肽序列

[0335] 实验步骤

[0336] MACS 筛选将来自给定文库的约 2.5×10^8 个细胞在补充了 50 μ g/ml 卡那霉素的 5ml LB 培养基中于 37℃ 培养 2 小时。通过加入 1mM IPTG 30 分钟诱导表达。用新鲜的 LB 培养基漂洗细胞两次并将大约 2×10^7 个细胞重悬于 100 μ l LB 培养基中并转移至 Eppendorf 管中。

[0337] 将纯化自血清的 10 μ g 生物素化人 IgG 加入细胞中并将此悬浮液在 4℃ 温和摇动保温过夜。加入 900 μ l LB 培养基,将悬浮液混和并随后在 4℃ 6,000rpm 离心 10 分钟(对于 IgA 筛选而言,使用了 10 μ g 的纯化 IgA 并用生物素化的抗人 IgG 二抗捕捉它们)。用 1ml LB 漂洗细胞一次,然后重悬于 100 μ l LB 培养基中。加入与链霉亲和素偶联的 MACS 微珠 (Miltenyi Biotech, Germany) 10 μ l 并在 4℃ 继续保温 20 分钟。之后加入 900 μ l LB 培养基并将 MACS 微珠细胞悬浮液加到固定于磁铁上且已平衡好的 MS 柱 (Miltenyi Biotech, Germany) 上。(所述 MS 柱通过用 1ml 70% 乙醇洗涤一次和 2ml LB 培养基洗涤两次被平衡。)

[0338] 然后用 3ml LB 培养基洗柱三次。去除磁铁后,用 2ml LB 培养基洗脱细胞。用 3ml LB 培养基洗柱后,将 2ml 洗脱液再次加到同一根柱子上并重复洗脱步骤。第三次进行上样、洗涤和洗脱操作,产生最终的洗脱液 2ml。

[0339] 第二轮筛选进行如下。离心收集最终的洗脱液中的细胞并重悬于补充了 50 μ g/ml 卡那霉素的 1ml LB 培养基中。将培养基在 37℃ 保温 90 分钟然后用 1mM IPTG 诱导 30 分钟。随后收集细胞,用 1ml LB 培养基洗涤一次并重悬于 10 μ l LB 培养基中。由于体积减少,加入 1 μ g 生物素化的人 IgG 并在悬浮液在 4℃ 温和摇动保温过夜。所有其它步骤完全与第一轮筛选相同。将两轮筛选后的细胞铺于补充有 50 μ g/ml 卡那霉素的 LB 琼脂平板上并于 37℃ 培养过夜。

[0340] 通过测序和蛋白质印迹分析评估选定的克隆。将选定的克隆在 37℃ 于补充有 50 μ g/ml 卡那霉素的 3ml LB 培养基中培养过夜以使用标准方法制备质粒 DNA。在 MWG (Germany) 或与 TIGR (U. S. A.) 协作完成测序。

[0341] 为了进行蛋白质印迹分析,用 10% SDS-PAGE 分离大约 10-20 μ g 总细胞蛋白并印迹至 HybondC 膜 (Amersham Pharmacia Biotech, England) 上。用大约 1:5,000 稀释的人血清作为一抗和 1:5,000 稀释的偶联至 HRP 的抗人 IgG 或 IgA 抗体作为二抗检测 LamB、BtuB 或 FhuA 融合蛋白。用 ECL 检测试剂盒 (Amersham Pharmacia Biotech, England) 完成检测。或者,用兔抗 FhuA 或小鼠抗 LamB 抗体作为一抗与偶联至 HRP 的相应二抗联合检测融合蛋白。

[0342] 结果

[0343] 使用生物素化的 Igs 通过磁性活化细胞分选术 (MACS) 筛选细菌表面展示文库。用来自患者血清或健康人血清的生物素化人 IgG 和 IgA 集合筛选 pMAL9.1 中的 LSPy70、pMAL10.1 中的 LSPy150 和 pHIE11 中的 LSPy300 (参阅实施例 1: 从人血清制备抗体)。如下实验流程所述完成选择步骤。图 3A 显示了用 LSPy-70 文库和 P4-IgG 进行筛选的一个代表性的例子。正如从 MACS 筛选的第一轮选择后的菌落数中可见的, 最终回收到的细胞总数从 3×10^7 个细胞急剧下降到大约 5×10^4 个细胞, 而不加入抗体进行的选择则显示下降到大约 2×10^3 个细胞 (图 3A)。第二轮选择后, 用 P4-IgG 回收到的细胞数目的细胞, 而当不加入来自人血清的 IgG 时回收到的细胞少于 10 个, 清楚的证明了选择依赖于化脓链球菌特异抗体。为了评估筛选的效果, 随机挑选大约 50 个选定的克隆并用相同的集合的血清进行蛋白质印迹分析 (图 3B)。此分析表明 70% 的选定克隆显示出对存在于相关血清中的抗体的反应性而表达 LamB 且无化脓链球菌特异性插入片段的对照菌株则不与同一种血清反应。通常, 观察到的反应比例在 35-75% 范围内。菌落 PCR 分析显示所有的选定克隆都含有预期大小范围内的插入片段。

[0344] 随后对大量随机挑选的克隆 (每次筛选 600-1200 个克隆) 进行测序, 鉴定用于筛选的人血清所特异识别的基因和相应的肽或蛋白质序列。特异克隆被选择的频率至少在部分程度上反映了用于选择的血清中识别该克隆所呈递表位的特异抗体的丰度和 / 或亲和力。在这点上令人吃惊的是, 来自某些 ORF (例如 Spy0433、Spy2025) 的克隆被选择次数超过 80 次, 证明了它们的高度免疫原性特性。表 1 总结了获自所有 15 个已完成筛选的资料。表 1 中的所有克隆已通过蛋白质印迹分析用来自单个克隆的全细胞提取物来显示与各个筛选中所用人血清集合的指征反应性而得以证实。正如从表 1 中可见的, 已鉴定 ORF 的单独区域被识别为是免疫原性的, 因为大小不定的蛋白质片段通过平台蛋白质展示其表面。

[0345] 更值得关注的是, 通过细菌表面展示筛选所鉴定的大部分基因编码附着于化脓链球菌表面和 / 或被分泌的蛋白质。这与化脓链球菌毒性中表面附着或分泌蛋白质的预期任务是一致的。

[0346] 实施例 4: 高度免疫原性肽序列对个体人血清的反应性的评估。

[0347] 分析中包括大约 100 名患者和 60 名健康成人。在选定克隆的生物信息分析之后, 设计并合成相应的肽。如果表位长度超过 28 个氨基酸残基, 就制备重叠肽。用 N 末端生物素标签合成所有的肽并将其用作链霉亲和素包被 ELISA 平板上的包被试剂。

[0348] 分析分两步进行。首先, 基于它们对个别血清的反应性选择肽, 所述个别血清包括在用于制备供细菌表面展示的 IgG 和 IgA 筛选试剂的血清集合内 (5 个个体血清)。未呈现阳性反应的肽不包括在进一步更详细的试验中。其次, 针对显示出对筛选血清具有特异的高反应性的肽检测来自急性咽炎患者或链球菌感染后疾病患者或来自健康成人和儿童的大量非预定个体血清。通过 ELISA 检测抗体水平并通过基于阳性血清数目和反应性程度计算的各肽得分进行比较。表 2 显示了代表来自 20 份人血清 (代表 5 种血清的 4 个不同集合) 的化脓链球菌表位的 174 个肽的血清反应性用于抗原鉴定的实例。所述的肽的范围包括高度反应性和广泛反应性的肽至弱阳性的肽。在最具有反应性的肽中包括已知的抗原, 其中一些还在有关鼻咽携带的动物攻击模型中起保护作用 (如 C5a 肽酶和 M 蛋白)。

[0349] 实施例 5: 用鉴定自化脓链球菌的高度免疫原性蛋白质进行的基因分布研究。

[0350] 通过 PCR 进行 A 族链球菌抗原的基因分布研究。理想的疫苗抗原应是存在于该疫

苗所定向的所有或绝大部分靶生物体株系内的抗原。为了确定编码已鉴定的化脓链球菌抗原是否广泛存在于化脓链球菌菌株内,用特异于目的基因的引物对一系列独立的化脓链球菌分离株进行 PCR。如图 4A 中所示,所获得的化脓链球菌覆盖了最频繁出现于患者中的 emm 类型。如果有可能覆盖所有的已鉴定免疫原性表位,就针对所有已鉴定的 ORF 设计作为引物的寡核苷酸序列,产生的产物大约为 1,000bp。如实施例 2 中所述制备所有的化脓链球菌菌株的基因组 DNA。按制造商的说明书 (Invitrogen, The Netherlands) 用 Taq 聚合酶 (1U)、200nM dNTPs、寡核苷酸各 10pMol 和试剂盒在 25 μ l 反应体积内进行 PCR。作为标准,进行 30 个循环 (1x :5 分钟, 95 $^{\circ}$ C, 30x :30 秒 95 $^{\circ}$ C, 30 秒 56 $^{\circ}$ C, 30 秒 72 $^{\circ}$ C, 1x4 分钟 . 72 $^{\circ}$ C), 除非条件需修改以适合个别引物对。

[0351] 结果

[0352] 用 PCR 检测编码免疫原性蛋白质的所有已鉴定基因是否存在于 50 个不同的化脓链球菌菌株中 (图 4A)。例如,图 4B 显示了对于 Spy0269 而言所有指定 50 个菌株的 PCR 反应。正如明显可见的,基因存在于所有被分析的菌株中。对来自第 8 号菌株 (M89) 的 PCR 片段进行测序显示,与化脓链球菌 M1 菌株 SF310 相比,917bp 中只有 2bp 是不同的,导致这两个分离株中只有一个氨基酸的差异。

[0353] 在总共 96 个被分析基因中,有 70 个存在于所有被测试的菌株中,而 22 个基因在被测试的 50 个菌株中的 10 个以上菌株中缺乏 (表 3)。数个基因 (Spy0433、Spy0681) 显示出大小的差异且不存在于所有的分离菌株。某些基因显示出了大小的变化,但却不同的保存于所有的被测试菌株中 (如 Spy1371)。对某一菌株的 PCR 片段进行测序并随后与 M1 菌株比较证实了 DNA 片段的正确扩增并显示了表 3 中所指出的序列差异程度。重要的是,所有的已鉴定抗原的序列和大小在所有菌株中都是很保守的且因此是预防 A 族链球菌感染的新候选疫苗。

[0354] 实施例 6 :从来自化脓链球菌并展示于大肠杆菌表面的高度免疫原性蛋白质 / 肽所免疫的小鼠中获得的免疫血清的鉴定。

[0355] 从小鼠中产生免疫血清

[0356] 将含有编码与化脓链球菌肽融合的平台蛋白的质粒的大肠杆菌克隆在补充有 50 μ g/ml 卡那霉素的 LB 培养基中于 37 $^{\circ}$ C 培养。1 : 10 稀释过夜培养物,培养至 OD₆₀₀ 为 0.5 并用 0.2mM IPTG 诱导 2 小时。将细菌细胞沉淀悬浮于 PBS 缓冲液中并在冰上通过超声波降解法分解,产生了粗细胞提取物。根据 OD₆₀₀ 检测值,将相当于 5×10^7 个细胞的等分试样静脉注射入 NMRI 小鼠中,2 周后加强免疫一次。第二次注射 1 周后提取血清。通过肽 ELISA 检测表位特异的抗体水平。

[0357] 抗原的体外表达

[0358] 通过免疫印迹检测体外培养的化脓链球菌 SF370/M1 的抗原表达。测试不同的生长培养基和培养条件以便检测总裂解产物和细菌培养上清液中抗原的存在。当检测到相当于预计分子量和电泳迁移率的特殊条带时认为表达被证实了。

[0359] 细胞表面染色

[0360] 流式细胞术分析进行如下。细菌在根据免疫印迹显示导致抗原表达的培养条件下生长。细胞在 Hanks 平衡盐溶液 (HBSS) 中漂洗两次并将细胞密度调整到 100 μ l HBSS, 0.5% BSA 中大约为 1×10^6 CFU。与 50-100 倍稀释的抗血清一起在 4 $^{\circ}$ C 保温 30-60 分钟后,

在过量 HBSS, 0.5% BSA 中离心洗去未结合抗体。将用荧光素 (FITC) 特异标记的山羊抗小鼠二抗 (F(ab')₂) 片段与细胞一起在 4℃ 保温 30–60 分钟。漂洗细胞后, 用 2% 多聚甲醛固定抗体。用 Becton Dickinson FACScan 流式细胞仪检测被结合的抗体并用计算机软件 CELLQuest 进一步分析数据。对照血清包括小鼠免疫前血清和用制备自 IPTG 诱导的大肠杆菌细胞的裂解产物产生的小鼠多克隆血清, 所述大肠杆菌是用编码基因 lamB 或 fhuA 且无化脓链球菌基因组插入片段的质粒转化的。

[0361] 调理噬菌作用检测

[0362] 在以 FACS 为基础的检测法中检测表位特异免疫血清诱发调理噬菌作用的活性。血清被热灭活, 然后通过完整的大肠杆菌细胞 (3x) 一起保温去除抗大肠杆菌抗体。在 2–10% 免疫血清存在且以 2% 仓鼠血清作为补充来源的条件下预调理 10⁷ 个 Alexa 488 标记的化脓链球菌细胞, 然后加入 10⁶ 个噬菌细胞 (RAW246.7 或 P388.D1 小鼠单核细胞系) 中。将细胞混合物在 37℃ 保温 30 分钟。时间、IgG 浓度和细菌的补体依赖型吸收表现为用荧光活化细胞分选器检测的噬菌细胞的平均荧光强度。

[0363] 杀菌 (处死) 试验

[0364] 将小鼠巨噬细胞 (RAW246.7 或 P388.D1) 和细菌保温并通过菌落计数测定 60 分钟后可生存细菌的损失。简而言之, 细菌用 Hanks 平衡盐溶液 (HBSS) 洗两次并将细胞密度调整到 50 μl HBSS 中大约 1×10⁵ CFU。将细菌与小鼠血清 (高达 25%) 和豚鼠补体 (高达 5%) 以总体积 100 μl 于 4℃ 保温 60 分钟。将预调理的细菌与巨噬细胞 (小鼠细胞系 RAW246.7 或 P388.D1; 每 100 μl 2×10⁶ 个细胞) 以 1:20 的比例混合并于 37℃ 旋转摇床上 500rpm 保温。将各样品的等分试样稀释于无菌水中并室温保温 5 分钟以裂解巨噬细胞。然后将系列稀释液铺到 Todd-Hewitt Broth 琼脂平板上。将平板在 37℃ 保温过夜, 并用 Counterstat 闪射计数器 (IUL Instruments) 对菌落进行计数。对照血清包括小鼠免疫前血清和用制备自 IPTG 诱导的大肠杆菌的裂解产物产生的小鼠多克隆血清, 所述大肠杆菌是用含有基因 lamB 或 fhuA 而无化脓链球菌基因组插入片段的质粒转化的。

[0365] 结果

[0366] 体外表达和细胞表面染色。通过利用所培养的抗大肠杆菌克隆的血清在体外分析化脓链球菌 SF370/M1 中的抗原蛋白质的表达, 所述大肠杆菌克隆含有编码与化脓链球菌肽融合的平台蛋白的质粒。为了通过 FACS 分析评估多肽的表面表达, 将此分析作为测定蛋白质是否完全表达的第一步。预期不是所有的蛋白质都会在体外条件下表达的, 但利用蛋白质印迹分析在全部细胞裂解物中检测到了数种蛋白质 (例如 Spy0012、Spy0112、Spy0416、Spy0437、Spy0872、Spy1032、Spy1315、Spy1798; 数据未显示)。随后用基于流式细胞术的检测法证实了数种抗原蛋白质的细胞表面可接触性。将链球菌与免疫前的和抗化脓链球菌裂解产物或大肠杆菌克隆的多克隆小鼠血清一起保温, 所述大肠杆菌克隆含有与化脓链球菌肽融合的平台蛋白的编码质粒, 然后用荧光素标记二抗进行检测。如图 5A 中所示, 针对化脓链球菌裂解产物培养的抗血清引起了化脓链球菌 SF370/M1 细胞群荧光的漂移。用针对抗原 Spy0012 (图 5B)、Spy1315 和 Spy1798 (图 5C) 的肽培养的多克隆血清观察到化脓链球菌 SF370/M1 细胞的相似细胞表面染色, 尽管只有该细菌的一个亚群被染色, 正如用两个峰的检测所指出的。这一现象可能是细菌生长期间基因产物分化表达或由其它表面分子引起的抗体结合的部分抑制作用的结果。

[0367] 这些实验证实了生物信息的预测,即,这些蛋白质由于其信号肽序列而被输出且还显示出它们锚定于化脓链球菌 SF370/M1 的细胞表面。它们还证实了这些蛋白质可被人抗体识别并使它们成为抗 A 族链球菌疾病疫苗开发的有价值候选分子。

[0368] 实施例 7:基于重组抗原免疫接种的抗 A 族链球菌感染的保护性免疫应答。

[0369] 实验步骤

[0370] 编码抗原蛋白质的基因的克隆

[0371] 用基因特异引物通过 PCR 扩增反应扩增来自化脓链球菌 SF370 基因组 DNA 的目的基因或 DNA 片段。除了基因特异序列之外,引物在 5' 端含有由限制性位点组成帮助扩增的 PCR 产物定向克隆的额外碱基。引物的基因特异序列长度范围在 15-24 个碱基之间。所包含的 PCR 产物用适当的限制性酶消化并克隆入适当的已消化 pET28b(+) 载体 (NOVAGEN) 中。重组质粒的构建被证实后,转化用作表达宿主的大肠杆菌 BL21 **STAR**[®] 细胞 (INVITROGEN)。将这些细胞最优化以有效表达 pET28b 质粒所编码的目的基因。

[0372] 大肠杆菌中的抗原表达

[0373] 将含重组质粒的大肠杆菌 BL21 **STAR**[®] 细胞在补充有 50 μ g/ml 卡那霉素的 LB 培养基中 37°C 培养至对数期。一旦 OD_{600nm} 达到 0.8,培养物用 1mM IPTG 在 37°C 诱导 3 小时。离心收获这些细胞,通过冻融方法及随后通过用来自 NOVAGEN 的 Bug-**buster**[®] 试剂破坏细胞联合裂解细菌。裂解产物通过离心分离成可溶性(上清液)和不溶性(沉淀)部分。

[0374] 来自大肠杆菌的重组蛋白质的纯化

[0375] 根据蛋白质的定位,采取不同的纯化策略。通过在细胞被破坏后将细胞裂解产物的上清液与镍-琼脂糖珠 (Ni-NTA-**Agarose**[®], QIAGEN) 结合来纯化可溶性组分里的蛋白质。由于在被表达蛋白质的 C 末端、N 末端或两个末端都存在的五组氨酸 (HIS),所述蛋白质与镍-琼脂糖结合而其它杂质蛋白质通过洗涤缓冲液从柱子上洗去和清除掉。用含有溶于适当缓冲液中的 100mM 咪唑的溶液洗脱所述蛋白质。浓缩洗脱液,用 Bradford 检测蛋白质浓度并通过 SDS-PAGE 和蛋白质印迹法进行分析。通过沉淀在含 8M 尿素的适当缓冲液中的增溶作用纯化不溶性部分的蛋白质。在用同一材料变性的条件下(在含 8M 尿素的缓冲液中)进行纯化,步骤与上述针对可溶性蛋白质的步骤一样。浓缩洗脱液并以渐进或逐步的方式透析除去所有尿素。浓缩最终的蛋白质溶液,通过 SDS-PAGE 进行分析并用 Bradford 方法进行检测。当检测到相当于预计分子量和电泳迁移率的特异条带时认为表达得以证实。对于在透析阶段由于变性试剂尿素的去除而沉淀的蛋白质而言,可将不溶性包含体漂洗数次并直接用于小鼠的免疫接种。

[0376] 用重组蛋白质免疫 NMRI 小鼠并用化脓链球菌 AP1 攻击

[0377] 利用 NMRI 小鼠和作为感染试剂的化脓链球菌菌株 AP1 在实验动物模型中检测蛋白质的免疫原性。用 50 μ g/剂重组蛋白质每 2 周免疫一次 7-8 周大的雌性 NMRI 小鼠,总共 3 次。最初施用的剂量以完全弗氏佐剂辅助,而余下的两剂用不完全弗氏佐剂辅助。免疫接种结束后将小鼠放血以检查抗体效价,随后静脉注射 (i. v.) 致死剂量的化脓链球菌 AP1 (5×10^7 个致病菌) 激发免疫。在激发免疫后对小鼠的生存状况记录 18-21 天。

[0378] 结果

[0379] 重组蛋白质的表达和纯化

[0380] 在选定用于重组蛋白质表达的 31 种蛋白质之中,29 种蛋白质在大肠杆菌中产生

的水平可足以纯化。而某些蛋白质可作为可溶性蛋白质产生（见表 4），某些蛋白质转变成不溶性的（如 Spy416B、Spy0872）或由于透析而沉淀，它们在诸如 Spy0031、Spy0292、Spy720 等不溶性蛋白质的增溶作用后试图去除变性试剂尿素。在这些情形中被漂洗的包含体直接被注射入小鼠中进行免疫接种。通常，亲和纯化产生了至少 85% 纯度的重组蛋白质制剂。

[0381] 在 NMRI 小鼠中用重组蛋白质免疫后的免疫应答

[0382] 表 4 列举了那些抗原，它们在小鼠中被测试并在实验动物内显示出某种程度的保护作用。同样在动物菌血症模型中检测到但在所述实验中未显示出任何水平的保护作用的重组蛋白质在此未列举，但包括诸如 Spy0012、Spy1063 和 Spy1494 等蛋白质。当致病细菌被直接注射入血液中时，上述菌血症模型评估了针对入侵疾病的候选疫苗的保护性意义。所诱导抗体能起到针对所述 A 族链球菌感染的保护作用的重组蛋白质被认为是用于研发抗 A 族链球菌疾病的疫苗的价值候选分子。与先前显示出提供了针对化脓链球菌攻击的保护作用的阳性对照 Spy2018（M1 蛋白）相比较而言，在 18-21 天后（表 4）评估激发实验的终点时许多抗原表现出相似程度的保护作用（Spy0416、Spy1607 或 Spy0292）。其它的蛋白质只显示出部分的保护效果（Spy0720、Spy0872），但当与其它抗原联合时可被证实非常有效（图 6）。

[0383] 令人吃惊的是，抗原筛选已鉴定了主要在两种较大蛋白质 Spy0416 和 Spy1972 前半部分的免疫原性表位。因此可推断保护区还可能包含于蛋白质的 N 末端部分。以 Spy0416 为例，抗原的两部分都作为重组蛋白质（Spy0416A 和 Spy0416B；见表 4）产生并测试于动物实验中。实验显示在动物模型中只有蛋白质 Spy0416（表 4；Spy0416A）的前半部分提供了保护作用，而蛋白质的后半部分（Spy0416B）根本无保护效果，明显勾画出作为候选疫苗的是蛋白质内的较小区域。对于抗原 Spy1972 而言，只有全长蛋白质的前半部分被作为重组蛋白质产生并在动物模型中进行测试。

[0384] 实施例 8：编码多种血清型化脓链球菌中抗原蛋白质的基因的可变性

[0385] 实验步骤

[0386] PCR 片段的测序和生物信息学分析

[0387] 化脓链球菌菌株的 PCR 分析参阅实施例 5。PCR 片段的测序提供了对基因可变性的一个估计，结果概述于表 3 中。来自五种化脓链球菌菌株（SF370：M1；MGAS8232：M18；SSI-1：M3；MGAS315：M3；Manfredo：M5）的基因组序列的可用性使得我们可以进一步的评估所述抗原的可变性。将所有序列与来自化脓链球菌 SF370 的相应抗原序列和经鉴定不同于化脓链球菌 SF370 抗原蛋白质氨基酸残基的那些氨基酸残基进行序列比对。在某些抗原蛋白质中检测到了插入或缺失序列，但不包括在此分析中。

[0388] 结果

[0389] 表 5 显示了经鉴定在四种化脓链球菌菌株（MGAS8232：M18；SSI-1：M3；MGAS315：M3；Manfredo：M5）之一或用于 PCR 扩增片段测序的菌株（见表 3）中指定的抗原内可变的的所有位置。生物信息分析显示，一些抗原蛋白质非常保守，在 6 种血清型株系 M1、M3、M5、M18 和 M89 中的任一种都无单一氨基酸交换。属于此类的蛋白质包括 Spy0103 和 Spy1536，而在较大蛋白质内的其它抗原蛋白质中的交换数目比在较小蛋白质内的要大，正如从其大小差异所预期的一样。尽管分析了各种菌株，但几乎从未观察到单个残基在其它菌株中被改变成一个以上的其它氨基酸。对更大数目不同血清型菌株中相应基因序列、临床指征或地理

定位的进一步分析可肯定地确认在所列举的那些氨基酸残基或另外的残基中可能发生的变化。

[0390] 通过 6 个基因序列比对分析的抗原蛋白质中只有一个显示出大小有相当程度的变化 (Spy1357 :SF370-217 氨基酸 ;MGAS8232-245 氨基酸 ;SSI-1-329 氨基酸 ;MGAS315-329 氨基酸 ;Manfredo-279 氨基酸)。因此显而易见,大部分被评估的抗原在序列以及大小上是非常保守的,且为疫苗研发提供了有希望的候选对象。

[0391] 参考文献

- [0392] Altschul, S., 等 (1990). Journal of Molecular Biology 215 :403-10.
- [0393] Bennett, D., 等 (1995). J Mol Recognit 8 :52-8.
- [0394] Bessen, D., 等 (1988). Infect Immun 56 :2666-2672.
- [0395] Bisno, A., 等 (1987). Infect Immun 55 :753-7.
- [0396] Bronze, M., 等 (1988). J Immunol 141 :2767-2770.
- [0397] Clackson, T., 等 (1991). Nature 352 :624-8.
- [0398] Cone, L., 等 (1987). New Engl J Med 317 :146-9.
- [0399] Cunningham, M. (2000). Clin Microbiol Rev 13 :470-511.
- [0400] Devereux, J., 等 (1984). Nucleic acids research 12 :387-95.
- [0401] Doherty, E., 等 (2001). Annu Rev Biophys Biomol Struct 30 :457-475.
- [0402] Eisenbraun, M., 等 (1993). DNA Cell Biol 12 :791-7.
- [0403] Enright M., 等 (2001) Inf. Immun. 69 :2416-27
- [0404] Etz, H., 等 (2001). J Bacteriol 183 :6924-35.
- [0405] Fenderson, P., 等 (1989). J Immunol 142 :2475-2481.
- [0406] Fischetti, V. (1989). Clin Microbiol Rev 2 :285-314.
- [0407] Ganz, T. (1999). Science 286 :420-421.
- [0408] Georgiou, G. (1997). Nature Biotechnology 15 :29-34.
- [0409] Guzman, C., 等 (1999). J Infect Dis 179 :901-6.
- [0410] Hashemzadeh-Bonehi, L., 等 (1998). Mol Microbiol 30 :676-678.
- [0411] Heinje, von G, (1987) e. g. 分子生物学中的序列分析 Sequence Analysis in Molecular Biology, Academic Press
- [0412] Hemmer, B., 等 (1999). Nat Med 5 :1375-82.
- [0413] Hoe N., 等 (2001) J. Inf. Dis. 183 :633-9
- [0414] Hope-Simpson, R. (1981). J Hyg(Lond) 87 :109-29.
- [0415] Ji, Y., 等 (1997). Infect Immun 65 :2080-2087.
- [0416] Johanson, K., 等 (1995). J Biol Chem 270 :9459-71.
- [0417] Jones, P., 等 (1986). Nature 321 :522-5.
- [0418] Kajava, A., 等 (2000). J Bacteriol 182 :2163-9.
- [0419] Kohler, G., 等 (1975). Nature 256 :495-7.
- [0420] Kolaskar, A., 等 (1990). FEBS Lett 276 :172-4.
- [0421] Lee, P. K. (1989). J Clin Microbiol 27 :1890-2.
- [0422] Lewin, A., 等 (2001). Trends Mol Med 7 :221-8.

- [0423] Marks, J., 等 (1992). Biotechnology(NY) 10 :779-83.
- [0424] McCafferty, J., 等 (1990). Nature 348 :552-4.
- [0425] Okano, H., 等 (1991). J Neurochem 56 :560-7.
- [0426] 作为基因表达反义抑制物的寡脱氧核苷酸 Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression ;CRC Press, Boca Ration, FL(1988) 用来描述这些分子
- [0427] Phillips-Quagliata, J., 等 (2000). J Immunol 165 :2544-55.
- [0428] Rammensee, H., 等 (1999). Immunogenetics 50 :213-9.
- [0429] Rosenshine, I., 等 (1992). Infect Immun 60 :2211-7.
- [0430] Seeger, C., 等 (1984). Proc Natl Acad Sci USA 81 :5849-52.
- [0431] Shibuya, A., 等 (2000). Nature Immunology 1 :441-6.
- [0432] Skerra, A. (1994). Gene 151 :131-5.
- [0433] Stevens, D. (1992). Clin Infect Dis 14 :2-11.
- [0434] Tang, D., 等 (1992). Nature 356 :152-4.
- [0435] Tempest, P., 等 (1991). Biotechnology(NY) 9 :266-71.
- [0436] Tourdot, S., 等 (2000). Eur J Immunol 30 :3411-21.
- [0437] Whitnack, E, 等 (1985). J Exp Med 162 :1983-97.
- [0438] Wiley, J., 等 (1987) Current Protocols in Molecular Biology
- [0439] Vitali, L., 等 (2002) J. Clin. Microbiol 40 :679-681
- [0440] 表 1 :通过细菌表面展示鉴定的免疫原性蛋白质
- [0441]

化脓链 球菌抗 原蛋白	推定的功能 (通过同源 性)	预测的免疫原性氨基酸**	每次 ORF 和筛选选 定克隆的 编号	已鉴定 免疫原 性区的 位置 (氨基 酸)	序列识 别号 (DNA , 蛋白 质)
Spy0012	假定的蛋白 质	4-44, 57-65, 67-98, 101-107, 109-125, 131-144, 146-159, 168-173, 181-186, 191-200, 206-213, 229-245, 261-269, 288-301, 304-317, 323-328, 350-361, 374-384, 388-407, 416-425	A :12, I :5, N :2	1-114	1, 151
		5-17, 49-64, 77-82, 87-98, 118-125, 127-140, 142-150, 153-159, 191-207, 212-218, 226-270, 274-287, 297-306, 325-331, 340-347, 352-369, 377-382, 390-395	F :2, I :16, K :24, N :29, P :12	29-226	2, 152
Spy0019	推定的分泌 蛋白质 (细 胞分裂和抗 生素耐受 性)				

[0442]

Spy0025	推定的磷酸 核糖甲酰甘 氨脒合酶 II	4-16, 20-26, 32-74, 76-87, 93-108, 116-141, 148-162, 165-180, 206-219, 221-228, 230-236, 239-245, 257-268, 313-328, 330-335, 353-359, 367-375, 394-403, 414-434,	D : 3	919-929	3, 153

化脓链 球菌抗 原蛋白	推定的功能 (通过同源 性)	预测的免疫原性氨基酸 ^{aa}	每次 ORF 和筛选选 定克隆的 编号	已鉴定 免疫原 性区的 位置 (氨基 酸)	序列识 别号 (DNA , 蛋白 质)
		437-444, 446-453, 456-464, 478-487, 526-535, 541-552, 568-575, 577-584, 589-598, 610-618, 624-643, 653-665, 667-681, 697-718, 730-748, 755-761, 773-794, 806-821, 823-831, 837-845, 862-877, 879-889, 896-919, 924-930, 935-940, 947-955, 959-964, 969-986, 991-1002, 1012-1036, 1047-1056, 1067-1073, 1079-1085, 1088-1111, 1130-1135, 1148-1164, 1166-1173, 1185-1192, 1244-1254			

[0443]

Spy0031	推定的胆碱结合蛋白	5-44, 62-74, 78-83, 99-105, 107-113, 124-134, 161-174, 176-194, 203-211, 216-237, 241-247, 253-266, 272-299, 323-349, 353-360	I : 3, K : 3, N : 3	145-305	4, 154
Spy0103	推定的感受态蛋白质	15-39, 52-61, 72-81, 92-97	A : 8	71-81	5, 155

化脓链 球菌抗 原蛋白	推定的功能 (通过同源 性)	预测的免疫原性氨基酸 **	每次 ORF 和筛选选 定克隆的 编号	已鉴定 免疫原 性区的 位置 (氨基酸)	序列识 别号 (DNA , 蛋白 质)
Spy0112	推定的吡咯 啉羧酸盐还 原酶	13-19, 21-31, 40-108, 115-122, 125-140, 158-180, 187-203, 210-223, 235-245	B : 4	173-186	6, 156
Spy0115	推定的谷氨 酰 - 氨基肽 酶	5-12, 19-27, 29-39, 59-67, 71-78, 80-88, 92-104, 107-124, 129-142, 158-168, 185-191, 218-226, 230-243, 256-267, 272-277, 283-291, 307-325, 331-344, 346-352	A : 3, C : 26	316-331	7, 157
Spy0166	假设的蛋白 质	6-28, 43-53, 60-76, 93-103	I : 22, K : 7, N : 17, O : 31, P : 5	21-99	8, 158
Spy0167	链球菌溶血 素 O	10-30, 120-126, 145-151, 159-169, 174-182, 191-196, 201-206, 214-220, 222-232, 254-272, 292-307, 313-323, 332-353, 361-369, 389-396, 401-415, 428-439, 465-481, 510-517, 560-568	A : 118, B : 14, C : 18, D : 37, F : 141, G : 79, H : 92, I : 97, K : 123, L : 5, M : 21, N : 225,	9-264	9, 159

[0444]

化脓链 球菌抗 原蛋白	推定的功能 (通过同源 性)	预测的免疫原性氨基酸**	每次 ORF 和筛选选 定克隆的 编号	已鉴定 免疫原 性区的 位置 (氨基 酸)	序列识 别号 (DNA , 蛋白 质)
			O :230, P :265		
Spy0168	假设的蛋白 质	5-29, 39-45, 107-128	K :4, N :7	1-112	10, 160
Spy0171	假设的蛋白 质	4-38, 42-50, 54-60, 65-71, 91-102	H :2	21-56	11, 161
Spy0183	推定的甘氨 酸甜菜碱 / 脯氨酸 ABC 转运 蛋白	4-13, 19-25, 41-51, 54-62, 68-75, 79-89, 109-122, 130-136, 172-189, 192-198, 217-224, 262-268, 270-276, 281-298, 315-324, 333-342, 353-370, 376-391	C :6	23-39	12, 162

Spy0230	推定的 ABC 转运 蛋白 (ATP - 结合蛋 白)	6-41, 49-58, 62-103, 117-124, 147-166, 173-194, 204-211, 221-229, 255-261, 269-284, 288-310, 319-325, 348-380, 383-389, 402-410, 424-443, 467-479, 496-517, 535-553, 555-565, 574-581, 583-591	C :46	474-489	13, 163
Spy0269	推定的表面 排阻蛋白	8-35, 52-57, 66-73, 81-88, 108-114, 125-131, 160-167, 174-180, 230-235, 237-249,	A :2, B :12, D :3, F :11, H :5, N :6	37-241 409-534 582-604	14, 164

[0445]

化脓链 球菌抗 原蛋白	推定的功能 (通过同源 性)	预测的免疫原性氨基酸**	每次 ORF 和筛选选 定克隆的 编号	已鉴定 免疫原 性区的 位置 (氨基酸)	序列识 别号 (DNA , 蛋白 质)
		254-262, 278-285, 308-314, 321-326, 344-353, 358-372, 376-383, 393-411, 439-446, 453-464, 471-480, 485-492, 502-508, 523-529, 533-556, 558-563, 567-584, 589-597, 605-619, 625-645, 647-666, 671-678, 690-714, 721-728, 741-763, 766-773, 777-787, 792-802, 809-823, 849-864		743-804	
Spy0287	保守的假设 蛋白质	4-17, 24-36, 38-44, 59-67, 72-90, 92-121, 126-149, 151-159, 161-175, 197-215, 217-227, 241-247, 257-264, 266-275, 277-284, 293-307, 315-321, 330-337, 345-350, 357-366, 385-416	K :1	202-337	15, 165
Spy0292	青霉素结合 蛋白 (D- 丙 氨酰 -D- 丙 氨酸 car	4-20, 22-46, 49-70, 80-89, 96-103, 105-119, 123-129, 153-160, 181-223, 227-233, 236-243, 248-255, 261-269, 274-279, 283-299, 305-313,	F :2	1-48	16, 166

[0446]

化脓链 球菌抗 原蛋白	推定的功能 (通过同源 性)	预测的免疫原性氨基酸**	每次 ORF 和筛选选 定克隆的 编号	已鉴定 免疫原 性区的 位置 (氨基 酸)	序列识 别号 (DNA , 蛋白 质)
		315-332, 339-344, 349-362, 365-373, 380-388, 391-397, 402-407			
Spy0295	寡肽透性酶	18-37, 41-63, 100-106, 109-151, 153-167, 170-197, 199-207, 212-229, 232-253, 273-297	A :3	203-217	17, 167
Spy0348	推定的氨基 脱氧 chorismate 裂合酶	20-26, 54-61, 80-88, 94-101, 113-119, 128-136, 138-144, 156-188, 193-201, 209-217, 221-229, 239-244, 251-257, 270-278, 281-290, 308-315, 319-332, 339-352, 370-381, 388-400, 411-417, 426-435, 468-482, 488-497, 499-506, 512-521	D :5, I :3, M :3, P :3	261-273	18, 168
Spy0416	推定的细胞 外壳丝氨酸 蛋白酶	6-12, 16-36, 50-56, 86-92, 115-125, 143-152, 163-172, 193-203, 235-244, 280-289, 302-315, 325-348, 370-379, 399-405, 411-417, 419-429, 441-449, 463-472, 482-490,	A :3, B :4, C :30, D :13, F :138, G :120, H :101, I :9,	1-414 443-614 997-139 2	19, 169

[0447]

化脓链 球菌抗 原蛋白	推定的功能 (通过同源 性)	预测的免疫原性氨基酸**	每次 ORF 和筛选选 定克隆的 编号	已鉴定 免疫原 性区的 位置 (氨基 酸)	序列识 别号 (DNA ,蛋白 质)

			500-516, 536-543, 561-569, 587-594, 620-636, 647-653, 659-664, 677-685, 687-693, 713-719, 733-740, 746-754, 756-779, 792-799, 808-817, 822-828, 851-865, 902-908, 920-938, 946-952, 969-976, 988-1005, 1018-1027, 1045-1057, 1063-1069, 1071-1078, 1090-1099, 1101-1109, 1113-1127, 1130-1137, 1162-1174, 1211-1221, 1234-1242, 1261-1268, 1278-1284, 1312-1317, 1319-1326, 1345-1353, 1366-1378, 1382-1394, 1396-1413, 1415-1424, 1442-1457, 1467-1474, 1482-1490, 1492-1530, 1537-1549, 1559-1576, 1611-1616, 1624-1641	K : 14, M : 2, N : 15, O : 8, P : 19		

[0448]

化脓链 球菌抗 原蛋白	推定的功能 (通过同源 性)	预测的免疫原性氨基酸 **	每次 ORF 和筛选选 定克隆的 编号	已鉴定 免疫原 性区的 位置 (氨基酸)	序列识 别号 (DNA , 蛋白 质)
Spy0430	假设的蛋白 质	14-42, 70-75, 90-100, 158-181	B : 7, I : 10, P : 18	1-164	20, 170
Spy0433	假设的蛋白 质	4-21, 30-36, 54-82, 89-97, 105-118, 138-147	A : 138, B : 8, C : 67, D : 11, E : 13, F : 35, G : 10, H : 5, M : 8	126-207	21, 171
Spy0437	假设的蛋白 质	4-21, 31-66, 96-104, 106-113, 131-142	A : 29, B : 10, C : 21, D : 24, E : 15	180-204	22, 172
Spy0469	推定的 42 kDa 蛋白	5-23, 31-36, 38-55, 65-74, 79-88, 101-129, 131-154, 156-165, 183-194, 225-237, 245-261, 264-271, 279-284, 287-297, 313-319, 327-336, 343-363, 380-386	B : 5, F : 77, I : 8, K : 15, M : 3, N : 17, O : 20	11-197 204-219 258-372	23, 173
Spy0488	假设的蛋白 质	4-20, 34-41, 71-86, 100-110, 113-124, 133-143, 150-158, 160-166, 175-182, 191-197, 213-223, 233-239, 259-278,	A : 17, B : 11, C : 23, D : 12, E : 4,	195-289	24, 174

[0449]

化脓链 球菌抗 原蛋白	推定的功能 (通过同源 性)	预测的免疫原性氨基酸 **	每次 ORF 和筛选选 定克隆的 编号	已鉴定 免疫原 性区的 位置 (氨基酸)	序列识 别号 (DNA , 蛋白 质)
		298-322	G : 4, H : 7		

Spy0515	推定的糖转移酶	4-10, 21-35, 44-52, 54-62, 67-73, 87-103, 106-135, 161-174, 177-192, 200-209, 216-223, 249-298, 304-312, 315-329	B ;5, I ;3	12-130	25, 175
Spy0580	保守的假定蛋白质	10-27, 33-38, 48-55, 70-76, 96-107, 119-133, 141-147, 151-165, 183-190, 197-210, 228-236, 245-250, 266-272, 289-295, 297-306, 308-315, 323-352, 357-371, 381-390, 394-401, 404-415, 417-425, 427-462, 466-483, 485-496, 502-507, 520-529, 531-541, 553-570, 577-588, 591-596, 600-610, 619-632, 642-665, 671-692, 694-707	C ;5	434-444	26, 176
Spy0621	保守的假定蛋白质	6-14, 16-25, 36-46, 52-70, 83-111, 129-138, 140-149, 153-166, 169-181, 188-206, 212-220, 223-259, 261-269,	C ;3	360-375	27, 177

[0450]

化脓链球菌抗原蛋白	推定的功能 (通过同源性)	预测的免疫原性氨基酸**	每次 ORF 和筛选选定克隆的 编号	已鉴定 免疫原性区的 位置 (氨基酸)	序列识别号 (DNA, 蛋白质)
		274-282, 286-293, 297-306, 313-319, 329-341, 343-359, 377-390, 409-415, 425-430			

Spy0630	推定的 PTS 依赖型 N- 乙酰 - 半乳 糖胺 -IIC	4-26, 28-48, 54-62, 88-121, 147-162, 164-201, 203-237, 245-251	C :2	254-260	28, 178
Spy0681	假设的蛋白 质, 与噬菌 体有关	12-21, 26-32, 66-72, 87-93, 98-112, 125-149, 179-203, 209-226, 233-242, 249-261, 266-271, 273-289, 293-318, 346-354, 360-371, 391-400	A :8	369-382	29, 179
Spy0683	假设的蛋白 质, 与噬菌 体有关	11-38, 44-65, 70-87, 129-135, 140-163, 171-177, 225-232, 238-249, 258-266, 271-280, 284-291, 295-300, 329-337, 344-352, 405-412, 416-424, 426-434, 436-455, 462-475, 478-487	B :11, D :4	270-312	30, 180
Spy0702	假设的蛋白 质	5-17, 34-45, 59-69, 82-88, 117-129, 137-142, 158-165, 180-195, 201-206, 219-226, 241-260, 269-279, 292-305,	L :2	486-598	31, 181

[0451]

化脓链 球菌抗 原蛋白	推定的功能 (通过同源 性)	预测的免疫原性氨基酸**	每次 ORF 和筛选选 定克隆的 编号	已鉴定 免疫原 性区的 位置 (氨基 酸)	序列识 别号 (DNA , 蛋白 质)
		312-321, 341-347, 362-381, 396-410, 413-432, 434-445, 447-453, 482-487, 492-499, 507-516, 546-552, 556-565, 587-604			
Spy0710	保守的假设 蛋白质, 与 噬菌体有关	4-15, 17-32, 40-47, 67-78, 90-98, 101-107, 111-136, 161-171, 184-198, 208-214, 234-245, 247-254, 272-279, 288-298, 303-310, 315-320, 327-333, 338-349, 364-374	B:10	378-396	32, 182
Spy0711	引起发热的外毒素前 体, 与噬菌体有关 (speC)	5-27, 33-49, 51-57, 74-81, 95-107, 130-137, 148-157, 173-184	K:2	75-235	33, 183
Spy0720	保守的假定 蛋白质	6-23, 47-53, 57-63, 75-82, 97-105, 113-122, 124-134, 142-153, 159-164, 169-179, 181-187, 192-208, 215-243, 247-257, 285-290, 303-310	D:2	30-51	34, 184
Spy0727	推定的 DNA 旋转 酶亚基 B	17-29, 44-52, 59-73, 77-83, 86-92, 97-110, 118-153, 156-166, 173-178, 182-200	M:26	208-219	35, 185

[0452]

化脓链 球菌抗 原蛋白	推定的功能 (通过同源 性)	预测的免疫原性氨基酸 ³⁸⁴	每次 ORF 和筛选选 定克隆的 编号	已鉴定 免疫原 性区的 位置 (氨基 酸)	序列识 别号 (DNA , 蛋白 质)
		156-166, 173-179, 192-209, 225-231, 234-240, 245-251, 260-268, 274-279, 297-306, 328-340, 353-360, 369-382, 384-397, 414-423, 431-436, 452-465, 492-498, 500-508, 516-552, 554-560, 568-574, 580-586, 609-617, 620-626, 641-647			

[0453]

Spy0737	推定的胞外 基质结合蛋 白	4-26, 32-45, 58-72, 111-119, 137-143, 146-159, 187-193, 221-231, 235-242, 250-273, 290-304, 311-321, 326-339, 341-347, 354-368, 397-403, 412-419, 426-432, 487-506, 580-592, 619-628, 663-685, 707-716, 743-751, 770-776, 787-792, 850-859, 866-873, 882-888, 922-931, 957-963, 975-981, 983-989, 1000-1008, 1023-1029, 1058-1064, 1089-1099, 1107-1114,	B : 5, E : 3, K : 11	396-533 1342-15 02 1672-19 20	36, 186

化脓链 球菌抗 原蛋白	推定的功能 (通过同源 性)	预测的免疫原性氨基酸 ^{seq}	每次 ORF 和筛选选 定克隆的 编号	已鉴定 免疫原 性区的 位置 (氨基 酸)	序列识 别号 (DNA , 蛋白 质)
		1139-1145, 1147-1156, 1217-1226, 1276-1281, 1329-1335, 1355-1366, 1382-1394, 1410-1416, 1418-1424, 1443-1451, 1461-1469, 1483-1489, 1491-1501, 1515-1522, 1538-1544, 1549-1561, 1587-1593, 1603-1613, 1625-1630, 1636-1641, 1684-1690, 1706-1723, 1765-1771, 1787-1804, 1850-1857, 1863-1894, 1897-1910, 1926-1935, 1937-1943, 1960-1983, 1991-2005, 2008-2014, 2018-2039			

[0454]

Spy0747	细胞外核酸 酶	4-25, 45-50, 53-65, 79-85, 87-92, 99-109, 126-137, 141-148, 156-183, 190-203, 212-217, 221-228, 235-242, 247-277, 287-293, 300-319,	A : 72, B : 17, H : 6, O : 3	1-113 210-232 250-423 536-564	37, 187

化脓链 球菌抗 原蛋白	推定的功能 (通过同源 性)	预测的免疫原性氨基酸**	每次 ORF 和筛选选 定克隆的 编号	已鉴定 免疫原 性区的 位置 (氨基酸)	序列识 别号 (DNA , 蛋白 质)
		321-330, 341-361, 378-389, 394-406, 437-449, 455-461, 472-478, 482-491, 507-522, 544-554, 576-582, 587-593, 611-621, 626-632, 649-661, 679-685, 696-704, 706-716, 726-736, 740-751, 759-766, 786-792, 797-802, 810-822, 824-832, 843-852, 863-869, 874-879, 882-905			
Spy0777	推定的 ATP- 依赖 型核酸外切 酶, 亚基 A	4-16, 33-39, 43-49, 54-85, 107-123, 131-147, 157-169, 177-187, 198-209, 220-230, 238-248, 277-286, 293-301, 303-315, 319-379, 383-393, 402-414, 426-432, 439-449, 470-478, 483-497, 502-535, 552-566, 571-582, 596-601, 608-620, 631-643, 651-656, 663-678, 680-699, 705-717, 724-732, 738-748, 756-763, 766-772, 776-791, 796-810,	C :4, E :2	617-635	38, 188

[0455]

化脓链 球菌抗 原蛋白	推定的功能 (通过同源 性)	预测的免疫原性氨基酸 ^{seq}	每次 ORF 和筛选选 定克隆的 编号	已鉴定 免疫原 性区的 位置 (氨基 酸)	序列识 别号 (DNA , 蛋白 质)
		819-827, 829-841, 847-861, 866-871, 876-882, 887-894, 909-934, 941-947, 957-969, 986-994, 998-1028, 1033-1070, 1073-1080, 1090-1096, 1098-1132, 1134-1159, 1164-1172, 1174-1201			
Spy0789	推定的 ABC- 转运 蛋白 (透性 酶蛋白)	7-25, 30-40, 42-64, 70-77, 85-118, 120-166, 169-199, 202-213, 222-244	A : 3	190-203	39, 189

[0456]

Spy0839	推定的甘油 磷酸二酯	4-11, 15-53, 55-93, 95-113, 120-159, 164-200, 210-243, 250-258, 261-283, 298-319, 327-340, 356-366, 369-376, 380-386, 394-406, 409-421, 425-435, 442-454, 461-472, 480-490, 494-505, 507-514, 521-527, 533-544, 566-574	A : 7, D : 2	385-398	40, 190
Spy0843	细胞表面蛋白	5-36, 66-72, 120-127, 146-152, 159-168, 172-184, 205-210, 221-232, 234-243, 251-275,	A : 11, B : 3, C : 5, D : 4, F : 50,	12-190 276-283 666-806	41, 191

化脓链 球菌抗 原蛋白	推定的功能 (通过同源 性)	预测的免疫原性氨基酸**	每次 ORF 和筛选选 定克隆的 编号	已鉴定 免疫原 性区的 位置 (氨基 酸)	序列识 别号 (DNA , 蛋白 质)
		295-305, 325-332, 367-373, 470-479, 482-487, 520-548, 592-600, 605-615, 627-642, 655-662, 664-698, 718-725, 734-763, 776-784, 798-809, 811-842, 845-852, 867-872, 879-888, 900-928, 933-940, 972-977, 982-1003	H :19, G :49, I :112, K :102, L :10, M :3, N :213, O :188, P :310		

Spy0872	推定被分泌 的 5' - 核苷 酸酶	4-38, 63-68, 100-114, 160-173, 183-192, 195-210, 212-219, 221-238, 240-256, 258-266, 274-290, 301-311, 313-319, 332-341, 357-363, 395-401, 405-410, 420-426, 435-450, 453-461, 468-475, 491-498, 510-518, 529-537, 545-552, 585-592, 602-611, 634-639, 650-664	A :6, D :2, F :5, H :14, I :9, K :10, L :1, N :16, O :12	30-80 89-105 111-151	42, 192
Spy0895	组氨酸蛋白 质激酶	7-29, 31-39, 47-54, 63-74, 81-94, 97-117, 122-127, 146-157, 168-192, 195-204, 216-240, 251-259	C :11	195-203	43, 193

[0457]

化脓链 球菌抗 原蛋白	推定的功能 (通过同源 性)	预测的免疫原性氨基酸**	每次 ORF 和筛选选 定克隆的 编号	已鉴定 免疫原 性区的 位置 (氨基 酸)	序列识 别号 (DNA , 蛋白 质)
Spy0972	推定的末端 酶大亚基 - 噬菌体	5-16, 28-34, 46-65, 79-94, 98-105, 107-113, 120-134, 147-158, 163-172, 180-186, 226-233, 237-251, 253-259, 275-285, 287-294, 302-308, 315-321, 334-344, 360-371, 399-412, 420-426	B :2	32-50	44, 194
Spy0981	假设的蛋白 质 - 噬菌体 相关	8-20, 30-36, 71-79, 90-96, 106-117, 125-138, 141-147, 166-174	A :7, B :2	75-90	45, 195
Spy1008	链球菌外毒 素 H 前体 (speH)	4-13, 15-33, 43-52, 63-85, 98-114, 131-139, 146-174, 186-192, 198-206, 227-233	C :11	69-88	46, 196

[0458]

Spy1032	细胞外透明 质酸裂合酶	4-22, 29-35, 59-68, 153-170, 213-219, 224-238, 240-246, 263-270, 285-292, 301-321, 327-346, 356-371, 389-405, 411-418, 421-427, 430-437, 450-467, 472-477, 482-487, 513-518, 531-538, 569-576, 606-614, 637-657, 662-667, 673-690, 743-753, 760-767,	B :3, K :3, M :5	96-230 361-491 572-585	47, 197

化脓链 球菌抗 原蛋白	推定的功能 (通过同源 性)	预测的免疫原性氨基酸**	每次 ORF 和筛选选 定克隆的 编号	已鉴定 免疫原 性区的 位置 (氨基酸)	序列识 别号 (DNA , 蛋白 质)
		770-777, 786-802			
Spy1054	推定的胶原 样蛋白质 (Sc1C)	4-12, 21-36, 48-55, 74-82, 121-127, 195-203, 207-228, 247-262, 269-278, 280-289	A :71, B :13, C :233, D :41, E :163, F :200, G :442, H :129, N :3	102-210	48, 198
Spy1063	推定的外周 - 铁 - 结合 蛋白	13-20, 23-31, 38-44, 78-107, 110-118, 122-144, 151-164, 176-182, 190-198, 209-216, 219-243, 251-256, 289-304, 306-313	A :4	240-248	49, 199
Spy1162	推定的核糖 核酸酶 HII	5-26, 34-48, 57-77, 84-102, 116-132, 139-145, 150-162, 165-173, 176-187, 192-205, 216-221, 234-248, 250-260	B :3, C :5	182-198	50, 200
Spy1206	推定的 ABC 转运 蛋白	10-19, 26-44, 53-62, 69-87, 90-96, 121-127, 141-146, 148-158, 175-193, 204-259,	A :2	41-56	51, 201

[0459]

化脓链 球菌抗 原蛋白	推定的功能 (通过同源 性)	预测的免疫原性氨基酸**	每次 ORF 和筛选 定克隆的 编号	已鉴定 免疫原 性区的 位置 (氨基 酸)	序列识 别号 (DNA , 蛋白 质)
		307-313, 334-348, 360-365, 370-401, 411-439, 441-450, 455-462, 467-472, 488-504			
Spy1228	推定的脂蛋 白	5-21, 36-42, 96-116, 123-130, 138-144, 146-157, 184-201, 213-228, 252-259, 277-297, 308-313, 318-323, 327-333	M: 33	202-217	52, 202
Spy1245	推定的磷酸 盐 ABC 转 运蛋白	6-26, 33-51, 72-90, 97-131, 147-154, 164-171, 187-216, 231-236, 260-269, 275-283	I: 3, K: 3	1-127	53, 203

[0460]

Spy1315	假定的蛋白 质	4-22, 24-38, 44-58, 72-88, 99-108, 110-117, 123-129, 131-137, 142-147, 167-178, 181-190, 206-214, 217-223, 271-282, 290-305, 320-327, 329-336, 343-352, 354-364, 396-402, 425-434, 451-456, 471-477, 485-491, 515-541, 544-583, 595-609, 611-626, 644-656, 660-681, 683-691, 695-718	B :4	297-458	54,204
Spy1357	蛋白质	5-43, 92-102, 107-116,	G :27, H :8,	24-135	55,205

化脓链 球菌抗 原蛋白	推定的功能 (通过同源 性)	预测的免疫原性氨基酸**	每次 ORF 和筛选 定克隆的 编号	已鉴定 免疫原 性区的 位置 (氨基酸)	序列识 别号 (DNA , 蛋白 质)
	GRAB(蛋 白质 G- 相 关 α 2M- 结 合 p	120-130, 137-144, 155-163, 169-174, 193-213	K :2, N :4		
SPy1361	推定的 internalin A 前体	4-25, 61-69, 73-85, 88-95, 97-109, 111-130, 135-147, 150-157, 159-179, 182-201, 206-212, 224-248, 253-260, 287-295, 314-331, 338-344, 365-376, 396-405, 413-422, 424-430, 432-449, 478-485, 487-494, 503-517, 522-536, 544-560, 564-578, 585-590, 597-613, 615-623, 629-636, 640-649, 662-671, 713-721	F :21, G :26, H :6, K :4, N :5	176-330	56, 206

[0461]

Spy1371	推定的 NADP- 依赖 型甘油醛 -3- 磷酸酯 脱氢酶	31-37, 41-52, 58-79, 82-105, 133-179, 184-193, 199-205, 209-226, 256-277, 281-295, 297-314, 322-328, 331-337, 359-367, 379-395, 403-409, 417-432, 442-447, 451-460, 466-472	D :14, H :3	46-62 296-341	57, 207
Spy1375	推定的核苷	23-29, 56-63, 67-74, 96-108,	A :2	667-684	58, 208

化脓链 球菌抗 原蛋白	推定的功能 (通过同源 性)	预测的免疫原性氨基酸**	每次 ORF 和筛选选 定克隆的 编号	已鉴定 免疫原 性区的 位置 (氨基 酸)	序列识 别号 (DNA , 蛋白 质)
	酸还原酶 α -c	122-132, 139-146, 152-159, 167-178, 189-196, 214-231, 247-265, 274-293, 301-309, 326-332, 356-363, 378-395, 406-412, 436-442, 445-451, 465-479, 487-501, 528-555, 567-581, 583-599, 610-617, 622-629, 638-662, 681-686, 694-700, 711-716			

[0462]

Spy1389	推定的丙氨 酰 - tRNA 合成酶	20-51, 53-59, 109-115, 140-154, 185-191, 201-209, 212-218, 234-243, 253-263, 277-290, 303-313, 327-337, 342-349, 374-382, 394-410, 436-442, 464-477, 486-499, 521-530, 536-550, 560-566, 569-583, 652-672, 680-686, 698-704, 718-746, 758-770, 774-788, 802-827, 835-842, 861-869	B :2, P :3	258-416	59, 209
Spy1390	推定的蛋白 酶成熟蛋白	7-25, 39-45, 59-70, 92-108, 116-127, 161-168, 202-211,	A :3, B :2, D :3	278-295	60, 210

化脓链 球菌抗 原蛋白	推定的功能 (通过同源 性)	预测的免疫原性氨基酸**	每次 ORF 和筛选选 定克隆的 编号	已鉴定 免疫原 性区的 位置 (氨基 酸)	序列识 别号 (DNA , 蛋白 质)
	质	217-227, 229-239, 254-262, 271-278, 291-300			
Spy1422	推定的重组 蛋白	4-20, 27-33, 45-51, 53-62, 66-74, 81-88, 98-111, 124-130, 136-144, 156-179, 183-191	C:2	183-195	61, 211
Spy1436	推定的脱氧 核糖核酸酶	12-24, 27-33, 43-49, 55-71, 77-85, 122-131, 168-177, 179-203, 209-214, 226-241	K:1	63-238	62, 212
Spy1494	假定的蛋白 质	4-19, 37-50, 120-126, 131-137, 139-162, 177-195, 200-209, 211-218, 233-256, 260-268, 271-283, 288-308	G:3, I:5, K:6, M:5, N:10, O:6, P:4	1-141	63, 213

[0463]

Spy1523	细胞分裂蛋白	11-17, 40-47, 57-63, 96-124, 141-162, 170-207, 223-235, 241-265, 271-277, 281-300, 312-318, 327-333, 373-379	I :2	231-368	64, 214
Spy1536	保守的假定蛋白质	9-33, 41-48, 57-79, 97-103, 113-138, 146-157, 165-186, 195-201, 209-215, 223-229, 237-247, 277-286, 290-297, 328-342	A :19, C :3	247-260	65, 215
Spy1564	保守的假定	7-15, 39-45, 58-64, 79-84,	C :4	64-72	66, 216

化脓链 球菌抗 原蛋白	推定的功能 (通过同源 性)	预测的免疫原性氨基酸**	每次 ORF 和筛选选 定克隆的 编号	已鉴定 免疫原 性区的 位置 (氨基酸)	序列识 别号 (DNA , 蛋白 质)
	蛋白质	97-127, 130-141, 163-176, 195-203, 216-225, 235-247, 254-264, 271-279			
Spy1604	保守的假定 蛋白质	4-12, 26-42, 46-65, 73-80, 82-94, 116-125, 135-146, 167-173, 183-190, 232-271, 274-282, 300-306, 320-343, 351-362, 373-383, 385-391, 402-409, 414-426, 434-455, 460-466, 473-481, 485-503, 519-525, 533-542, 554-565, 599-624, 645-651, 675-693, 717-725, 751-758, 767-785, 792-797, 801-809, 819-825, 831-836, 859-869, 890-897	B :2, K :2	222-362 756-896	67, 217
Spy1607	保守的假定 蛋白质	11-17, 22-28, 52-69, 73-83, 86-97, 123-148, 150-164, 166-177, 179-186, 188-199, 219-225, 229-243, 250-255	D :5	153-170	68, 218
Spy1615	推定的晚感 受态蛋白质	4-61, 71-80, 83-90, 92-128, 133-153, 167-182, 184-192, 198-212	C :4	56-73	69, 219

[0464]

化脓链 球菌抗 原蛋白	推定的功能 (通过同源 性)	预测的免疫原性氨基酸 ^{**}	每次 ORF 和筛选选 定克隆的 编号	已鉴定 免疫原 性区的 位置 (氨基 酸)	序列识 别号 (DNA , 蛋白 质)
Spy1666	保守的假定 蛋白质	4-19, 26-37, 45-52, 58-66, 71-77, 84-92, 94-101, 107-118, 120-133, 156-168, 170-179, 208-216, 228-238, 253-273, 280-296, 303-317, 326-334	D :2	298-312	70, 220
Spy1727	保守的假定 蛋白质	7-13, 27-35, 38-56, 85-108, 113-121, 123-160, 163-169, 172-183, 188-200, 206-211, 219-238, 247-254	B :5	141-157	71, 221

Spy1785	推定的 ATP- 依赖 型 DNA 螺旋酶	23-39, 45-73, 86-103, 107-115, 125-132, 137-146, 148-158, 160-168, 172-179, 185-192, 200-207, 210-224, 233-239, 246-255, 285-334, 338-352, 355-379, 383-389, 408-417, 423-429, 446-456, 460-473, 478-503, 522-540, 553-562, 568-577, 596-602, 620-636, 640-649, 655-663	D :3	433-440 572-593	72,222
Spy1798	假定的蛋白 质	4-42, 46-58, 64-76, 118-124, 130-137, 148-156, 164-169, 175-182, 187-194, 203-218,	A :12, I :12, K :7, N :17, O :13, P :8	17-319 417-563	73,223

[0465]

化脓链 球菌抗 原蛋白	推定的功能 (通过同源 性)	预测的免疫原性氨基酸**	每次 ORF 和筛选 定克隆的 编号	已鉴定 免疫原 性区的 位置 (氨基 酸)	序列识 别号 (DNA ,蛋白 质)
		220-227, 241-246, 254-259, 264-270, 275-289, 296-305, 309-314, 322-334, 342-354, 398-405, 419-426, 432-443, 462-475, 522-530, 552-567, 593-607, 618-634, 636-647, 653-658, 662-670, 681-695, 698-707, 709-720, 732-742, 767-792, 794-822, 828-842, 851-866, 881-890, 895-903, 928-934, 940-963, 978-986, 1003-1025, 1027-1043, 1058-1075, 1080-1087, 1095-1109, 1116-1122, 1133-1138, 1168-1174, 1179-1186, 1207-1214, 1248-1267			

[0466]

Spy1801	免疫原性分泌蛋白前体类似物	6-19, 23-33, 129-138, 140-150, 153-184, 190-198, 206-219, 235-245, 267-275, 284-289, 303-310, 322-328, 354-404, 407-413, 423-446, 453-462,	H :2, I :8, K :6, N :11	46-187	74, 224

化脓链 球菌抗 原蛋白	推定的功能 (通过同源 性)	预测的免疫原性氨基酸 ^{**}	每次 ORF 和筛选选 定克隆的 编号	已鉴定 免疫原 性区的 位置 (氨基 酸)	序列识 别号 (DNA , 蛋白 质)
		467-481, 491-500			
Spy1813	假定的蛋白 质	4-34, 39-57, 78-86, 106-116, 141-151, 156-162, 165-172, 213-237, 252-260, 262-268, 272-279, 296-307, 332-338, 397-403, 406-416, 431-446, 448-453, 464-470, 503-515, 519-525, 534-540, 551-563, 578-593, 646-668, 693-699, 703-719, 738-744, 748-759, 771-777, 807-813, 840-847, 870-876, 897-903, 910-925, 967-976, 979-992	I :16, K :12, N :6	21-244 381-499 818-959	75, 225
Spy1821	推定的翻译 延伸因子 EF-P	19-29, 65-75, 90-109, 111-137, 155-165, 169-175	C :6	118-136	76, 226

Spy1916	推定的磷酸 - β -D- 半乳 糖苷酶	15-20, 30-36, 55-63, 73-79, 90-117, 120-127, 136-149, 166-188, 195-203, 211-223, 242-255, 264-269, 281-287, 325-330, 334-341, 348-366, 395-408, 423-429, 436-444,	C :8	147-155	77, 227

[0467]

化脓链 球菌抗 原蛋白	推定的功能 (通过同源 性)	预测的免疫原性氨基酸**	每次 ORF 和筛选选 定克隆的 编号	已鉴定 免疫原 性区的 位置 (氨基 酸)	序列识 别号 (DNA , 蛋白 质)
		452-465			
Spy1972	支链淀粉酶	11-18, 21-53, 77-83, 91-98, 109-119, 142-163, 173-181, 193-208, 216-227, 238-255, 261-268, 274-286, 290-297, 308-315, 326-332, 352-359, 377-395, 399-406, 418-426, 428-438, 442-448, 458-465, 473-482, 488-499, 514-524, 543-553, 564-600, 623-632, 647-654, 660-669, 672-678, 710-723, 739-749, 787-793, 820-828, 838-860, 889-895, 901-907, 924-939, 956-962, 969-976, 991-999, 1012-1018, 1024-1029, 1035-1072, 1078-1091, 1142-1161	A :6, I :2, K :5, N :9	74-438	78, 228

[0468]

Spy 1979	链激酶 A 前 体	4-31, 41-52, 58-63, 65-73, 83-88, 102-117, 123-130, 150-172, 177-195, 207-217, 222-235, 247-253, 295-305, 315-328, 335-342, 359-365,	I :6, M :3, N :10	156-420	79, 229

化脓链 球菌抗 原蛋白	推定的功能 (通过同源 性)	预测的免疫原性氨基酸 ^{aa}	每次 ORF 和筛选选 定克隆的 编号	已鉴定 免疫原 性区的 位置 (氨基 酸)	序列识 别号 (DNA , 蛋白 质)
		389-394, 404-413			
Spy1983	胶原样表面 蛋白 (Sc1D)	4-42, 56-69, 98-108, 120-125, 210-216, 225-231, 276-285, 304-310, 313-318, 322-343	A :81, B :24, F :19, G :41, I :2, K :2	79-348	80, 230
Spy1991	氨基苯甲酸 酯合酶组分 II	12-21, 24-30, 42-50, 61-67, 69-85, 90-97, 110-143, 155-168	D :2	53-70	81, 231

[0469]

Spy2000	表面脂蛋白	4-26, 41-54, 71-78, 88-96, 116-127, 140-149, 151-158, 161-175, 190-196, 201-208, 220-226, 240-247, 266-281, 298-305, 308-318, 321-329, 344-353, 370-378, 384-405, 418-426, 429-442, 457-463, 494-505, 514-522	B :3, N :2	183-341	82, 232
Spy2006	假定的蛋白 质	4-27, 69-77, 79-101, 117-123, 126-142, 155-161, 171-186, 200-206, 213-231, 233-244, 258-263, 269-275, 315-331, 337-346, 349-372, 376-381,	A :15, B :9, C :5, D :3, F :18, G :25, H :5, M :10, N :5	92-231 618-757	83, 233

化脓链 球菌抗 原蛋白	推定的功能 (通过同源 性)	预测的免疫原性氨基酸**	每次 ORF 和筛选选 定克隆的 编号	已鉴定 免疫原 性区的 位置 (氨基 酸)	序列识 别号 (DNA , 蛋白 质)
		401-410, 424-445, 447-455, 463-470, 478-484, 520-536, 546-555, 558-569, 580-597, 603-618, 628-638, 648-660, 668-683, 717-723, 765-771, 781-788, 792-806, 812-822			
Spy2009	假定的蛋白 质	11-47, 63-75, 108-117, 119-128, 133-143, 171-185, 190-196, 226-232, 257-264, 278-283, 297-309, 332-338, 341-346, 351-358, 362-372	B : 2, I : 7, K : 7, P : 2	41-170	84, 234

Spy2010	C5A 肽酶 前体	6-26, 50-56, 83-89, 108-114, 123-131, 172-181, 194-200, 221-238, 241-259, 263-271, 284-292, 304-319, 321-335, 353-358, 384-391, 408-417, 424-430, 442-448, 459-466, 487-500, 514-528, 541-556, 572-578, 595-601, 605-613, 620-631, 634-648, 660-679, 686-693, 702-708, 716-725, 730-735, 749-755, 770-777,	A :47, B :10, D :3, F :48, G :20, H :4, I :6, K :13, M :5, N :10, P :6	20-487 757-115 3	85,235

化脓链 球菌抗 原蛋白	推定的功能 (通过同源 性)	预测的免疫原性氨基酸 ^{**}	每次 ORF 和筛选选 定克隆的 编号	已鉴定 免疫原 性区的 位置 (氨基 酸)	序列识 别号 (DNA , 蛋白 质)
		805-811, 831-837, 843-851, 854-860, 863-869, 895-901, 904-914, 922-929, 933-938, 947-952, 956-963, 1000-1005, 1008-1014, 1021-1030, 1131-1137, 1154-1164, 1166-1174			
Spy2016	补体抑制物 (Sic)	10-34, 67-78, 131-146, 160-175, 189-194, 201-214, 239-250, 265-271, 296-305	A :11, B :38, C :16, F :56, G :27, H :13, K :5, N :2, O :3, P :14	26-74 91-100 105-303	86, 236

Spy2018	M1- 蛋白质	9-15, 19-32, 109-122, 143-150, 171-180, 186-191, 209-217, 223-229, 260-273, 302-315, 340-346, 353-359, 377-383, 389-406, 420-426, 460-480	A :316, B :26, C :107, D :12, E :49, F :88, G :118, H :6, I :7,	10-223 231-251 264-297 312-336	87, 237

[0470]

化脓链 球菌抗 原蛋白	推定的功能 (通过同源 性)	预测的免疫原性氨基酸**	每次 ORF 和筛选选 定克隆的 编号	已鉴定 免疫原 性区的 位置 (氨基 酸)	序列识 别号 (DNA ,蛋白 质)
			K :2, M :48, N :4		
Spy2025	免疫原性分 泌蛋白前体	5-28, 76-81, 180-195, 203-209, 211-219, 227-234, 242-252, 271-282, 317-325, 350-356, 358-364, 394-400, 405-413, 417-424, 430-436, 443-449, 462-482, 488-498, 503-509, 525-537	F :7, G :16, H :7, K :63, L :2, N :18, O :42	22-344	88, 238
Spy2039	发热产生的 外毒素 B	5-28, 42-54, 77-83, 86-93, 98-104, 120-127, 145-159, 166-176, 181-187, 189-197, 213-218, 230-237, 263-271, 285-291, 299-305, 326-346, 368-375, 390-395	I :15, K :3, N :12	1-151	89, 239

[0471]

Spy2043	有丝分裂因子 MF1 (speF)	6-34, 48-55, 58-64, 84-101, 121-127, 143-149, 153-159, 163-170, 173-181, 216-225, 227-240, 248-254, 275-290, 349-364, 375-410, 412-418, 432-438, 445-451, 465-475, 488-496, 505-515, 558-564,	K : 1	91-263	90,240

化脓链 球菌抗 原蛋白	推定的功能 (通过同源 性)	预测的免疫原性氨基酸**	每次 ORF 和筛选选 定克隆的 编号	已鉴定 免疫原 性区的 位置 (氨基酸)	序列识 别号 (DNA , 蛋白 质)
		571-579, 585-595, 604-613, 626-643, 652-659, 677-686, 688-696, 702-709, 731-747, 777-795, 820-828, 836-842, 845-856, 863-868, 874-882, 900-909, 926-943, 961-976, 980-986, 992-998, 1022-1034, 1044-1074, 1085-1096, 1101-1112, 1117-1123, 1130-1147, 1181-1187, 1204-1211, 1213-1223, 1226-1239, 1242-1249, 1265-1271, 1273-1293, 1300-1308, 1361-1367, 1378-1384, 1395-1406, 1420-1428, 1439-1446, 1454-1460, 1477-1487, 1509-1520, 1526-1536, 1557-1574, 1585-1596, 1605-1617, 1621-1627, 1631-1637, 1648-1654, 1675-1689, 1692-1698,			

[0472]

化脓链 球菌抗 原蛋白	推定的功能 (通过同源 性)	预测的免疫原性氨基酸**	每次 ORF 和筛选选 定克隆的 编号	已鉴定 免疫原 性区的 位置 (氨基酸)	序列识 别号 (DNA , 蛋白 质)

		1700-1706, 1712-1719, 1743-1756			
Spy2059	青霉素结合 蛋白 2a	4-16, 75-90, 101-136, 138-144, 158-164, 171-177, 191-201, 214-222, 231-241, 284-290, 297-305, 311-321, 330-339, 352-369, 378-385, 403-412, 414-422, 428-435, 457-473, 503-521, 546-554, 562-568, 571-582, 589-594, 600-608, 626-635, 652-669, 687-702, 706-712, 718-724, 748-760, 770-775	D :2, E :2	261-272	91, 241
Spy2110	推定的厌氧 核糖核苷 - 三磷酸酯还 原酶	4-19, 30-41, 46-57, 62-68, 75-92, 126-132, 149-156, 158-168, 171-184, 187-194, 210-216, 218-238, 245-253, 306-312, 323-329, 340-351, 365-373, 384-391, 399-405, 422-432, 454-465, 471-481, 502-519, 530-541, 550-562, 566-572, 576-582, 593-599,	E :7	541-551	92, 242

[0473]

化脓链 球菌抗 原蛋白	推定的功能 (通过同源 性)	预测的免疫原性氨基酸**	每次 ORF 和筛选 定克隆的 编号	已鉴定 免疫原 性区的 位置 (氨基 酸)	序列识 别号 (DNA , 蛋白 质)
		620-634, 637-643, 645-651, 657-664, 688-701			
Spy2127	假定的蛋白 质	6-11, 17-25, 53-58, 80-86, 91-99, 101-113, 123-131, 162-169, 181-188, 199-231, 245-252	I : 6, P : 2	84-254	93, 243
Spy2191	假定的蛋白 质	13-30, 71-120, 125-137, 139-145, 184-199	C : 20, E : 3, M : 5	61-78	94, 244

Spy2211	跨膜蛋白	9-30, 38-53, 63-70, 74-97, 103-150, 158-175, 183-217, 225-253, 260-268, 272-286, 290-341, 352-428, 434-450, 453-460, 469-478, 513-525, 527-534, 554-563, 586-600, 602-610, 624-640, 656-684, 707-729, 735-749, 757-763, 766-772, 779-788, 799-805, 807-815, 819-826, 831-855	A :3	568-580	95, 245
ARF045 0	无同源性	11-21, 29-38	A :11	5-17	96, 246
ARF056 9	无同源性	无	A :2	2-9	97, 247

[0474]

化脓链 球菌抗 原蛋白	推定的功能 (通过同源 性)	预测的免疫原性氨基酸 ^{aa}	每次 ORF 和筛选选 定克隆的 编号	已鉴定 免疫原 性区的 位置 (氨基 酸)	序列识 别号 (DNA , 蛋白 质)
ARF069 4	无同源性	4-10, 16-28	B : 7, D : 3, M : 3	7-18 26-34	98, 248
ARF070 0	无同源性	10-16	M : 11	1-15	99, 249
ARF100 7	无同源性	无	B : 2	4-11	100, 250
ARF114 5	无同源性	4-40, 42-51	C : 9	37-53	101, 251
ARF120 8	无同源性	4-21	C : 1	22-29	102, 252
ARF126 2	无同源性	无	D : 2	2-11	103, 253

ARF129 4	与 SA0131 39% 同源性 (67 个氨基 酸蛋白质的 头 28 个氨 基酸)	9-17, 32-44	D :2	1-22	104, 254
ARF131 6	无同源性	19-25, 27-32	E :19	15-34	105, 255
ARF135 2	与 SA1142 38% 同源性 (358 蛋白 质的第 265-295 位 氨基酸)	4-12, 15-22	D :4	11-33	106, 256
ARF148 1	无同源性	10-17, 24-30, 39-46, 51-70	C :2	51-61	107, 257
ARF155 7	无同源性	无	C :2	6-19	108, 258

[0475]

化脓链 球菌抗 原蛋白	推定的功能 (通过同源 性)	预测的免疫原性氨基酸**	每次 ORF 和筛选选 定克隆的 编号	已鉴定 免疫原 性区的 位置 (氨基 酸)	序列识 别号 (DNA ,蛋白 质)
ARF162 9	与 SP0069 具 36% 的 同源性 (211 个氨基酸蛋 白质的第 139-169 位)	6-11, 21-27, 31-54	A :4, B :6	11-29	109, 259
ARF165 4	无同源性	4-10, 13-45	A :2	11-35	110, 260
ARF202 7	无同源性	4-14, 23-32	D :2	11-35	111, 261
ARF209 3	推定的延伸 因子 TS	14-39, 45-51	C :3	15-29	112, 262
ARF220 7	与 SP1006 具 38% 的 同源性 (67 个氨基酸蛋 白质的第 7 -37 位)	4-11, 14-28	A :117	4-17	113, 263
CRF003 8	无同源性	4-16	C :6	2-16	114, 264
CRF012 2	无同源性	4-10, 12-19, 39-50	C :2	6-22	115, 265
CRF040 6	无同源性	none	D :5, E :11	2-13	116, 266
CRF041 6	无同源性	4-11, 22-65	C :42	3-19	117, 267

CRF050 7	无同源性	17-23, 30-35, 39-46, 57-62	B :3, C :4	30-49	118, 268
CRF054	无同源性	4-19	C :6	14-22	119, 269

[0476]

化脓链 球菌抗 原蛋白	推定的功能 (通过同源 性)	预测的免疫原性氨基酸**	每次 ORF 和筛选选 定克隆的 编号	已鉴定 免疫原 性区的 位置 (氨基酸)	序列识 别号 (DNA , 蛋白 质)
9					
CRF056 9	无同源性	无	N :35	2-9	120, 270
CRF062 8	与 P. aeruginosa 的保守假定 蛋白质具 34%的同源 性 (41 个氨 基酸中有 14 个)	7-18, 30-43	A :3	4-12	121, 271
CRF072 7	与肺炎链球 菌的转录调 节子 (70 个 氨基酸, SP0584) 具 40%的同源 性 (40 个氨 基酸中有 16 个同源)	4-30, 39-47	N :6	5-22	122, 272

CRF074 2	与 SA0422 具 33% 的同源性 (42 个氨基酸蛋白质的第 11-37 位, 作为 280 个氨基酸的蛋白质列出)	6-15	D :7, E :12	14-29	123, 273

[0477]

化脓链 球菌抗 原蛋白	推定的功能 (通过同源 性)	预测的免疫原性氨基酸**	每次 ORF 和筛选选 定克隆的 编号	已鉴定 免疫原 性区的 位置 (氨基 酸)	序列识 别号 (DNA , 蛋白 质)
CRF078 4	无同源性	4-34	N :9	23-35	124, 274
CRF085 4	无同源性	4-36, 44-57, 65-72	N :14	14-27	125, 275
CRF087 5	无同源性	4-18	A :4, D :1	11-20	126, 276
CRF090 7	与溶酶体流 通调节子 LYST[人] 同源	无	A :39	5-19	127, 277
CRF097 9	无同源性	18-36	D :21	6-20	128, 278
CRF106 8	无同源性	4-10, 19-34, 41-84, 96-104	C :1, D :3	50-63	129, 279
CRF115 2	无同源性	4-9, 19-27	C :15	8-21	130, 280

CRF120 3	无同源性	4-16, 18-28	N ;3	22-30	131, 281
CRF122 5	无同源性	4-15	C ;8	21-35	132, 282
CRF123 6	无同源性	4-17	N ;3	3-13	133, 283
CRF136 2	无同源性	4-12	C ;6	4-18	134, 284
CRF152 4	无同源性	4-24, 31-36	D ;3	29-45	135, 285
CRF152 5	无同源性	12-22, 34-49	C ;2	21-32	136, 286
CRF152 7	无同源性	4-17	D ;4, E ;1	22-32	137, 287

[0478]

化脓链 球菌抗 原蛋白	推定的功能 (通过同源 性)	预测的免疫原性氨基酸 ^{aa}	每次 ORF 和筛选选 定克隆的 编号	已鉴定 免疫原 性区的 位置 (氨基 酸)	序列识 别号 (DNA , 蛋白 质)
CRF158 8	无同源性	4-16, 25-42	C : 2	7-28	138, 288
CRF164 9	无同源性	4-10	C : 3	7-20	139, 289
CRF174 9	无同源性	4-11, 16-36, 39-54	C : 15	28-44	140, 290
CRF190 3	无同源性	5-20, 29-54	A : 14	14-29	141, 291
CRF196 4	无同源性	24-33	A : 8	10-22	142, 292
CRF205 5	无同源性	10-51, 54-61	B : 1, F : 12, H : 14	43-64	143, 293

CRF209 1	无同源性	7-13	C :2	2-17	144, 294
CRF209 6	无同源性	11-20	C :4	6-20	145, 295
CRF210 4	无同源性	4-30, 34-41	C :2	19-28	146, 296
CRF211 6	无同源性	未检测到		11-21	147, 297
CRF215 3	无同源性	4-16, 21-26	F :2	9-38	148, 298
NRF000 1	在寡 ABC 转运蛋白中的 ARF (未被 TIGR 注释), 与 SA0643 具 33% 的同源性 (469 个氨	4-12, 15-27, 30-42, 66-72	A :7, B :1	10-24	149, 299

[0479]

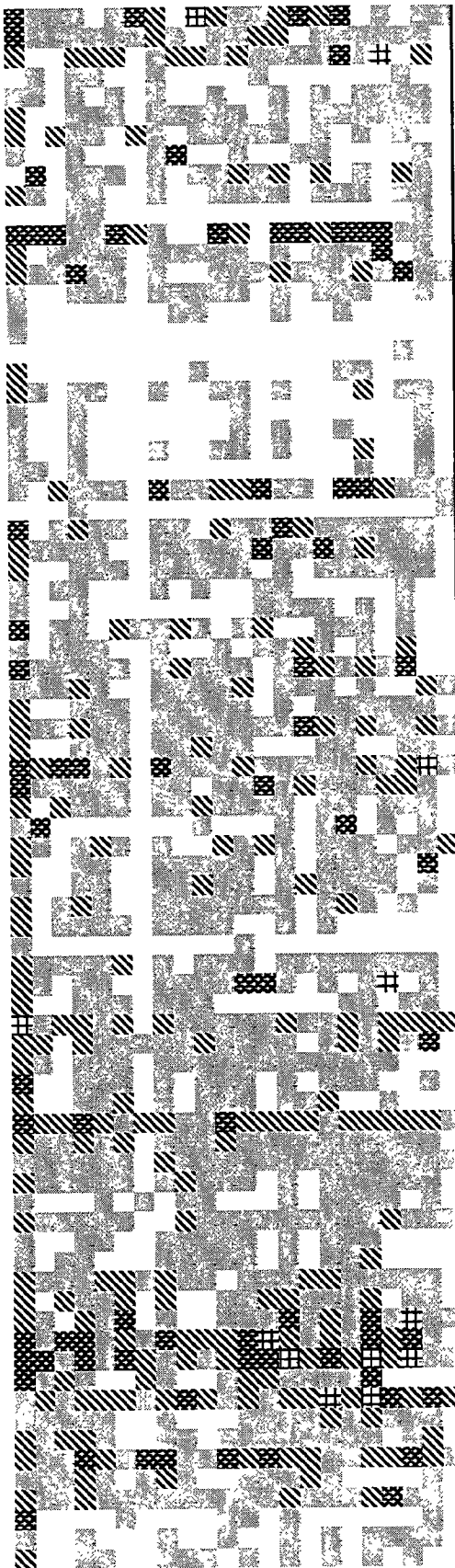
化脓链 球菌抗 原蛋白	推定的功能 (通过同源 性)	预测的免疫原性氨基酸 **	每次 ORF 和筛选选 定克隆的 编号	已鉴定 免疫原 性区的 位置 (氨基 酸)	序列识 别号 (DNA , 蛋白 质)
	基酸蛋白质的第 107- 162 位)				
NRF000 3	无同源性	8-17	A :23	11-20	150, 300

[0480] 表 2-1

[0481]

Peptide	Peptide sequence	Seq. ID	location in protein (aa)	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	N1	N2	N3	N4	N5	N6	N7	N8	N9	N10	N11	N12	N13	N14	N15	N16	N17	N18	N19	N20	N21	N22	N23	N24	N25	N26	N27	N28	N29	N30	N31	N32	N33	N34	N35	N36	N37	N38	N39	N40	N41	N42	N43	N44	N45	N46	N47	N48	N49	N50	N51	N52	N53	N54	N55	N56	N57	N58	N59	N60	N61	N62	N63	N64	N65	N66	N67	N68	N69	N70	N71	N72	N73	N74	N75	N76	N77	N78	N79	N80	N81	N82	N83	N84	N85	N86	N87	N88	N89	N90	N91	N92	N93	N94	N95	N96	N97	N98	N99	N100	N101	N102	N103	N104	N105	N106	N107	N108	N109	N110	N111	N112	N113	N114	N115	N116	N117	N118	N119	N120	N121	N122	N123	N124	N125	N126	N127	N128	N129	N130	N131	N132	N133	N134	N135	N136	N137	N138	N139	N140	N141	N142	N143	N144	N145	N146	N147	N148	N149	N150	N151	N152	N153	N154	N155	N156	N157	N158	N159	N160	N161	N162	N163	N164	N165	N166	N167	N168	N169	N170	N171	N172	N173	N174	N175	N176	N177	N178	N179	N180	N181	N182	N183	N184	N185	N186	N187	N188	N189	N190	N191	N192	N193	N194	N195	N196	N197	N198	N199	N200	N201	N202	N203	N204	N205	N206	N207	N208	N209	N210	N211	N212	N213	N214	N215	N216	N217	N218	N219	N220	N221	N222	N223	N224	N225	N226	N227	N228	N229	N230	N231	N232	N233	N234	N235	N236	N237	N238	N239	N240	N241	N242	N243	N244	N245	N246	N247	N248	N249	N250	N251	N252	N253	N254	N255	N256	N257	N258	N259	N260	N261	N262	N263	N264	N265	N266	N267	N268	N269	N270	N271	N272	N273	N274	N275	N276	N277	N278	N279	N280	N281	N282	N283	N284	N285	N286	N287	N288	N289	N290	N291	N292	N293	N294	N295	N296	N297	N298	N299	N300	N301	N302	N303	N304	N305	N306	N307	N308	N309	N310	N311	N312	N313	N314	N315	N316	N317	N318	N319	N320	N321	N322	N323	N324	N325	N326	N327	N328	N329	N330	N331	N332	N333	N334	N335	N336	N337	N338	N339	N340	N341	N342	N343	N344	N345	N346	N347	N348	N349	N350	N351	N352	N353	N354	N355	N356	N357	N358	N359	N360	N361	N362	N363	N364	N365	N366	N367	N368	N369	N370	N371	N372	N373	N374	N375	N376	N377	N378	N379	N380	N381	N382	N383	N384	N385	N386	N387	N388	N389	N390	N391	N392	N393	N394	N395	N396	N397	N398	N399	N400	N401	N402	N403	N404	N405	N406	N407	N408	N409	N410	N411	N412	N413	N414	N415	N416	N417	N418	N419	N420	N421	N422	N423	N424	N425	N426	N427	N428	N429	N430	N431	N432	N433	N434	N435	N436	N437	N438	N439	N440	N441	N442	N443	N444	N445	N446	N447	N448	N449	N450	N451	N452	N453	N454	N455	N456	N457	N458	N459	N460	N461	N462	N463	N464	N465	N466	N467	N468	N469	N470	N471	N472	N473	N474	N475	N476	N477	N478	N479	N480	N481	N482	N483	N484	N485	N486	N487	N488	N489	N490	N491	N492	N493	N494	N495	N496	N497	N498	N499	N500	N501	N502	N503	N504	N505	N506	N507	N508	N509	N510	N511	N512	N513	N514	N515	N516	N517	N518	N519	N520	N521	N522	N523	N524	N525	N526	N527	N528	N529	N530	N531	N532	N533	N534	N535	N536	N537	N538	N539	N540	N541	N542	N543	N544	N545	N546	N547	N548	N549	N550	N551	N552	N553	N554	N555	N556	N557	N558	N559	N560	N561	N562	N563	N564	N565	N566	N567	N568	N569	N570	N571	N572	N573	N574	N575	N576	N577	N578	N579	N580	N581	N582	N583	N584	N585	N586	N587	N588	N589	N590	N591	N592	N593	N594	N595	N596	N597	N598	N599	N600	N601	N602	N603	N604	N605	N606	N607	N608	N609	N610	N611	N612	N613	N614	N615	N616	N617	N618	N619	N620	N621	N622	N623	N624	N625	N626	N627	N628	N629	N630	N631	N632	N633	N634	N635	N636	N637	N638	N639	N640	N641	N642	N643	N644	N645	N646	N647	N648	N649	N650	N651	N652	N653	N654	N655	N656	N657	N658	N659	N660	N661	N662	N663	N664	N665	N666	N667	N668	N669	N670	N671	N672	N673	N674	N675	N676	N677	N678	N679	N680	N681	N682	N683	N684	N685	N686	N687	N688	N689	N690	N691	N692	N693	N694	N695	N696	N697	N698	N699	N700	N701	N702	N703	N704	N705	N706	N707	N708	N709	N710	N711	N712	N713	N714	N715	N716	N717	N718	N719	N720	N721	N722	N723	N724	N725	N726	N727	N728	N729	N730	N731	N732	N733	N734	N735	N736	N737	N738	N739	N740	N741	N742	N743	N744	N745	N746	N747	N748	N749	N750	N751	N752	N753	N754	N755	N756	N757	N758	N759	N760	N761	N762	N763	N764	N765	N766	N767	N768	N769	N770	N771	N772	N773	N774	N775	N776	N777	N778	N779	N780	N781	N782	N783	N784	N785	N786	N787	N788	N789	N790	N791	N792	N793	N794	N795	N796	N797	N798	N799	N800	N801	N802	N803	N804	N805	N806	N807	N808	N809	N810	N811	N812	N813	N814	N815	N816	N817	N818	N819	N820	N821	N822	N823	N824	N825	N826	N827	N828	N829	N830	N831	N832	N833	N834	N835	N836	N837	N838	N839	N840	N841	N842	N843	N844	N845	N846	N847	N848	N849	N850	N851	N852	N853	N854	N855	N856	N857	N858	N859	N860	N861	N862	N863	N864	N865	N866	N867	N868	N869	N870	N871	N872	N873	N874	N875	N876	N877	N878	N879	N880	N881	N882	N883	N884	N885	N886	N887	N888	N889	N890	N891	N892	N893	N894	N895	N896	N897	N898	N899	N900	N901	N902	N903	N904	N905	N906	N907	N908	N909	N910	N911	N912	N913	N914	N915	N916	N917	N918	N919	N920	N921	N922	N923	N924	N925	N926	N927	N928	N929	N930	N931	N932	N933	N934	N935	N936	N937	N938	N939	N940	N941	N942	N943	N944	N945	N946	N947	N948	N949	N950	N951	N952	N953	N954	N955	N956	N957	N958	N959	N960	N961	N962	N963	N964	N965	N966	N967	N968	N969	N970	N971	N972	N973	N974	N975	N976	N977	N978	N979	N980	N981	N982	N983	N984	N985	N986	N987	N988	N989	N990	N991	N992	N993	N994	N995	N996	N997	N998	N999	N1000	N1001	N1002	N1003	N1004	N1005	N1006	N1007	N1008	N1009	N1010	N1011	N1012	N1013	N1014	N1015	N1016	N1017	N1018	N1019	N1020	N1021	N1022	N1023	N1024	N1025	N1026	N1027	N1028	N1029	N1030	N1031	N1032	N1033	N1034	N1035	N1036	N1037	N1038	N1039	N1040	N1041	N1042	N1043	N1044	N1045	N1046	N1047	N1048	N1049	N1050	N1051	N1052	N1053	N1054	N1055	N1056	N1057	N1058	N1059	N1060	N1061	N1062	N1063	N1064	N1065	N1066	N1067	N1068	N1069	N1070	N1071	N1072	N1073	N1074	N1075	N1076	N1077	N1078	N1079	N1080	N1081	N1082	N1083	N1084	N1085	N1086	N1087	N1088	N1089	N1090	N1091	N1092	N1093	N1094	N1095	N1096	N1097	N1098	N1099	N1100	N1101	N1102	N1103	N1104	N1105	N1106	N1107	N1108	N1109	N1110	N1111	N1112	N1113	N1114	N1115	N1116	N1117	N1118	N1119	N1120	N1121	N1122	N1123	N1124	N1125	N1126	N1127	N1128	N1129	N1130	N1131	N1132	N1133	N1134	N1135	N1136	N1137	N1138	N1139	N1140	N1141	N1142	N1143	N1144	N1145	N1146	N1147	N1148	N1149	N1150	N1151	N1152	N1153	N1154	N1155	N1156	N1157	N1158	N1159	N1160	N1161	N1162	N1163	N1164	N1165	N1166	N1167	N1168	N1169	N1170	N1171	N1172	N1173	N1174	N1175	N1176	N1177	N1178	N1179	N1180	N1181	N1182	N1183	N1184	N1185	N1186	N1187	N1188	N1189	N1190	N1191	N1192	N1193	N1194	N1195	N1196	N1197	N1198	N1199	N1200	N1201	N1202	N1203	N1204	N1205	N1206	N1207	N1208	N1209	N1210	N1211	N1212	N1213	N1214	N1215	N1216	N1217	N1218	N1219	N1220	N1221	N1222	N1223	N1224	N1225	N1226	N1227	N1228	N1229	N1230	N1231	N1232	N1233	N1234	N1235	N1236	N1237	N1238	N1239	N1240	N1241	N1242	N1243	N1244	N1245	N1246	N1247	N1248	N1249	N1250	N1251	N1252	N1253	N1254	N1255	N1256	N1257	N1258	N1259	N1260	N1261	N1262	N1263	N1264	N1265	N1266	N1267	N1268	N1269	N1270	N1271	N1272	N1273	N1274	N1275	N1276	N1277	N1278	N1279	N1280	N1281	N1282	N1283	N1284	N1285	N1286	N1287	N1288	N1289	N1290	N1291	N1292	N1293	N1294	N1295	N1296	N1297	N1298	N1299	N1300	N1301	N1302	N1303	N1304	N1305	N1306	N1307	N1308	N1309	N1310	N1311	N1312	N1313	N1314	N1315	N1316	N1317	N1318	N1319	N1320	N1321	N1322	N1323	N1324	N1325	N1326	N1327	N1328	N1329	N1330	N1331	N1332	N1333	N1334	N1335	N1336	N1337	N1338	N1339	N1340	N1341	N1342	N1343	N1344	N1345	N1346	N1347	N1348	N1349	N1350	N1351	N1352	N1353	N1354	N1355	N1356	N1357	N1358	N1359	N1360	N1361	N1362	N1363	N1364	N1365	N1366	N1367	N1368	N1369	N1370	N1371	N1372	N1373	N1374	N1375	N1376	N1377	N1378	N1379	N1380	N1381	N1382	N1383	N1384	N1385	N1386	N1387	N1388	N1389	N1390	N1391	N1392	N1393	N1394	N1395	N1396	N1397	N1398	N1399	N1400	N1401	N1402	N1403	N1404	N1405	N1406	N1407	N1408	N1409	N1410	N1411	N1412	N1413	N1414	N1415	N1416	N1417	N1418	N1419	N1420	N1421	N1422	N1423	N1424	N1425	N1426	N1427	N1428	N1429	N1430	N1431	N1432	N1433	N1434	N1435	N1436	N1437	N1438	N1439	N1440	N1441	N1442	N1443	N1444	N1445	N1446	N1447	N1448	N1449	N1450	N1451	N1452	N1453	N1454	N1455	N1456	N1457	N1458	N1459	N1460	N1461	N1462	N1463	N1464	N1465	N1466	N1467	N1468	N1469	N1470	N1471	N1472	N1473	N14

SPO0572.1	QAGMAGSPANALL	192	90-105
SPO0572.2	KVFNMKFEYGLNHFEBGLDEF	192	112-136
SPO066.1	FEQVWNLALHLKPFEESE	193	290-209
SPO0972.1	KSGRSKSKTALNFIV	194	33-50
SPO0881.1	DELQELANKETNEL	195	76-90
SPO1008.1	MLKYSVKRNLVFEEDW	196	70-88
SPO1032.1	TRUKLWQGNFVFNWNLKTYHD	197	418-442
SPO1032.2	ASKGANTISALA	197	574-585
SPO1054.1	EKGQGTSGQGERGTG	198	87-104
SPO1054.2	VGPAGKQNGKGLFKGKGTGN	198	124-148
SPO1054.3	KLGKSGNGKOG	198	141-152
SPO1053.1	IKHAPMT	199	241-248
SPO1052.1	QSLKAGNLSIAAA	200	183-198
SPO1206.1	GLVANGKSTENSTVI	201	40-57
SPO1228.1	GSFGAKGCTARAAQ	202	202-217
SPO1245.1	DEFGSSNLGKTVNVOGSSGTLGSG	203	50-74
SPO1245.2	GIGLSVQSGAVQIGNSVFAEKD	203	69-93
SPO1245.3	FAEEKGIDASKVDHQAVALAV	203	88-112
SPO1245.4	VAGLAVNENKVKSLSSQD	203	107-127
SPO1357.1	KHQIENKATTAIEAASS	205	74-92
SPO1361.1	PGLHFTSDGFQNGQVGVTKDSI	206	207-232
SPO1361.2	VTKISLWHDGHUHLSEADLAQGG	206	227-252
SPO1361.3	DIRGG@HVAODDPAKAKKPAET	206	247-272
SPO1371.1	ASAKLSKWRALS	207	47-60
SPO1371.2	ADQIAETK	207	297-305
SPO1371.3	SVGNEDADITPLITSAADVEGL	207	312-337
SPO1375.1	LYNKTKSKTIRLSIL	208	667-384
SPO1380.1	IKANWIKKAFANLLA	210	278-295
SPO1422.1	DIEYADEVILLRAIDENREL	211	179-198
SPO1484.1	NHLSKGTGKIDSTIVKIDAPKD	213	27-51
SPO1484.2	DWAKGELMLWASSQITDARD	213	46-70
SPO1484.3	DDADRYEFLDSDFTSQNETORM	213	65-89
SPO1484.4	NEITHFENLIDRLNLELLEEE	213	84-108
SPO1484.5	DEEDTEPEQWIMPSLELFLATAVEIR	213	112-141
SPO1536.1	ITIKELRGRITA	215	248-260
SPO1564.1	IFVQNERIEDFSLEKRF	216	59-78
SPO1607.1	ACKVSKLEKCYQKLP	218	154-170
SPO1615.1	CRDCLWENKGVNHR	219	57-73
SPO1665.1	EDMKKELVSRKPLPS	220	297-314
SPO1727.1	KRSLPEFSEVATIVH	221	142-157
SPO1765.1	ELEAGKPIIDRWKHEQLS	222	428-447
SPO1765.2	VJANRKTDSKRMITMET	222	573-593
SPO1768.1	EKVLKLGOLDGGLSKTEL	223	523-544
SPO1768.2	RLJWELGPIASQAGKYYKHIEK	223	46-70
SPO1768.3	YKHIEKLVNDVYKSLGERTFD	223	65-89
SPO1768.4	GERTFDNGQKINLILDSGKE	223	84-108
SPO1768.6	TFITKIKVITLSAQKLGITTHQDSUKGV	223	122-151
SPO1801.1	KPTDQSPSPKVDIAPASS	224	123-142
SPO1813.1	DIKIRYSSPWPELOILG	225	903-921
SPO1821.1	EVIGVITVITVELIVAEI	226	119-136
SPO1916.1	AECFKKESEVNIWITPNEI	227	142-161
SPO1972.1	ANEIGLILKSKTGDAKW	228	258-277
SPO1972.2	GORDVQPKYLEKELDINHQTVEKIDP	228	272-300
SPO1972.3	VKIDENWNPYIQLVSLGEGQTP	228	295-322
SPO1979.1	DEPDRSHLKFTIKVDVNIILSKSROLTAS	229	311-343
SPO1979.2	YEINRKTIGKTNVILVSEKYYVLK	229	278-304
SPO1983.1	ERGAGKIGRGCGIQGKA	230	131-150
SPO1983.2	FGEKENGIRGCTGQGVGQGE	230	195-218
SPO1991.1	WPEANQKRLIGDFYOT	231	53-70
SPO2000.1	NPIGSGPWAKEKAGEQATVNP	232	184-208
SPO2000.3	WVLLENPLAIESGDMITVATP	232	222-246
SPO2000.4	MTVATPELADKVKGTLLIDFEND	232	241-265
SPO2000.5	DIPSNDARGLSLFYKGVITDSFD	232	260-284
SPO2000.6	ITISPDGIFVGNVTSDEAIRKALT	232	279-303
SPO2000.8	NGYKGPVSIIDKTFWNRKTAIKD	232	317-341
SPO2005.1	VIAKETFWKGSFSLRA	233	678-696
SPO2010.1	ENKATUTPATSQITRLNDPSQVATL	235	88-114
SPO2010.2	PTASGTLKLRSSGLDIA	235	464-481
SPO2016.1	KRIDMGFGVATDPYITPPY	236	153-172
SPO2016.2	DPYITPPYGAIGTGIERD	236	137-155, 166-184
SPO2016.3	POWNGFGLSPERS	236	215-228
SPO2018.1	TEWANGDNREVI	237	37-51
SPO2018.2	DIAANNPAIQNIRLRYENKOLKA	237	53-75
SPO2018.3	EEKQISLPSRGLRDLIA	237	232-251
SPO2018.4	EKQISLPSRGLRDLIAS	237	318-336
SPO2025.1	AFDINENAKK	238	305-315
SPO2025.3	QNKDGRPTSPDQKQDITPKYPEKS	238	131-156
SPO2029.1	QNGWGVEDASQYTGIT	241	258-275
SPO2127.1	ARKVADMDITKYSQGYDQMLKAVDEE	243	107-137



33	9	4	4	1
24	15	3	1	
30	7	8	1	1
6	6			
12	12			
17	15	1		
17	11	3		
10	7		1	
20	9	4	1	
16	13	1		
3	3			
41	5	3	10	
19	14	1	1	
26	14	3	2	
14	14			
5	5			
5	3	1		
13	9	2		
4	4			
6	6			
9	7	1		
5	5			
31	11	4	4	
2	2			
25	13	3	2	
25	15	2	2	
17	15	1		
8	8			
7	7			
19	10	3	1	
14	10	2		
27	12	3	3	
19	13	3		
17	15	1		
28	15	5	1	
15	11	2		
37	11	5	4	1
26	14	3	2	
20	14	3		
12	6	2		
21	11	5		
17	12	1	1	
17	13	2		
18	12	3		
12	10	1		
2	2			
20	16	2		
23	11	1	2	1
12	10	1		
34	10	10		1
27	12	6	1	
18	16	1		
15	12	1		
15	9	3		
42	5	14	3	
24	16	4		
22	20	1		
20	16	2		
11	11			
23	17	3		
12	12			
13	11	1		
26	14	6		
23	9	7		
30	11	3	3	1
45	6	7	7	1
51	6	6	7	3
22	8	4	2	
46	5	12	3	2
18	14	2		
21	13	4		
42	8	8	6	
17	11	3		
27	16	4	1	
8	5	1		
10	10			
12	10	1		

[0483] 表 2-2

[0484] 表 3 : 化脓链球菌菌株中的基因分布

[0485]

ORF	常用名	基因分布 (50 的出 现率)	氨基酸替 代 (在菌株 M89 中)	同源性 (SP/EC)	序列 ID (DNA, 蛋白质)
Spy0012	假定的蛋白质	50	3/302	SP0010-40% / 无	1, 151
Spy0019	推定的分泌蛋白 质 (细胞分裂和抗 生素耐受)	50	0/300	SP2216-44-49% / 无	2, 152
Spy0025	推定的磷酸核糖 甲酰甘氨酸合酶 II	38	0/303	SP0045-85% / 24%	3, 153
Spy0031	推定的胆碱结合 蛋白	50	0/297	SP2201-42% (cbpD) / 无	4, 154
Spy0103	推定的免疫蛋白 质	50	0/81	SP2051-41% / 无	5, 155
Spy0112	推定的吡咯啉羧 酸盐还原酶	50	3/235	SP0933-32% / 34%	6, 156
Spy0115	推定的谷氨酰-氨 基肽酶	50	6/306	SP1865-76% / 30%	7, 157
Spy0166	假定的蛋白质	50	未检测到	无 / 无	8, 158
Spy0167	链球菌溶血素 O	50	7/300	SP1923-40% (肺炎 球菌溶血素) / 无	9, 159
Spy0168	假定的蛋白质	8	19/126	无 / 无 e	10, 160
Spy0171	假定的蛋白质	18	8/95	无 / 无	11, 161
Spy0183	推定的甘氨酸甜 菜碱 / 脯氨酸 ABC 转运蛋白	50	0/297	SP0151-39% / 48%	12, 162
Spy0230	推定的 ABC 转运 蛋白 (ATP- 结合蛋 白)	50	1/299	SP2073-64% / 32%	13, 163
Spy0269	推定的表面排阻 蛋白	50	1/303	无 / 无	14, 164
Spy0287	保守的假定蛋白 质	50	1/307	SP0868-71% / 19%	15, 165

Spy0292	青霉素结合蛋白质 (D- 丙氨酰 -D-	50	1/359	SP0872-47% /27%	16,166

[0486]

	丙氨酸 car				
Spy0295	寡肽透性酶	50	2/269	SP1889-69% /24%	17,167
Spy0348	推定的氨基脱氧 chorismate 裂合酶	50	1/307	SP1518-47% /25%	18,168
Spy0416	推定的细胞外壳丝氨酸蛋白酶	50	4/314	SP0641-22% / 无	19,169
Spy0430	假定的蛋白质	13	0/165#	无 / 无	20,170
Spy0433	假定的蛋白质	21 (27/49) 1	2/174#	无 / 无	21,171
Spy0437	假定的蛋白质	19 (34/49) ¹	0/106#	无 / 无	22,172
Spy0469	推定的 42kDa 蛋白质	50	6/313	SP2063-44% (LysM 蛋白) / 无	23,173
Spy0488	假定的蛋白质	50	9/178	无 / 无	24,174
Spy0515	推定的糖转移酶	50	未检测到	SP1075-26% /None	25,175
Spy0580	保守的假定蛋白质	50	0/297	SP0908-72% /43%	26,176
Spy0621	保守的假定蛋白质	50	未检测到	SP1290-72% / 无	27,177
Spy0630	推定的 PTS 依赖型 N- 乙酰 - 半乳糖胺 -IIC	50	未检测到	SP0324-79% /30%	28,178
Spy0681	假定的蛋白质, 噬菌体相关	27	2/303#	无 / 无	29,179
Spy0683	推定的小衣壳蛋白质, 噬菌体相关	25	1/233	无 / 无	30,180
Spy0702	假定的蛋白质	22	未检测到	无 / 无	31,181
Spy0710	保守的假定蛋白质, 噬菌体相关	32	51/286#	无 /36% 在 313 个氨基酸中的 122 个中	32,182

Spy0711	发热产生的外毒素 C 前体, 噬菌体相关 (speC)	17	1/225	无 / 无	33,183
Spy0720	保守的假定蛋白质	50	2/270	SP1298-60% (DHH 1 蛋白) / 无	34,184
Spy0727	推定的 DNA 旋转酶, 亚基 B	未检测到	未检测到	SP0806-80% / 46%	35,185

[0487]

Spy0737	推定的胞外基质结合蛋白	29 (48/49) ¹	0/466#	无 / 27% 在 421 个氨基酸中的 340 个中	36,186
Spy0747	细胞外核酸酶	50	0/179	无 / 无	37,187
Spy0777	推定的 ATP- 依赖型核酸外切酶, 亚基 A	50	2/306	SP1152-48% / 22%	38,188
Spy0789	推定的 ABC- 转运蛋白 (透性酶蛋白)	50	1/231	无 / 无	39,189
Spy0839	推定的甘油磷酸二酯	50	1/301	SP0994-24% / 31% 在 358 个氨基酸的 121 个中	40,190
Spy0843	细胞表面蛋白质	50	3/312	无 / 无	41,191
Spy0872	推定的被分泌 5' - 核苷酸酶	50	2/309	无 / 27% 在 647 个氨基酸的 274 个中	42,192
Spy0895	组氨酸蛋白质激酶	50	0/244	无 / 无	43,193
Spy0972	推定的末端酶, 大亚基 - 噬菌体	28	1/314#	无 / 无	44,194
Spy0981	假定的蛋白质 - 噬菌体相关	23	未检测到	无 / 无	45,195
Spy1008	链球菌外毒素 H 前体 (speH)	15 (14/49) ¹	1/223#	无 / 无	46,196

Spy1032	细胞外透明质酸盐裂合酶	50(175 的 175, Hynes 2000)	3/311	SP0314-51% / 无	47,197
Spy1054	推定的胶原样蛋白 (Sc1C)	26, (45/49) ¹ (50 的 50, 但重复片段的数目变化 ; Lukomski, 2001)	未检测到	无 / 无	48,198
Spy1063	推定的外周铁结合蛋白	49/50 (49/49) ¹	2/292#	SP0243-52%, 铁 ABC 转运蛋白 /26% 在 348 个氨基酸的 161 个中	49,199

[0488]

Spy1162	推定的核糖核酸酶 HII	50	3/240	SP1156-67% /46%	50,200
Spy1206	推定的 ABC 转运蛋白	50	1/302	SP0770-81% /30%	51,201
Spy1228	推定的脂蛋白	49	未检测到	SP0845-57% /None	52,202
Spy1245	推定的 ABC 转运蛋白	50	未检测到	SP1400-64% /None	53,203
Spy1315	假定的蛋白质	50	4/305	SP1241-64% /32%	54,204
Spy1357	蛋白质 GRAB(蛋白质 G- 相关 α 2M- 结合蛋白)	49 ;12 个菌株中的 11 个 (Rasmussen, 1999)	9/226 ;28 个氨基酸的插入	无 / 无	55,205
Spy1361	推定的 internalin A 前体	50	7/295	SP1004-26% 在 1039 个的 283 个中 / 无	56,206

Spy1371	推定的 NADP- 依 赖型甘油醛-3- 磷 酸盐脱氢酶	50	2/308	SP1119-71% /34%	57,207
Spy1375	推定的核糖核苷 酸还原酶 α -c	50	4/304	SP1179-85% /49%	58,208
Spy1389	推定的丙氨酰 -tRNA 合成酶	50	0/309	SP1383-74% /40%	59,209
Spy1390	推定的蛋白酶成 熟蛋白	50	0/232	SP0981-42% / 无	60,210
Spy1422	推定的重组蛋白	未检测到	未检测到	SP1672-88% /64%	61,211
Spy1436	推定的脱氧核糖 核酸酶	25	0/243#	SP1964-29%在 274 个氨基酸的 181 个中 / 无	62,212
Spy1494	假定的蛋白质	50	13/282	无 / 无	63,213
Spy1523	细胞分裂蛋白	49	2/329	SP0690-27% / 无	64,214
Spy1536	保守的假定蛋白 质	50	9/280	SP1967-57% / 无	65,215
Spy1564	保守的假定蛋白 质	39	未检测到	无 / 无	66,216
Spy1604	保守的假定蛋白 质	50	1/233	SP2143-47% /28%	67,217
Spy1607	保守的假定蛋白	50	0/241	SP1902-55% / 无	68,218

[0489]

	质				
Spy1615	推定的晚感受蛋 白	50	2/204	SP2207-41% / 无	69,219
Spy1666	保守的假定蛋白 质	50	2/305	SP0334(γ 11C)- 78% /40%	70,220
Spy1727	保守的假定蛋白 质	50	0/237	SP0549-53% / 无	71,221

Spy1785	推定的 ATP- 依赖型 DNA 螺旋酶	50	1/306	SP1697-71% /37%	72,222
Spy1798	假定的蛋白质	50	2/128	无 / 无	73,223
Spy1801	免疫原性分泌蛋白前体类似物	50	6/313 ;6 个氨基酸的 插入 aa	SP2216-33%在 392 个氨基酸的 119 个中 / 无	74,224
Spy1813	假定的蛋白质	46	47/433 ;9 个氨基酸的插入,1 个氨基酸的删除	无 / 无	75,225
Spy1821	推定的翻译延伸因子 EF-P	未检测到	未检测到	SP0435-94% /45%	76,226
Spy1916	推定的磷酸 -β-D- 半乳糖苷酶	未检测到	未检测到	SP1184-91% /83%	77,227
Spy1972	支链淀粉酶	50	1/233	SP0268-53%, SP1118-29% /25% 在 657 个氨基酸的 352 个中	78,228
Spy1979	链激酶 A 前体	50	309# 的 20.1% 同一性	无 / 无	79,229
Spy1983	胶原样表面蛋白 (Sc1D)	50, (50of 50, 但大小 变异符合 Lukowski, 2000	未检测到	无 / 无	80,230
Spy1991	氨基苯甲酸盐和酶组分 II	50	1/170	SP1816-58% /47%	81,231
Spy2000	表面脂蛋白	50	0/307	None/27%在 524 个 氨基酸的 389 个中	82,232

[0490]

Spy2006	假定的蛋白质	50	0/234	SP1003-36%, SP1174-37%, SP1004-33%, SP1175-48% / 无	83,233

Spy2009	假定的蛋白质	39(38/49) 1	58/344 ; insertion of 36, deletion of 4aa	无 / 无	84,234
Spy2010	C5A 肽酶前体	未检测到	未检测到	SP0641-23%在 2140 个氨基酸的 783 个中 / 无	85,235
Spy2016	补体抑制物 (Sic)	47 ;主要在 M1 菌株 中 (Reid 2001)	11/269#	无 / 无	86,236
Spy2018	M1- 蛋白	未检测到	未检测到	无 / 无	87,237
Spy2025	免疫原性分泌蛋 白前体	50	3/296	SP2216-31%在 392 个氨基酸的 138 个中 / 无	88,238
Spy2039	发热产生的外毒 素 B	未检测到	未检测到	无 / 无	89,239
Spy2043	有丝分裂因子 MF1 (speF)	50	0/247	无 / 无	90,240
Spy2059	青霉素结合蛋白 2a	50	0/293	SP2010-55% (pbp2A)/30%在 844 个氨基酸的 539 个中	91,241
Spy2110	推定的厌氧核糖 核苷 - 三磷酸盐还 原酶	50	0/311	SP0202-80% (nrdD)/50%	92,242
Spy2127	假定的蛋白质	1	未检测到	无 / 无	93,243
Spy2191	假定的蛋白质	50	1/175	无 / 无	94,244
Spy2211	跨膜蛋白质	50	2/281	SP2231-43% / 无	95,245
ARF0450	假定的蛋白质	50	5/191	无 / 无	96,246
ARF0569	假定的蛋白质	未检测到	未检测到	无 / 无	97,247
ARF0694	假定的蛋白质	23	1/122#	无 / 无	98,248
ARF0700	假定的蛋白质	未检测到	未检测到	无 / 无 e	99,249
ARF1007	假定的蛋白质	未检测到	未检测到	无 / 无	100,250

[0491]

ARF1145	假定的蛋白质	未检测到	未检测到	无 / 无	101,251
ARF1208	假定的蛋白质	未检测到	未检测到	无 / 无	102,252
ARF1262	假定的蛋白质	未检测到	未检测到	无 / 无	103,253
ARF1294	假定的蛋白质	50	1/186	与 SA0131 具 39% 的同源性 (67 个氨基酸蛋白质的头 28 个氨基酸)	104,254
ARF1316	假定的蛋白质	未检测到	未检测到	无 / 无	105,255
ARF1352	假定的蛋白质	未检测到	未检测到	与 SA1142 具 38% 的同源性 (358 个氨基酸蛋白质的第 265-295 位氨基酸)	106,256
ARF1481	假定的蛋白质	未检测到	未检测到	无 / 无	107,257
ARF1557	假定的蛋白质	未检测到	未检测到	无 / 无	108,258
ARF1629	假定的蛋白质	未检测到	未检测到	与 SP0069 具 36% 的同源性 (211 个氨基酸蛋白质的第 139-169 位氨基酸)	109,259
ARF1654	假定的蛋白质	未检测到	未检测到	无 / 无	110,260
ARF2027	假定的蛋白质	未检测到	未检测到	无 / 无	111,261
ARF2093	假定的蛋白质	未检测到	未检测到	无 / 无	112,262
ARF2207	假定的蛋白质	50	未检测到	与 SP1006 具 38% 的同源性 (67 个氨基酸蛋白质的第 7-37 位氨基酸)	113,263
CRF0038	假定的蛋白质	未检测到	未检测到	无 / 无	114,264
CRF0122	假定的蛋白质	未检测到	未检测到	无 / 无	115,265
CRF0406	假定的蛋白质	未检测到	未检测到	无 / 无	116,266
CRF0416	假定的蛋白质	未检测到	未检测到	无 / 无	117,267

CRF0507	假定的蛋白质	未检测到	未检测到	无 / 无	118,268
CRF0549	假定的蛋白质	未检测到	未检测到	无 / 无	119,269
CRF0569	假定的蛋白质	未检测到	未检测到	无 / 无	120,270
CRF0628	假定的蛋白质	未检测到	未检测到	无 / 无	121,271
CRF0727	假定的蛋白质	未检测到	未检测到	与 SP0584 具 40% 的同源性 (70 个氨基酸蛋白质的第 21-60 位氨基酸)	122,272
CRF0742	假定的蛋白质	未检测到	未检测到	与 SA0422 具 33% 的	123,273

[0492]

				同源性 (42 个氨基酸蛋白质的第 11-37 位氨基酸, 作为 280 个氨基酸的蛋白质列出)	
CRF0784	假定的蛋白质	未检测到	未检测到	无 / 无	124,274
CRF0854	假定的蛋白质	未检测到	未检测到	无 / 无	125,275
CRF0875	假定的蛋白质	未检测到	未检测到	无 / 无	126,276
CRF0907	假定的蛋白质	未检测到	未检测到	与溶酶体流通调节子 LYST[人] 同源	127,277
CRF0979	假定的蛋白质	未检测到	未检测到	无 / 无	128,278
CRF1068	假定的蛋白质	50	0/148	无 / 无	129,279
CRF1152	假定的蛋白质	未检测到	未检测到	无 / 无	130,280
CRF1203	假定的蛋白质	未检测到	未检测到	无 / 无	131,281
CRF1225	假定的蛋白质	未检测到	未检测到	无 / 无	132,282
CRF1236	假定的蛋白质	未检测到	未检测到	无 / 无	133,283
CRF1362	假定的蛋白质	未检测到	未检测到	无 / 无	134,284
CRF1524	假定的蛋白质	未检测到	未检测到	无 / 无	135,285

CRF1525	假定的蛋白质	未检测到	未检测到	无 / 无	136,286
CRF1527	假定的蛋白质	未检测到	未检测到	无 / 无	137,287
CRF1588	假定的蛋白质	未检测到	未检测到	无 / 无	138,288
CRF1649	假定的蛋白质	未检测到	未检测到	无 / 无	139,289
CRF1749	假定的蛋白质	未检测到	未检测到	无 / 无	140,290
CRF1903	假定的蛋白质	50	0/140	无 / 无	141,291
CRF1964	假定的蛋白质	未检测到	未检测到	无 / 无	142,292
CRF2055	假定的蛋白质	未检测到	未检测到	无 / 无	143,293
CRF2091	假定的蛋白质	未检测到	未检测到	无 / 无	144,294
CRF2096	假定的蛋白质	未检测到	未检测到	无 / 无	145,295
CRF2104	假定的蛋白质	未检测到	未检测到	无 / 无	146,296
CRF2116	假定的蛋白质	未检测到	未检测到	无 / 无	147,297
CRF2153	假定的蛋白质	未检测到	未检测到	无 / 无	148,298
NRF0001	假定的蛋白质	50	0/130	在寡 ABC 转运蛋白中的 ARF (未用 TIGR 注释), 与 SA0643 具 33% 的同源性 (469 个氨基酸蛋白质的第 107-162	149,299

[0493]

				位氨基酸)	
NRF0003	假定的蛋白质	未检测到	未检测到	无 / 无	150,300

[0494] 表 4 :用于 NMRI 小鼠中免疫接种实验的重组蛋白质。

[0495]

ORF	长度 (氨基酸)	氨基酸 ^A		溶解性	保护作用 ^B	被克隆片段的总长 (Kbp)
		从	至			
Spy0031	374	39	374	不可溶	20 % (10 %, 40 %)	1.008
Spy0103	108	2	108		50% (10%, 80%)	0.321
Spy 0269	873	36	873	可溶	40% (40%, 70%) ^C	2.511
<i>Spy 0292</i>	<i>410</i>	22	410	不可溶	70% (10%, 80%)	1.164
<i>Spy0416A</i>	<i>1647</i>	33	867	可溶	50 % (10 %, 40 %)	2.502
Spy0416 B	1647	736	1617	增溶	0 % (0%, 40 %)	2.646
<i>Spy0720</i>	<i>313</i>	2	313	不可溶	60% (10%, 80%)	0.939
<i>Spy0872</i>	<i>670</i>	27	640	增溶	60% (10%, 80%)	1.839
Spy1245	288	49	288	可溶	20 % (10 %, 40 %)	0.717
Spy1357	217	33	186	可溶	40 % (30%, 90 %)	0.459
Spy1361	792	22	792	可溶	60 % (30%, 90 %)	2.31
<i>Spy1390</i>	<i>351</i>	21	351		60% (10%, 80%)	0.99
Spy1536	345	31	345		20 % (0%, 40 %)	0.942
Spy1607	258	2	258		40 % (10 %, 40 %)	0.771
Spy1666	337	22	337	可溶	50 % (30%, 90 %)	0.945

[0496]

Spy1972	1165	45	500		40 % (30%, 90 %)	1.365
Spy2000	542	24	542	可溶	20 % (30%, 90 %)	1.554
Spy2025	541	27	541		40 % (40%, 70%)	1.542
Spy2191	204	36	204		50% (10%, 80%)	0.504

[0497] 表 5 :化脓链球菌菌株的抗原可变性。

抗原名称	序列号	抗原中的残基 ^A	残基数目	氨基酸变化 ^B
Spy0031	154	G	126	D
		A	192	S
		V	233	I
		D	328	N
		I	338	T
Spy0103	155	无		
Spy0269	164	H	97	N
		A	150	V
		A	168	V
		H	482	R
		N	485	K
		Q	577	E
		A	610	V
		L	636	M
		E	640	K
		P	752	S
		I	764	V
		D	765	E
		K	873	R
Spy0292	166	A	214	D
		Y	309	S
		T	317	N
		V	318	C

[0498]

[0499]

Spy0416	169	K	319	Q
		V	1	M
		F	25	M
		L	26	M
		V	27	M
		S	38	T
		M	40	T
		A	49	T
		S	68	P
		L	76	P
		S	85	P
		D	87	G
		S	104	P
		S	110	P
		D	151 ^C	A, S, T, G
		S	164	P
		E	215	G
		H	279 ^C	A, S, T, G
		T	395	I
		D	452	N
		N	478	K
		G	484	D
		A	547	V
		S	617 ^C	A, S, T, G
		D	723	A
		H	749	R
		R	770	K
		P	787	S
		D	804	A
		T	874	M
		N	913	S
		H	991	Y
		N	1080	S
		V	1238	A
		D	1313	G
		V	1349	M
		A	1393	V
		N	1479	K
		I	1487	M
		D	1516	G
		N	1555	D
		T	1560	A

[0500]

		S	1599	F
		S	1605	T
		T	1617	A
Spy0720	184	A	61	T
		I	63	M
		K	99	Q
		K	109	Q
		N	295	S
Spy0872	192	K	178	N
		P	181	S
		V	253	I
		A	393	V
		T	600	I
		V	605	I
Spy1063	199	N	168	S
		A	169	S
		D	170	E
		A	173	E
		M	175	V
		V	180	L
		N	181	S
		E	192	D
		Q	195	E
		K	228	D
		H	243	K
		P	245	K
		N	246	A
		T	248	K
		L	252	Q
		M	257	I
		R	260	S
		Q	277	R
		D	284	E
		A	287	P
		E	289	D
		T	290	E
		A	292	I
		A	299	G
		K	303	R
		V	309	L
		A	310	N
		Q	314	R

[0501]

		R	316	H
		Q	317	E
		R	318	A
		K	321	R
		A	322	G
Spy1245	203	L	72	M
		A	97	P
		Q	213	K
Spy1357	205	C	9	Y
		G	48	R
		I	87	L
		S	91	A
		T	102	A
		L	105	V
		A	111	S
		N	117	T
		E	139	S, A
		Q	142	K
		S	143	A
		N	145	T
		W	151	L
		A	155	D
		T	156	N
		P	157	A
		I	158	T, A
		A	159	S
		L	160	D
		D	161	A
		V	162	L
		K	163	E
		K	164	A
		T	165	L
		K	166	A
		T	168	Q
		K	169	T
		P	170	S, D
		V	171	A
		K	173	Q
		K	174	S
		G	187	S
		A	197	T
		V	207	A

[0502]

Spy1361	206	R	129	Q
		P	141	S
		L	197	P
		A	201	V
		D	230	A
		S	231	N
		D	235	N
		P	262	L
		T	272	N
		Q	274	H
		T	302	A
		T	308	I
		A	346	V
		V	354	F
		P	389	L
		M	391	K
		I	427	L
		P	431	L
		P	503	S
		D	645	N
Spy1390	210	S	696	P
		K	738	N
		T	757	A
		N	3	Q
		S	4	M
		A	9	T
		S	10	G
		S	16	T
		M	18	V
		A	19	T
		A	21	S
		T	26	S
		N	27	H
		D	28	N
		V	32	L
		I	33	V
		S	41	T
		V	54	L
		S	55	A
		N	61	S
		A	70	T
		G	73	A

[0503]

D	74	N
K	78	D
H	86	K
K	87	Q
E	90	D
A	94	T
S	97	K
A	98	T
A	99	V
S	104	G
F	110	Y
R	112	K
S	116	L
S	117	T
A	127	Q
K	130	N
L	132	I
T	133	S
T	134	K
Q	135	K
E	136	D
K	138	R
K	139	Q
E	142	D
S	143	A
A	149	T
V	150	A
M	152	I
I	153	M
T	154	Q
L	155	F
D	156	E
N	157	K
E	158	D
T	160	D
S	163	A
V	164	A
T	176	A
T	184	I
T	185	A
P	186	A
E	187	D

[0504]

V	190	T
K	193	T
A	198	E
T	199	I
N	200	T
V	201	L
T	203	A
D	204	E
K	207	R
S	211	G
N	213	K
G	216	N
I	217	R
D	219	E
V	220	I
S	222	T
V	223	A
T	227	A
S	228	T
Y	229	S
Q	230	K
K	231	R
K	232	T
F	233	Y
Y	234	H
V	236	I
E	243	T
S	246	A
Q	249	K
E	250	A
E	252	A
A	257	D
I	260	V
A	261	T
E	262	G
S	264	L
M	267	P
N	268	D
N	276	K
Y	297	F
N	299	K
L	300	P

[0505]

		G	301	N
		T	304	Q
		K	305	P
		A	307	Q
		S	308	K

[0506] Spy1536 215 无

[0507]

Spy1607	218	E	21	D
		A	91	P
		H	194	R
		D	204	G, N
Spy1666	220	K	90	Q
		K	302	T
		肺炎链球菌 TIGR4		
		V	37	I
		I	42	V
		S	56	A
		A	60	E
		G	67	S
		E	69	K
		C	74	A
		K	80	N
		V	87	K
		T	88	R
		K	90	A, Q
		S	91	P
		D	94	E
		Q	97	M
		K	109	Q
		R	111	C
		T	113	R
		A	114	E
		L	115	A
		D	118	Q
		L	124	C
		E	136	Q
		Q	145	K
		D	154	N
		R	155	Q
		Q	156	D
		S	157	A
		L	158	S

[0508]

		T	167	N
		P	169	D
		F	170	Y
		N	171	H
		K	175	R
		A	198	E
		I	199	V
		L	211	I
		A	214	L
		A	215	V
		D	252	Q
		E	255	D
		L	256	M
		D	287	E
		L	294	F
		E	297	D
		M	299	L
		F	303	M
		H	315	A
		S	316	E
		T	319	E
		K	322	N
		A	324	S
		A	332	V
		K	333	R
		R	336	H
Spy1972	228	V	32	M
		L	70	F
		M	98	I
		K	182	R
		F	224	S
		D	226	E
		H	245	P
		P	300	L
		R	363	K
		K	365	T
		T	369	A
		A	376	T
		R	443	K
		V	445	L
		A	460	T
		V	467	I

[0509]

		D	510	V
		A	496	T
		T	611	K
		T	718	A
		G	831	S
		A	913	V
		Q	930	K
		V	1053	A
		E	1079	D
		N	1094	D
		T	1102	I
		D	1103	G
		I	1149	V
Spy2000	232	K	27	N
		S	101	L
		V	151	I
		D	250	S
		P	335	S
		A	338	P
		V	519	I
Spy2025	238	S	33	N
		D	46	A
		D	49	A
		P	54	A
		T	78	N
		D	107	N
		K	109	N
		D	112	N
		P	119	S
		Q	147	P
		T	160	I
		D	170	E
		I	183	N
		I	194	A
		G	297	E
		S	528	R
Spy2191	244	A	70	V
		V	93	A

序列表

SPy0012

Seq ID 1

ATGAGAAAATTATTAGCGGCTATGTTAATGACTTTTTTCTGACTCCTTTACCAGTGATTAGTACAGAAAAA
AACTTATATTTTCA
AAAAATGCTGTTTATCAATTGAAACAAGATGTCGTTCAATCAACACAATTCTATAATCAAATACCCTCTAATC
CAAATCTTTATCAA
GAAACGTGTGCCTATAAAGACAGTGATTTAACTCTACCAGCAGGAAGATTAGGTGTAAATCAACCATTACTTA
TTAAATCGCTTG
TGCTTAACAAAGAATCTTTACCGGTTTTTGAGTTAGCTGATGGTACCTATGTTGAGGCTAATCGACAATTGAT
TTATGACGATATT
GTACTTAATCAAGTAGATATAGATAGCTATTTTTGGACACAAAAGAACTTAGGCTTTATTAGCCCCCTTATG
TTTTAGGTACGCA
ACAATTCCTTCTTCTTTTTTATTTGCTCAAAAAGTTCATGCCACTCAAATGGCACAACAAACCATGGAAC
TATTATCTTATTGA
TGATAAGGGCTGGGCATCACAAGAAGATCTAGTTCAATTTGATAACCGCATGTTAAAAGTCCAGGAAATGCTC
TTACAAAAATAT
AATAACCCAAATTATTCAATTTTTGTAAAGCAACTCAACACACAAACAAGTGCTGGTATTAATGCTGATAAAA
AAATGTATGCTGC
AAGTATCTCGAAGTTAGCACCCTTTATATTGTTCAAAAACAATTACAAAAAAGAAATTAGCAGAGAATAAA
ACTTTGACTTATA
CTAAAGATGTTAATCATTTTTTATGGAGACTATGATCCATTGGGAAGTGGTAAAATTAGTAAAATAGCTGATAA
TAAAGATTATCGT
GTTGAAGACCTACTGAAAGCTGTAGCACAAACAATCGGATAATGTAGCAACTAATTTTTAGGTTATTATCTAT
GTCATCAGTATGA
TAAAGCTTTCCGCTCAGAGATAAAAGCTTTATCAGGTATCGATTGGGATATGGAGCAGCGCTTATTAACCTCT
CGTTCAGCTGCA
AATATGATGGAAGCTATTTATCATCAAAAAGGCCAAATTTTCTTACCTTTCAAATACCGAATTTGATCAAC
AACGTATCACAAAA
AATATTACTGTTCCAGTTGCACATAAAATTGGTGATGCTTATGATTATAAACATGACGTTGCTATTGTTTACG
GTAATACTCCATTT
ATTTTGTCTATTTTTACAAATAAATCAACCTATGAAGATATTACGGCTATTGCAGATGACGTTTATGGTATTT
TAAAATGA

SPy0019

Seq ID 2

ATGAAAAAAGAATTTTATCAGCAGTTCTTGTAAGTGGTGTTACCCTCGGAGCAGCTACAACTGTAGGAGCGG
AGGATTTAAGT

ACTAAGATTGCTAAGCAAGATTCTATTATCTCAAACTGACTACAGAGCAAAAAGCTGCACAGAATCAAGTTT
CAGCGTTACAGG
CTCAAGTAAGTTCACTACAATCTGAACAAGATAAACTGACCGCAAGAAATACAGAACTTGAGGCGCTTTCAA
GCGATTTGAGCA
AGAAATTAAGGCTCTAACAAGTCAAATTGTTGCTCGTAATGAAAAATTAAAAATCAAGCTCGTAGTGCTTAT
AAAAACAATGAAA
CTTCTGGTTATATTAATGCACTTTTGAATTCTAAATCAATTTCTGATGTTGTAAACCGTTTAGTAGCAATTAA
TAGAGCTGTCTCTG
CTAACGCTAAATTGTTAGAACAACAAAAAGCTGATAAAGTTTCCCTTGAAGAAAAGCAAGCTGCTAACCAAAC
AGCTATTAATAC
CATTGCCGCTAATATGGCAATGGCTGAAGAAAACAAAAATACATTACGTACTCAACAAGCTAATTTGGTAGCT
GCAACTGCAAAT
TTAGCTCTCCAATTAGCATCTGCTACTGAAGATAAAGCTAATTTGGTAGCTCAAAAAGAAGCTGCAGAAAAAG
CTGCTGCTGAAG
CCTTAGCACAAGAACAGGCTGCTAAAGTTAAGGCACAAGAACAGGCTGCACAACAAGCAGCATCTGTTGAAGC
AGCAAAATCT
GCTATTACTCCAGCACCACAAGCTACTCCGGCAGCGCAAAGTAGTAATGCTATTGAACCAGCTGCACTCACGG
CTCCGGCAGC
TCCTTCTGCAGGACCACAAACATCATATGATTCTTCTAATACTTATCCAGTTGGACAATGCACATGGGGAGCT
AAATCTTTAGCT
CCTTGGGCAGGAAATAATTGGGGAAATGGTGGTCAATGGGCTTATAGTGCTCAAGCAGCTGGTTATCGTACTG
GTTCAACGCC
GATGGTAGGTGCGATTGCCGTTTGGAACGATGGTGGTTATGGACATGTCGCCGTTGTAGTTGAGGTTCAAAGT
GCCTCAAGTAT
TCGTGTGATGGAGTCTAACTACAGTGGTAGACAGTACATTGCTGACCACCGTGGCTGGTTTAATCCAACAGGT
GTTACATTTATT
TATCCACACTAA

SPy0025

Seq ID 3

ATGTCCTCCTATTTTCCAGTCGCTCCCTTGTCGGACTTGGTATCTTATATGAATAAACGTATTTTGTGAGA
AAAAGGCTGACTT
TGGTATTAAATCGGCTAGTCTTGTGAAAGAGTTGACGCATAATCTACAACTGACCTCTTTGAAGGCTTTGCGT
ATTGTGCAGGTC
TATGATGTCTTCAATTTGGCTGAGGATTTGCTGGCGCGTGCTGAGAAGCATATTTTCTCTGAGCAGGTGACAG
ACTGTCTTTTGA
CGGAACTGAAATCACTGCGGAGCTTGATAAGGTTGCCTTCTTTGCCATTGAGGCGCTTCCTGGTCAATTTGA
CCAACGTGCTG
CTAGTTCGCAAGAAGCTTTGCTATTATTTGGAAGTGACAGTCAGGTAAAGTCAATACAGCCCAGCTATACTT

GGTCAATAAGGA

TATTACAGAAGCAGAGCTTGAAGCCGTTAAGAACTATCTTTGAACCCTGTTGATTCGCGTTTCAAGGACATT
ACTTTGCCGCTT

GAAGAGCAGGCTTTCTCTGTATCTGATAAGACGATCCCTAATCTTGATTCTTTGAAACTTATCAAGCTGACG
ATTTTGCGACTTA

TAAGGCAGAGCAGGGCTTGGCTATGGAGGTCGATGACCTTCTCTTCATCCAAAATTATTTCAAATCAATCGGA
TGTGTGCCAAC

TGAGACTGAGTTGAAAGTTTTGGATACTTACTGGTCAGACCACTGCCGTCACACAACCTTTGAAACTGAATTG
AAGAACATTGAT

TTTTCAGCTTCTAAATTCCAAAAACAA TTGCAGACAACTTATGACAAATATATCGCCATGCGTGATGAGCTTG
GTCGTTCTGAAAA

GCCACAAACACTTATGGATATGGCGACTATTTTTGGTCGTTATGAGCGTGCCAACGGTCGTCTGGACGATATG
GAAGTCTCAGA

TGAAATCAATGCCTGCTCAGTTGAGATTGAAGTAGATGTTGATGGTGTGAAAGAGCCTTGGCTCCTCATGTTT
AAGAACGAGAC

TCACAATCACCCAACAGAAATTGAGCCATTCGGTGGAGCGGCGACTTGTATCGGTGGTGCTATTCGTGACCCA
TTGTCAGGAC

GTTTCATACGTTTATCAGGCTATGCGTATTTCAGGCGCAGGCGATATCACGACTCCGATTGCGGAAACACGTGC
TGGTAAATTGC

CACAACAAGTTATTTCTAAAACTGCGGCGCACGGCTATTCTTCATATGGTAACCAAATTGGGCTTGCGACAAC
TTATGTGCGCG

AGTACTTCCACCCTGGCTTCGTAGCCAAACGTATGGAGCTTGGAGCTGTGGTTGGTGCTGCACCTAAGGAAAA
TGTGGTTCGT

GAAAAACCAGAAGCAGGCGATGTGGTCATCTTGCTCGGTGGTAAACAGGTCGTGATGGTGTCGGTGGTGCGA
CAGGTTTCATC

TAAGGTTCAAACGGTTGAATCTGTGGAACAGCTGGCGCAGAGGTACAAAAAGGGAATGCCATTGAAGAACGT
AAGATTCAAC

GTCTTTTCCGTGATGGCAATGTCACTCGTCTTATTAAGAAATCAAATGACTTCGGTGCAGGTGGTGTCTGTGT
TGCCATCGGTG

AATTGGCTGACGGTCTTGAAATCGATTTGGACAAGGTGCCTCTTAAATACCAAGGTCTTAATGGTACTGAAAT
TGCAATCTCAGA

ATCTCAAGAGCGTATGTCAGTCGTTGTTTCGTCCAAATGATGTGGATGCCTTCATCGCAGCCTGCAACAAGGAA
AATATCGATGC

AGTCGTTGTTGCGACCGTTACTGAAAAACCAAATCTTGTCATGACTTGGAATGGCGAAATCATCGTTGATTTG
GAACGCCGTTTC

CTTGATACCAATGGTGTCCGTGTCGTTGTTGATGCTAAAGTCGTTGACAAGGACTTGACAGTTCCAGAAGCAC
GCACAACATCA

GCAGAGACACTTGAAGCAGATACGCTTAAGGTCTTGTCTGACCTCAACCACGCTAGTCAAAAAGGTCTTCAAA
CTATCTTTGACT

CATCTGTTGGTCGTTCAACCGTTAACCACCCAATCGGTGGTCGTTACCAAATCACACCGACAGAAAGTTCTGT
TCAAAAATTGCC
AGTTCAACATGGTGTGACAACAACTGCATCTGTTATGGCTCAAGGTTACAATCCTTATATTGCAGAGTGGTCA
CCTTATCACGGT
GCTGCCTATGCTGTCATTGAAGCGACAGCTCGCTTGGTAGCAACGGGTGCTGACTGGTCTCGTGACGTTTCT
CTTACCAAGA
GTACTTTGAGCGTATGGATAAACAGGCAGAGCGTTTTGGTCAGCCAGTATCAGCTCTTCTTGGTTCTATTGAG
GCTCAGATTCA
ACTTGGTTTTGCCATCAATCGGCGGTAAGGACTCTATGTCTGGTACTTTCGAAGACTTGACAGTACCACCAACC
TTGGTAGCTTT
CGGCGTGACAACAGCGGACAGCCGCAAGGTTCTCTCTCCTGAGTTTAAAGCGGCTGGCGAAAACATTTACTAT
ATCCCAGGTC
AAGCTATTTTCAGAAGATATTGATTTTGACCTTATCAAGGATAACTTTAGCCAGTTTGAAGCTATTCAAGCTCA
ACATAAGATTACA
GCTGCCTCAGCCGCTAAATACGGTGGTGTCTAGAAAAGTCTTGCTCTCATGACTTTTGGTAACCGTATCGGTG
CTTCTGTTGAA
ATTGCAGAGCTTGACAGCAGCTTGACAGCTCAACTCGGAGGTTTGTCTTTACATCAGCTGAGGAAATTGCTG
ACGCGGTGAAA
ATCGGTCAAACCTCAGGCAGACTTTACAGTCACTGTCAATGGAAATGACCTTGCTGGCGCTAGCCTTCTAGCAG
CCTTCGAAGGC
AAATTGGAAGAGGTTTACCCAACAGAATTTGAGCAGACAGATGTTCTTGAAGAAGTTCCTGCTGTGGTATCAG
ATACTGTTATCA
AGGCTAAGGAAACAATTGAAAAACCAGTGGTTTACATTCAGTCTTCCCTGGTACCAACTCAGAATACGATTC
AGCTAAGGCCTT
TGAACAGGTTGGAGCTAGTGTCAACTTGGTACCATTTGTAACCTTGAATGAGGTTGCTATTGCTGAGTCAGTT
GACACTATGGTT
GCTAATATTGCTAAGGCAAATATCATCTTCTTTGCTGGAGGTTTCTCAGCAGCGGATGAACCAGATGGGTCTG
CTAAGTTTATCG
TCAATATCTTGCTTAACGAGAAGGTCCGCGCAGCTATTGACAGCTTCATCGAAAAAGGTGGCCTTATCATCGG
TATCTGTAATG
GTTTCCAAGCCCTTGTTAAATCAGGTCTTCTTCCATACGGAACTTCGAGGAAGCTGGTGAGACAAGTCCAAC
TCTCTTCTATAA
TGATGCTAATCAGCATGTTGCCAAGATGGTTGAGACTCGTATCGCAAATACCAACTCACCTTGGTTGGCAGGA
GTTGAGGTCCG
CGATATTCATGCCATTCCAGTTTCACATGGTGAAGGTAAACTTGTTGTCAGCGCTTCTGAATTTGCAGAGCTA
AGAGACAATGGT
CAAATCTGGAGCCAATATGTGGACTTTGACGGACAACCATCTATGGATTCTAAATACAA TCCAAACGGCTCTG
TCAATGCCATCG
AAGGGATTACCAGCAAGAATGGTCAAATCATCGGTAAGATGGGACACTCAGAACGCTGGGAAGACGGACTCTT

CCAAAATATC

CCTGGTAACAAAGACCAAAATCCTCTTTGCAAGTGCTGTAAAATACTTTACAGGGAAGTAA

SPy0031

Seq ID 4

ATGAAAAAATTTTCATCGTTTTTTGGTCTCAGGAGTAATCCTTTTAGGTTTTAATGGTCTAGTACCTACTATGC
CATCTACACTTATT

TCGCAACAGGAAAAATCTTGTTTCATGCAGCTGTTTTAGGCGATAACTATCCGAGTAAGTGGAAGGCAATG
GAATCGATTCTG

TGGAACATGTATATCCGCCAATGCACTTCTTTTGCAGCTTTTCGTTTAAGCTCTGCTAATGGTTTTTCAGTTAC
CTAAAGGCTACG

GTAATGCCTGCACGTGGGGACATATCGCGAAAAATCAGGGTTATCCTGTGAATAAGACACCAAGCATAGGGGC
TATCGCTTGG

TTTGATAAAAAACGCTTATCAGTCAAATGCTGCTTACGGTCATGTAGCATGGGTAGCTGATATCCGTGGAGACA
CTGTCACTATCG

AAGAGTATAATTACAACGCTGGACAAGGCCCTGAAAGATACCATAAGCGTCAAATTCCTAAAATCTCAGGTAAG
TGTTTATATCCA

TTTTAAAGACTTATCATCTCAGACAAGTCATTCCCTACCCAAGACAACCTAAAACACATTTCTCAAGCTTCATTT
GACCCCTCTGGAA

CTTATCACTTTACAACCAGATTACCAGTCAAAGGACAAACCAGTATCGATAGCCCTGATCTTGCTTACTATGA
AGCAGGTCAATC

TGTTTATTACGATAAAGTCGTGACTGCTGGAGGTTATACATGGCTTAGCTACCTCAGTTTTTCTGGAACCGA
CGCTATATTCCC

ATTAAAGAGCCCGCACAGTCTGTGGTTCAAAATGACAATACAAAACCTTCCATTAAGGTCGGTGATACTGTTA
CCTTCCCTGGC

GTTTTTCGTGTAGATCAGCTTGTTAATAATTTGATCGTTAATAAAGAATTAGCCGGAGGAGACCCAACTCCAC
TAAACTGGATTG

ATCCACACCATTAGATGAAACAGATAACCAAGGAAAAGTTTTAGGAGATCAAATTCCTCCGTGTGGGTGAATA
TTTTATCGTCAC

TGGTAGTTATAAAGTATTAAAAATTGATCAACCAAGTAATGGTATTTATGTTCAAATCGGATCTCGTGGAACA
TGGGTAAATGCTG

ATAAAGCTAACAAATTATAG

SPy0103

Seq ID 5

ATGATTAATCAATGGAACAACTTACGACACAAGAAGCTAAAAGGATTTACTCTTCTAGAAATGTTATTGGTGA
TTCTTGTCATCAG

TGTTTTGATGCTATTATTTGTGCCTAATTTAAGCAAGCAAAAAGACAGGGTTACAGAAACAGGTAATGCCGCT
GTTGTTAAATTA

GTGGAGAATCAAGCAGAACTATATGAATTATCTCAAGGCTCAAAACCAAGTTTGAGCCAGTTAAAGGCAGATG
GTAGTATCACT

GAGAAACAAGAAAAAGCTTATCAAGACTATTATGACAAACATAAAAAATGAAAAAGCCCGTCTTAGCAATTAA

SPy0112

Seq ID 6

ATGAAAATTGGCATTATTGGTGTGGCAAAATGGCTAGCGCTATCATCAAAGGCCTTAAACAAACACCCCATG
AACTTATCATTT

CAGGATCATCTTTAGAACGGTCCAAGGAAATTGCGGAGCAGTTAGCACTGCCTTATGCTATGTCCCACCAAGA
CCTTATTGACC

AGGTTGATCTTGTTATTTTAGGCATCAAGCCTCAACTATTTGAAACGGTACTCAAACCGCTTCACTTCAAACA
GCCTATTATATCT

ATGGCAGCAGGCATTTCCCTTCAACGACTAGCAACATTCGTAGGACAAGACCTTCCGCTGCTACGTATCATGC
CAAACATGAAT

GCACAAATTCTCCAAAGCAGTACCGCTTTAACGGGAAATGCTTTGGTGTCCCAGGAATTACAAGCACGTGTTC
GAGACTTAACA

GATAGCTTTGGTAGCACATTTGATATTAGTGAAAAGGATTTTGACACCTTTACCGCTTTAGCAGGGTCAAGTC
CTGCCTATATTT

ATCTCTTTATTGAGGCTTTGGCTAAGGCTGGCGTCAAGAATGGCATACCTAAAGCAAAGGCGCTGGAGATTGT
TACTCAAACAG

TATTGGCTAGCGCCAGCAATCTCAAGACCAGTTCTCAAAGTCCGCACGATTTTCATTGACGCTATTTGTAGCCC
CGGTGGCACAA

CTATTGCTGGTCTGATGGAGTTAGAACGCCTTGGCCTCACAGCTACTGTCAGCTCTGCCATTGACAAAACCAT
CGATAAAGCTA

AAAGCTTGTA

SPy0115

Seq ID 7

ATGACAGACTTATTCTCAAAAAATCAAAGAAGTTACCGAACTGGATGGCATTGCGGGCTATGAACATAGCGTTC
GTGACTACCTA

CGCACCAAAATAACCCCGCTGGTTGACCGTGTGAAACAGACGGGCTTGGTGGCATTTTTGGTATCAGAGATA
GTAAAGCTGAA

AAAGCCCCCGTATTTTAGTAGCTGCGCACATGGACGAAGTCGGTTTTATGGTCAGTGATATCAAAGTTGACG
GAACGCTACGC

GTGGTTGGTATCGGTGGTTGGAACCCACTTGTGTGAGTTCACAACGGTTTACCCTTTACACACGCACTGGCC
AAGTTATTCCC

CTTATTTTCAGGATCGGTACCTCCCCATTTTTTACGTGGGGCAAATGGCTCTGCTAGTCTACCACATATCGAAG
ATATTGTGTTTG

ATGGTGGCTTTACGGATAAGGCAGAAGCTGAAAGATTTGGTATTACACCGGGTGATATTATTATCCCTCAATC

TGAAACGATCCT

AACAGCCAATCAAAAAAATATTATTTCAAAAGCTTGGGACAATCGCTATGGCGTTCTCATGATAACAGAAATG
CTTGAAGCGTTA

AAAGGACAAGACCTTAACAACACCCTAATTGCAGGTGCTAACGTTCAAGAAGAAGTTGGTCTGCGCGGAGCCC
ACGTCTCAAC

CACCAAGTTCGACCCTGAACTCTTTTTTCGCAGTAGATTGTTGCGCTGCTGGTGATATTTATGGCAATCCTGGA
ACAATCGGAGAT

GGTACCTTGTTGCGTTTCTACGACCCAGGCCATGTCATGCTCAAAGATATGCGCGACTTCTTACTGACTACTG
CTGAGGAAGCT

GGTGTCAATTTCCAATACTATTGTGGCAAGGGAGGCACAGATGCAGGTGCTGCACACCTTCAAAATGGTGGTG
TCCCATCAACA

ACCATCGGAGTCTGTGCACGCTACATTCACTCTCATCAAACCCTCTACGCTATGGATGATTTCTGTAAGAGCCC
AAGCCTTCTTAC

AAGCCATTATCAAAAAACTGGATCGCTCAACCGTTGACTTGATTAAATGTTACTAA

SPy0166

Seq ID 8

ATGGAAGATATTTCTGATCCAGAAGTTATTTTAGAGTATGGGGTTTACCCTGCTTTCATAAAAGGCTATACCC
AATTGAAAGCTAA

CATCGAAGAAGCATTATTAGAAATGTCAAATAGCGGTCAAGCATTAGACATTTACCAAGCAGTTCAAACCCTA
AACGCTGAAAAC

ATGTTATTAAATTATTACGAAAGCTTGCCATTTTATTTAAACCGTCAAAGCATACTAGCTAATATGACCAAAG
CGTTAAAAGATGC

GCATATTAGAGAGGCTATGGCACATTACAAATTAGGAGAATTTGCTCACTATCAAGATACTATGCTTGATATG
GTCGAAAGAACA

ATAAAAACATTTTAG

SPy0167

Seq ID 9

ATGTCTAATAAAAAAACATTTAAAAAATACAGTCGCGTCGCTGGGCTACTGACGGCAGCTCTTATCATTGGTA
ACCTTGTTACTG

CTAATGCTGAATCGAACAAACAAAACTGCTAGTACAGAAACCACAACGACAAATGAGCAACCAAAGCCAGA
AAGTAGTGAGC

TAATACTGAAAAAGCAGGTGAGAAAACGGATGATATGCTTAACTCTAACGATATGATTAAGCTTGCTCCCAA
AGAAATGCCACT

AGAATCTGCAGAAAAAGAAGAAAAAAGTCAGAAAGACAAAAAAGAGCGAAGAAGATCACACTGAAGAAATC
AATGACAAGAT

TTATTCATAAATTATAATGAGCTTGAAGTACTTGCTAAAAATGGTGAAACCATTGAAAATTTTGTTCCTAAA
GAAGGCGTTAAGA

AAGCTGATAAATTTATTGTCATTGAAAGAAAGAAAAAATATCAACACTACACCAGTCGATATTTCCATTAT
TGACTCTGTCACT
GATAGGACCTATCCAGCAGCCCTTCAGCTGGCTAATAAAGGTTTTACCGAAAACAAACCAGACGCGGTAGTCA
CCAAGCGAAA
CCCACAAAAATCCATATTGATTTACCAGGTATGGGAGACAAAGCAACGGTTGAGGTCAATGACCCTACCTAT
GCCAATGTTTCA
ACAGCTATTGATAATCTTGTTAACCAATGGCATGATAATTATTCTGGTGGTAATACGCTTCCTGCCAGAACAC
AATATACTGAATC
AATGGTATATTCTAAGTCACAGATTGAAGCAGCTCTAAATGTTAATAGCAAAATCTTAGATGGTACTTTAGGC
ATTGATTTCAAGT
CGATTTCAAAAGGTGAAAAGAAGGTGATGATTGCAGCATACAAGCAAATTTTTTACACCGTATCAGCAAACCT
TCCTAATAATCC
TGCGGATGTGTTTGATAAATCGGTGACCTTTAAAGAGTTGCAACGAAAAGGTGTCAGCAATGAAGCTCCGCCA
CTCTTTGTGAG
TAACGTAGCCTATGGTCGAACTGTTTTGTCAAACCTAGAAAACAAGTTCTAAAAGTAATGATGTTGAAGCGGCC
TTTAGTGCAGCT
CTAAAAGGAACAGATGTTAAACTAATGGAAAAATTTCTGATATCTTAGAAAATAGCTCATTTACAGCTGTCTG
TTTTAGGAGGAGA
TGCTGCAGAGCACAATAAGGTAGTCACAAAAGACTTTGATGTTATTAGAAACGTTATCAAAGACAATGCTACC
TTCAGTAGAAAA
AAGCCAGCTTATCCTATTTTCATACACCAGTGTTTTCCTTAAAAATAATAAAATTGCGGGTGTCAATAACAGAA
CTGAATACGTTGA
AACAACATCTACCGAGTACACTAGTGGAAAAATTAACCTGTCTCATCAAGGCGCGTATGTTGCTCAATATGAA
ATCCTTTGGGAT
GAAATCAATTATGATGACAAAGGAAAAAGAAGTGATTACAAAACGACGTTGGGACAACAACCTGGTATAGTAAGA
CATCACCATTTA
GCACAGTTATCCCACTAGGAGCTAATTCACGAAATATCCGTATCATGGCTAGAGAGTGCAGTGGCTTAGCTTG
GGAATGGTGGC
GAAAAGTGATCGACGAAAGAGATGTGAAACTGTCTAAAGAAATCAATGTCAATATCTCAGGATCAACCTTGAG
CCCATATGTTT
GATTACTTATAAGTAG

SPy0168

Seq ID 10

ATGAAACAACAATCTTACCAGCCTCTACGCTTCGTCTACCTCTTGGTGGCTCTATTTGCTGCTCTGTTGCTTA
TAGCAAGACCTG
TTATGGCAGATGAGGGAACAAATAGTGCTGATGCGGCGTATTATAAAGGGCAAAGTGCTGGAAAAAAGCAGG
GAAAAAAGCT
GGAAAAGAAGCTACTTGGACTGATTTGACCCCAACTGTCCCAACTAATCCAGAAACACCTAGTGACATCGGAG

AGACTACTAAT

AAACAGCTCTATAAAGAAGGGTATAAAGATGGGTACAAAGAGGGTTATAATGAAGGCTGGAAATCTCAGTATC
CCGTTTTGACT
CCGGTCAAGGTTATATGGGATTTGATCTCTTATTGGCTACAGCGATTATCCCCAATAATCAGTCAAGTACCG
CAGCACAAAGCA
TGTCATAA

SPy0171

Seq ID 11

GTGAAAAACAAATTATTTTTAGTTGCCCTTGCGACCGTAACTGTCCTAGGGCCGTCTTTAGCAACCCCTCATC
ACCAGACCGTG
CATGCTAGTGATGTAACATTAAGTGAACATGTGATAAAAAACGGAACAGTATGTTTTGGCTACGAAAACGTAG
ATGGTGAAGTAT
GTAAATTAAACAGCTGACGGAAGGGAACCATTTGTGTGGGTTACGAAAATAGAGACATAAAAGAGAGTGAAAC
TTCTAGCACCA
AAAATGATTGTTCTAATTGGTTTTGGTGCTTTTTAAATTATCTTTGGACTACAATAAAAAGCTGGGTTTCGTA
A

SPy0183

Seq ID 12

ATGGAAACAATTTTAGAAGTCAAAACATCTCAGTAAAAATTTTGGCAAAAAACAAAAAGCAGCTCTTGAGATGG
TAAAGACTGGCA
AAAATAAGAGTGAGATTTTAAAGAAAAACAGGCGCTACTGTAGGTGTCTATGACGCTAGTTTTGAGGTCAAAAA
AGGTGAAATCTT
TGTTATTATGGGGCTATCAGGAAGTGGGAAATCAACCCTTGTCGCATGCTAAATCGTTTGATTGAACCTTCA
GCAGGATCTATC
TTGCTGGAAGGTAAAGACATCTCAACCATGTCAGCAGATCAGCTGCGTGAGGTGCGCCGCCATGACATTAACA
TGGTTTTCCAA
AGCTTTGCCCTCTTTCCTCATAAAACCATTTTGAAAAATACCGAATTTGGTTTGGAATTACGTGGCGTTCCCA
AAGAAGAAGCC
AGCGATTGGCAGAAAAAGCCCTTGATAATTCAGGCCTATTAGATTTTAAAGACCAGTACCCAAACCAACTATC
TGGTGGGATGC
AGCAGCGTGTGCGCCTAGCCCGTGCGCTAGCTAATAGCCCTAAAAATCTCTTAATGGACGAGGCATTCTCAGC
GCTTGATCCTT
TGATTCGTGCTGAGATGCAAGATGAATTACTTGATTTGCAAGACAGCATGAAACAAACCATCATCTTTATCAG
TCATGACTTGAA
TGAAGCCTTGCGGATTGGTGATCGGATTGCTTTGATGAAAGACGGACAAATTATGCAAATTGGTACTGGTGAG
GAAATCTTGAC
TAACCCAGCCAATGACTTTGTGCGTGAATTCGTTGAGGATGTGGACCGTTCTAAAGTCTTGACAGCACAAAAT

ATCATGATCAAA

CCGTTAACAACACTACTGTTGAATTAGATGGACCTCAAGTTGCCTTGAACCGTATGCACAACGAAGAGGTGTCTA
TGTTGATGGCG

ACGAATCGCCGCCGCCAATTAGTCGGTAGTTTGACGGCCGATGCCGCTATAGAGGCGCGCAAAAAAGGGTTAC
CGCTATCAGA

AGTGATTGATCGCGATGTGAGAACTGTCTCAAAAGATACTATTATTACAGATATTTTGCCTCTTATCTATGAT
TCATCTGCTCCGA

TTGCAGTGACAGATGATAATAATCGTCTGTTAGGTGTCATTATTCGAGGACGAGTGATTGAAGCCTTGGCTAA
TATCTCAGACGA

AGACCTTAACTAA

SPy0230

Seq ID 13

ATGAAAACAGCACGTTTTTTCTGGTTTTATTTTAAACGCTATCGTTTCTCATTTACTGTCATTGCTGTTGCCG
TTATCTTAGCAACT

TATTTACAAGTAAAAGCTCCTGTCTTCTTAGGAGAGTCCTTGACTGAGTTGGGAAAAATCGGTCAGGCTTATT
ACGTTGCTAAGA

TGAGTGGCCAGACACATTTTAGCCCTGATTTATCAGCTTTTAATGCCGTGATGTTTAAGCTTTTGATGACTTA
TTTCTTTACTGTT

TTAGCTAATCTAATATATAGTTTCTTACTTACACGTGTTGTCTCACATTCGACTAACCGCATGCGCAAGGGCT
TATTTGGTAAATT

AGAACGTTTAAACGTCGCCTTTTTTGACCGCCATAAAGATGGGGAGATTCTTTCTCGTTTCACGAGTGATTG
GATAATATTCAA

AACTCGCTGAACCAATCCTTGATTCAAGTGGTGACTAATATTGCCCTTTACATCGGCCTGGTCTGGATGATGT
TTAGGCAAGATA

GCCGTTTAGCTTTGTAAACCATCGCATCAACCCAGTTGCTCTCATTTTTTTTAGTGATTAACATCCGTTTGGC
AAGAAAATACACC

AATATCCAACAGCAAGAAGTCAGTGCTTTAAATGCTTTTATGGATGAAACCATTTTCAGGACAAAAGGCTATTA
TTGTACAAGGTG

TCCAAGAAGATACGATGACAGCCTTTTTAAAGCATAATGAAAGGGTTCGACAAGCCACCTTCAAACGCCGTCT
GTTCTCAGGAC

AATTATTTCCAGTCATGAATGGAATGAGCCTTATTAACACGGCTATCGTGATTTTTGTGCGTTCAACAATTGT
CCTCAGTGACAAA

TCTATGCCAGCAGCGGCAGCGCTTGGTTTAGTGGTTACTTTTGTACAATATTTCCAGCAATATTACCAACCCA
TGATGCAAATCG

CGTCTAGTTGGGGAGAATTGCAGCTGGCCTTTACCGGTGCTCACCGTATTCAAGAAATGTTTGATGAAACCGA
AGAAGTTCGTC

CACAAAATGCACCAGCGTTCACCAGCTTAAAAAGAAGCAGTGGCGATTAAACCACGTCGATTTTGGGTATCTTCC
TGGGCAAAAAG

TTTTATCAGATGTGTCAATCGTTGCACCCAAGGGCAAAATGATTGCCGTGGTTGGACCGACAGGTTCTGGAAA
GACCACTATTA

TGAACTTGATTAACCGTTTCTACGATGTGGATGCAGGTCGATTACCTTTGATGGCCGTGATATTCGTGACTA
CGATTTGGATAG

TCTTCGTCAAAAGGTAGGGATTGTGTTGCAAGAGTCAGTTCTTTTTTCAGGAACCATTACGGATAATATTCGT
TTTGGTGATCAG

ACCATTAGTCAAGACATGGTTGAAACTGCTGCGCGTGCGACCCATATTCATGACTTTATCATGTCCTTACCAA
AAGGTACAATA

CCTATGTCTCAGATGATGACAATGTCTTTTCAACAGGTCAAAAGCAGTTGATTTCTATTGCTAGGACGCTACT
GACTGACCCTGA

AGTGTTGATTTTGGATGAGGCCACTTCAAATGTTGATACGGTTACCGAAAGTAAAATTCAACGGGCCATGGAA
GCTATCGTGGC

AGGTCGAACTAGCTTTGTCATTGCTCACCGCCTCAAAACCATTTTAAATGCCGATCACATTATTGTGTTGAAA
GATGGCAAGGTC

ATTGAGCAAGGAAATCATCATGAGCTATTGCATCAAAAAGGCTTTTATGCCGAATTGTATCACAATCAATTTG
TCTTTGAATAG

SPy0269

Seq ID 14

ATGGACTTAGAACAAACGAAGCCAAACCAAGTTAAGCAGAAAAATTGCTTTAACCTCAACAATTGCTTTATTGA
GTGCCAGTGTAG

GCGTATCTCACCAAGTCAAAGCAGATGATAGAGCCTCAGGAGAAACGAAGGCGAGTAATACTCACGACGATAG
TTTACCAAAAC

CAGAAACAATTCAAGAGGCAAAGGCAACTATTGATGCAGTTGAAAAAACTCTCAGTCAACAAAAAGCAGAACT
GACAGAGCTTG

CTACCGCTCTGACAAAACTACTGCTGAAATCAACCACTTAAAAGAGCAGCAAGATAATGAACAAAAAGCTTT
AACCTCTGCACA

AGAAATTTACACTAATACTCTTGCAAGTAGTGAGGAGACGCTATTAGCCCAAGGAGCCGAACATCAAAGAGAG
TTAACAGCTAC

TGAAACAGAGCTTCATAATGCTCAAGCAGATCAACATTCAAAAGAGACTGCATTGTCAGAACAAAAAGCTAGC
ATTTCAGCAGAA

ACTACTCGAGCTCAAGATTTAGTGGAACAAGTCAAAACGTCTGAACAAAATATTGCTAAGCTCAATGCTATGA
TTAGCAATCCTG

ATGCTATCACTAAAGCAGCCAAACGGCTAATGATAATACAAAAGCATTAAGCTCAGAATTGGAGAAGGCTAAA
GCTGACTTAGA

AAATCAAAAAGCTAAAGTTAAAAAGCAATTGACTGAAGAGTTGGCAGCTCAGAAAGCTGCTCTAGCAGAAAAA
GAGGCAGAACT

TAGTCGTCTTAAATCCTCAGCTCCGTCTACTCAAGATAGCATTGTGGGTAATAATACCATGAAAGCACCGCAA
GGCTATCCTCTT

GAAGAACTTAAAAAATTAGAAGCTAGTGGTTATATTGGATCAGCTAGTTACAATAATTATTACAAAGAGCATG
CAGATCAAATTAT
TGCCAAAGCTAGTCCAGGTAATCAATTAAATCAATACCAAGATATTCCAGCAGATCGTAATCGCTTTGTTGAT
CCCATAATTG
ACACCAGAAGTGCAAAATGAGCTAGCGCAGTTTGCAGCTCACATGATTAATAGTGTGAAGAAGACAATTAGGTC
TACCACCAGTT
ACTGTTACAGCAGGATCACAAGAATTTGCAAGTACTTAGTACCAGCTATAAGAAAACATGGTAATACAA
GACCATCATTTG
TCTACGGACAGCCAGGGTATCAGGGCATTATGGTGTGGGCCTCATGATAAACTATTATTGAAGACTCTGC
CGGAGCGTCA
GGGCTCATTCGAAATGATGATAACATGTACGAGAATATCGGTGCTTTTAACGATGTGCATACTGTGAATGGTA
TTAAACGTGGTA
TTTATGACAGTATCAAGTATATGCTCTTTACAGATCATTTACACGGAAATACATACGGCCATGCTATTAACTT
TTTACGTGTAGAT
AAACATAACCCTAATGCGCCTGTTTACCTTGGATTTTCAACCAGCAATGTAGGATCTTTGAATGAACACTTTG
TAATGTTTCCAGA
GTCTAACATTGCTAACCATCAACGCTTTAATAAGACCCCTATAAAAGCCGTTGGAAGTACAAAAGATTATGCC
CAAAGAGTAGGC
ACTGTATCTGATACTATTGCAGCGATCAAAGGAAAAGTAAGCTCATTAGAAAATCGTTTGTGGCTATTTCATC
AAGAAGCTGATA
TTATGGCAGCCCAAGCTAAAGTAAGTCAACTTCAAGGTAAATTAGCAAGCACACTTAAGCAGTCAGACAGCTT
AAATCTCCAAGT
GAGACAATTAAATGATACTAAAGGTTCTTTGAGAACAGAATTACTAGCAGCTAAAGCAAAACAAGCACAACTC
GAAGCTACTCGT
GATCAATCATTAGCTAAGCTAGCATCGTTGAAAGCCGCACTGCACCAGACAGAAGCCTTAGCAGAGCAAGCCG
CAGCCAGAGT
GACAGCACTGGTGGCTAAAAAGCTCATTTGCAATATCTAAGGGACTTTAAATTGAATCCTAACCGCCTTCAA
GTGATACGTGAG
CGCATTGATAATACTAAGCAAGATTTGGCTAAAACTACCTCATCTTTGTTAAATGCACAAGAAGCTTTAGCAG
CCTTACAAGCTAA
ACAAAGCAGTCTAGAAGCTACTATTGCTACCACAGAACACCAGTTGACTTTGCTTAAACCTTAGCTAACGAA
AAGGAATATCGC
CACTTAGACGAAGATATAGCTACTGTGCCTGATTTGCAAGTAGCTCCACCTCTTACGGGCGTAAACCGCTAT
CATATAGTAAGA
TAGATACTACTCCGCTTGTTCAAGAAATGGTTAAAGAAACGAAACAACTATTAGAAGCTTCAGCAAGATTAGC
TGCTGAAAATAC
AAGTCTTGTAGCAGAAGCGCTTGTTGGCCAAACCTCTGAAATGGTAGCAAGTAATGCCATTGTGTCTAAAATC
ACATCTTCGATT
ACTCAGCCCTCATCTAAGACATCTTATGGCTCAGGATCTTCTACAACGAGCAATCTCATTTCTGATGTTGATG

AAAGTACTCAAA

GAGCTCTTAAAGCAGGAGTCGTCATGTTGGCAGCTGTCGGCCTCACAGGATTTAGGTTCCGTAAGGAATCTAA
GTGA

SPy0287

Seq ID 15

ATGACAAAAGAAAACTAGTGGCTTTTTCGCAAGCCCACGCTGAGCCTGCTTGGCTGCAAGAACGGCGTTTAG
CGGCATTAGA

AGCCATTCCAAATTTGGAATTACCAACCATCGAAAAGGTTAAATTTACCGTTGGAATCTAGGAGATGGTACC
TTAACAGAAAAT

GAAAGTCTAGCTAGTGTTCAGATTTTATAGCTATTGGAGATAACCCAAAGCTTGTTTCAGGTAGGCACGCAAA
CAGTCTTAGAAC

AGTTACCAATGGCGTTAATTGACAAGGGAGTTGTTTTTCAGTGATTTTTATACGGCGCTTGAGGAAATCCCAGA
AGTAATTGAAGC

TCATTTTGGTCAGGCATTAGCTTTTGATGAAGACAACTAGCTGCCTACCACACTGCTTATTTTAATAGCGCA
GCCGTGCTCTAC

GTTCTGATCACTTGGAATCACAACCTCCTATTGAAGCTATTTTCTTACAAGATAGTGACAGTGACGTTTCCTT
TTAACAAGCATGT

TCTAGTGATTGCAGGAAAAGAAAGTAAGTTCACCTATTTAGAGCGTTTTGAATCTATTGGCAATGCCACTCAA
AAGATCAGCGCT

AATATCAGTGTAGAAGTGATTGCTCAAGCAGGCAGCCAGATTAAATTTCTCGGCTATCGACCGCTTAGGTCCTT
CAGTGACAACC

TATATTAGCCGTCGAGGACGTTTAGAGAAGGATGCCAACATTGATTGGGCCTTAGCTGTGATGAATGAAGGCA
ATGTCATTGCT

GATTTTGACAGTGATTTGATTGGTCAGGGCTCACAAGCTGATTTGAAAGTTGTTGCAGCCTCAAGTGGTCGTC
AGGTACAAGGT

ATTGACACGCGGTGACCAACTATGGTCAACGTACGGTCGGTCATATTTTACAGCATGGTGTGATTTTGGAAC
GTGGCACCTTA

ACGTTTAAACGGGATTGGTCATATTCTAAAAGACGCTAAGGGAGCTGATGCTCAACAAGAAAGCCGTGTTTTGA
TGCTTTCTGAC

CAAGCAAGAGCCGATGCCAATCCAATCCTCTTAATTGATGAAAAATGAAGTAACAGCAGGTCATGCAGCTTCTA
TCGGTCAGGTT

GACCCTGAAGATATGTATTACTTGATGAGTCGAGGACTGGATCAAGAAACAGCAGAACGATTGGTTATTAGAG
GATTCCTAGGA

GCGGTTATCGCTGAAATTCCTATTCCATCAGTCCGCCAAGAGATTATTAAGGTTTTAGATGAGAAATTGCTTA
ATCGTTAA

SPy0292

Seq ID 16

ATGATCAAACGATTAATTTCCCTAGTGGTCATCGCCTTATTTTTTGCAGCAAGCACTGTTAGCGGTGAAGAGT
ATTCGGTAACTG
CTAAGCATGCGATTGCCGTTGACCTTGAAAGTGGCAAAGTTTTATACGAAAAAGATGCTAAAGAAGTTGTCCC
AGTCGCCTCAG
TCAGTAAGCTCTTGACAACCTATCTGGTTTACAAAGAAGTTTCTAAGGGCAAGCTAAATTGGGATAGTCCTGT
AACTATTTCTAA
CTACCCTTATGAACTCACTACAACTATACTATTAGTAACGTTCCCTCTTGATAAGAGAAAATATACCGTTAAA
GAACTTTTAAAGTG
CGTTAGTTGTTAATAACGCCAATAGCCCCGCTATTGCTTTAGCTGAAAAAATAGGCGGAACCGAACCCAAATT
TGTTGACAAAAT
GAAAAACAATTAAGACAATGGGGCATTTCGGATGCAAAGGTCGTCAATTCAACTGGCTTAACTAACCATTTT
TTAGGAGCTAAT
ACTTATCCTAATACAGAACCAGATGATGAAAAATTGTTTTTGCGCCACTGATTTAGCTATTATTGCCAGGCATC
TCTTATTAGAATT
TCCAGAAGTACTGAAATTATCTAGCAAATCCTCCACTATTTTTGCTGGACAAACCATTTACAGTTATAATTAC
ATGCTTAAAGGCA
TGCCTTGTTATCGAGAAGGCGTGGATGGTCTTTTTGTTGGTTATTCTAAAAAAGCCGGTGCTTCTTTGTAGC
TACTAGTGTGCA
AAATCAAATGAGGGTTATTACAGTAGTTTTAAATGCTGATCAAAGCCACGAGGATGATTTAGCTATATTTAAA
ACAACCAATCAAT
TGTTGCAGTACCTTTTAATTAATTTTCAAAAAAGTCCAGTTAATTGAAAATAATAAACAGTAAAAACGTTATA
TGTCTTAGACAGTC
CTGAAAAAAGTGTCAAACCTGTAGCCCCAAAATAGTTTATTTTTTATCAAACCAATACATACAAAGACCAAAAA
TACCGTCCATATTA
CTAAGAAATCATCCACAATGATCGCACCTCTATCAAAGGGACAAGTCTTAGGTAGAGCAACCCTTCAAGATAA
ACATCTTATTGG
ACAAGTTATCTGGATACTCCTCCTTCTATCAATCTTATCCTTCAAAAAACATTTCTAAAAGTTTCTTTTAA
AAGGTCTGGTGAA
CCGTTTTGTGAGGTATGTCAATACCTCTTTATAG

SPy0295

Seq ID 17

ATGGAATCGATTGATAAATCTAAATTTGATTTGTTGAGCGGATAGTGAAGCCTCCGAAGTGATTGATACCC
CTGCTTATTCTT
ACTGGAATCAGTGTTTCGTCAGTTTTTTTCTAAAAAATCTACAGTCTTTATGCTCGTAATTTTAGTGACAGT
CTTGATGATGAGC
TTTATTTATCCAATGTTTGCCAACTACGACTTTAATGACGTTAGTAATATCAATGACTTTTCAAAGCGTTATA
TTTGGCCAAATGCA
GAGTACTGGTTTGAACCGACAAAAATGGGCAATCTCTGTTTGATGGTGTGGTATGGGGCACGTAATTCTA

TTTTAATCTCAG

TTATAGCGACACTAATTAATATCACCATTGGGGTAGTGTTAGGAGCCATATGGGGAGTTTCTAAAGCATTGATAAAGTTATGAT

TGAAATTTATAACATTATCTCAAATATCCCTTCTATGCTTATTATCATTGTTTTGACCTATTCATTAGGTGCA GGATTTTGGAAATTT

GATTCTAGCTTTCTGTATCACTGGATGGATTGGTGTGCGCTACTCCATCCGTGTTCAAATCTTGC GTTACCGT GATTTAGAATAC

AACCTTGCTAGTCAAACCTTTGGGAACACCAATGTACAAGATTGCTGTTAAGAACCTCCTGCCTCAATTGGTTT CAGTTATCATGA

CTATGTTGTCACAAATGCTACCAGTTTATGTATCTTCTGAGGCCTTCTTATCCTTCTTTGGGATTGGTTTACC AACCACCACTCCA

AGTTTAGGACGTTTTATTGCTAATTATTCAAGCAACTTAACAACAAATGCCTACCTCTTTTGGATTCCCTTAG TAACATTGATTTTA

GTATCGTTACCACTATACATTGTCGGACAAAACCTTGGCTGATGCCAGTGACCCACGTTACATAGATAG

SPy0348

Seq ID 18

TTGGCTTTGACCGATTTTAAGGATAAAGACCAACAAGATCAGCAGCGAAGCTTTAAAGAGCAGATTCTTGCTG AGTTGGAAAAA

GCAAATCAAATCAGAAAAGAAAAAGAGGAAGAACTTTTTCAAAAAGAGTTGGAAGCTAAAGAAGCAGCTAGGA GAACGGCCCAG

CTATATGCTGAATATAAAAGACAAGATGCTTTTCAAAAAGAGTCTATAGCACATAACAATAAGACAGCTAAAC ACTTTCAAGCTAT

AAAAGGTGCGGTAATGACTTCAGAAGCGCTTAAACCGACTTTACTTTCTGAAAAAGAAAACCTCATCTTTAAAA ACGACAAATAAG

AGAGTCGTGCAGGCAAATGAGCTTCAAGAGACTGCCTCTAAAGAATCTCAAGTACCGTTAACTATTGAGAAAG GTCATTCAGTG

AGACGAAAATTAAGCAAACGCCAACAGACTGAGCGAGCTGCTAAAAAGATTTCAACCGTTTTGATTAGTTCTA TTATTATAACCC

TTTTGGCTGTTACTCTAGCAGGAGCAGGCTATGTTTATAGTGCTTTAAATCCTGTTGATAAAAAATAGTGATGC CTTTGTTCAGTT

GAGATTCCATCTGGGTCAGGCAATAAATTGATTGGTCAAATTCCTCAAAAAAAGGTTTAATCAAGAATAGCA CTGTTTTTTAGTTT

TTATACAAAATTTAAAACTTTACAAATTTTCAGAGCGGGTATTATAATCTGCAAAAAAGTATGAGTCTAGAA GAAATTGCTAGTG

CTTTACAAGAAGGTGGTACAGCAGAACCTACCAAGCCATCTCTGGGAAGATCTTGATTCCAGAAGGATACAC GATTAAACAAA

TAGCTAAAGCTGTTGAGCATAATAGCAAGGGAAAGACCAAAAAAGCTAAAACACCTTTTAACGAGAAGGATTT TTTGGATTTAGT

CACGGATGAGGCTTTTATTCAAGATATGGTAAAAAGATATCCAAAATTATTAGCAACTATCCCAACTAAAGAA
AAAGCTATTTACC

GTTTGGAAGGTTACCTTTTCCCAGCAACCTATAACTATTACAAAGAACTACCATGAGAGAACTTGTAGAGGA
CATGCTGGCAG

CTATGGATGCTACTTTGGTACCCTATTATGATAAAATTGCTGCTAGTGGTAAGACAGTCAACGAGGTATTGAC
CTTGGCCTCTTT

GGTTGAAAAAGAAGGTTCAACAGACGATGACAGACGTCAAATTGCAAGTGTCTTTTATAACCGCCTTAATAGC
GGAATGGCACT

ACAATCTAATATAGCTATTTTGTATGCGATGGGGAAACTTGGTGAGAAAACAACCTTGGCTGAGGATGCTACT
ATTGACACCACC

ATTAATTCTCCTTATAATATTTATACCAATACAGGTCTGATGCCTGGTCCAGTTGCTAGCTCGGGGGTTTCTG
CAATTGAAGCAA

CCCTAAATCCAGCCTCAACGGATTATTTATACTTTGTGGCCAATGTCCATACTGGTGAAGTTTACTATGCAAA
AACATTTGAAGAA

CACTCTGCAAATGTTGAAAAATATGTGAATAGTCAAATTCAGTAA

SPy0416

Seq ID 19

GTGGAGAAAAAGCAACGTTTTTCCCTTAGAAAAATACAAATCAGGAACGTTTTTCGGTCTTAATAGGAAGCGTTT
TCTTGGTGATGA

CAACAACAGTAGCAGCAGATGAGCTAAGCACAAATGAGCGAACCAACAATCACGAATCACGCTCAACAACAAGC
GCAACATCTCA

CCAATACAGAGTTGAGCTCAGCTGAATCAAAATCTCAAGACACATCACAAATCACTCTCAAGACAAATCGTGA
AAAAGAGCAATC

ACAAGATCTAGTCTCTGAGCCAACCACAACCTGAGCTAGCTGACACAGATGCAGCATCAATGGCTAATACAGGT
TCTGATGCGAC

TCAAAAAAGCGCTTCTTTACCGCCAGTCAATACAGATGTTACGATTGGGTAAAAACCAAAGGAGCTTGGGAC
AAGGGATACAA

AGGACAAGGCAAGGTTGTCGCAGTTATTGACACAGGGATCGATCCGGCCCATCAAAGCATGCGCATCAGTGAT
GTATCAACTG

CTAAAGTAAATCAAAAAGAAGACATGCTAGCACGCCAAAAAGCCGCCGTATTAATTATGGGAGTTGGATAAA
TGATAAAGTTGT

TTTTGCACATAATTATGTGAAAAATAGCGATAATATCAAAGAAAAATCAATTCGAGGATTTTGATGAGGACTGG
GAAAACTTTGAGT

TTGATGCAGAGGCAGAGCCAAAAGCCATCAAAAAACACAAGATCTATCGTCCCCAATCAACCCAGGCACCGAA
AGAAACTGTTA

TCAAAACAGAAGAAACAGATGGTTCACATGATATTGACTGGACACAAACAGACGATGACACCAAATACGAGTC
ACACGGTATGC

ATGTGACAGGTATTGTAGCCGGTAATAGCAAAGAAGCCGCTGCTACTGGAGAACGCTTTTTAGGAATTGCACC

AGAGGCCCAA

GTCATGTTTCATGCGTGTTTTTGCCAACGACATCATGGGATCAGCTGAATCACTCTTTATCAAAGCTATCGAAG
ATGCCGTGGCTT

TAGGAGCAGATGTGATCAACCTGAGTCTTGGAACCGCTAATGGGGCACAGCTTAGTGGCAGCAAGCCTCTAAT
GGAAGCAATT

GAAAAAGCTAAAAAAGCCGGTGTATCAGTTGTTGTAGCAGCAGGAAATGAGCGCGTCTATGGATCTGACCATG
ATGATCCATTG

GCGACAAATCCAGACTATGGTTTGGTCGGTCTCCCTCAACAGGTCGAACACCAACATCAGTGGCAGCTATAA
ACAGTAAGTGG

GTGATTCAACGTCTAATGACGGTCAAAGAATTAGAAAACCGTGCCGATTTAAACCATGGTAAAGCCATCTATT
CAGAGTCTGTGCG

ACTTTAAAGACATAAAAGATAGCCTAGGTTATGATAAATCGCATCAATTTGCTTATGTCAAAGAGTCAACTGA
TGCGGGTTATAAC

GCACAAGACGTTAAAGGTAAAATTGCTTTAATTGAACGTGATCCCAATAAAACCTATGACGAAATGATTGCTT
TGGCTAAGAAAC

ATGGAGCTCTGGGAGTACTTATTTTTTAATAACAAGCCTGGTCAATCAAACCGCTCAATGCGTCTAACAGCTAA
TGGGATGGGGA

TACCATCTGCTTTCATATCGCACGAATTTGGTAAGGCCATGTCCCAATTAAATGGCAATGGTACAGGAAGTTT
AGAGTTTGACAG

TGTGGTCTCAAAAGCACCGAGTCAAAAAGGCAATGAAATGAATCATTTTTTCAAATTGGGGCCTAACTTCTGAT
GGCTATTTAAAA

CCTGACATTACTGCACCAGGTGGCGATATCTATTCTCCTAATAACGATAACCACTATGGTAGCCAAACAGGAA
CAAGTATGGCC

TCTCCTCAGATTGCTGGCGCCAGCCTTTTGGTCAAACAATACCTAGAAAAGACTCAGCCAAACTTGCCAAAAG
AAAAAATTGCT

GATATCGTTAAGAACCTATTGATGAGCAATGCTCAAAATTCATGTTAATCCAGAGACAAAAACGACCACCTCAC
CGCGTCAGCAA

GGGGCAGGATTACTTAATATTGACGGAGCTGTCACTAGCGGCCTTTATGTGACAGGAAAAGACAACTATGGCA
GTATATCATTA

GGCAACATCACAGATACGATGACGTTTGATGTGACTGTTACAAACCTAAGCAATAAAGACAAAACATTACGTT
ATGACACAGAAT

TGCTAACAGATCATGTAGACCCACAAAAGGGCCGCTTCACTTTGACTTCTCACTCCTTAAAAACGTACCAAGG
AGGAGAAGTTA

CAGTCCCAGCCAATGGAAAAGTGACTGTAAGGGTTACCATGGATGTCTCACAGTTCACAAAAGAGCTAACAAA
ACAGATGCCAA

ATGGTTACTATCTAGAAGGTTTTGTCCGCTTTAGAGATAGTCAAGATGACCAACTAAATAGAGTAAACATTCC
TTTTGTTGGTTTT

AAAGGGCAATTTGAAAACCTTAGCAGTTGCAGAAGAGTCCATTTACAGATTAAAATCTCAAGGCAAAACTGGTT
TTTACTTTGATG

AATCAGGTCCAAAAGACGATATCTATGTCGGTAAACACTTTACAGGACTTGTCACCTCTGGTTCAGAGACCAA
TGTGTCAACCAA
AACGATTTCTGACAATGGTCTACACACACTTGGCACCTTTAAAAATGCAGATGGCAAATTTATCTTAGAAAAA
AATGCCCAAGGA
AACCCTGTCTTAGCCATTTCTCCAAATGGTGACAACAACCAAGATTTTGCAGCCTTCAAAGGTGTTTTCTTGA
GAAAATATCAAG
GCTTAAAAGCAAGTGTCTACCATGCTAGTGACAAGGAACACAAAAATCCACTGTGGGTGAGCCAGAAAGCTT
TAAAGGAGATA
AAAACCTTTAATAGTGACATTAGATTTGCAAAATCAACGACCCTGTTAGGCACAGCATTTTCTGAAAAATCGTT
AACAGGAGCTGA
ATTACCAGATGGGCATTATCATTATGTGGTGTCTTATTACCCAGATGTGGTGGTGCCTGCAACGTCAAGAAATG
ACATTTGACATG
ATTTTAGACCGACAAAAACCGGTACTATCACAAGCAACATTTGATCCTGAAACAAACCGATTCAAACCGAAGC
CCCTAAAAGACC
GTGGATTAGCTGGTGTTCGCAAGACAGTGTCTTTATCTAGAAAGAAAAGACAACAAGCCTTATACAGTTAC
GATAAACGATAG
CTACAAATATGTCTCAGTAGAAGACAATAAAACATTTGTGGAGCGACAAGCTGATGGCAGCTTTATCTTGCCG
CTTGATAAAGCA
AAATTAGGGGATTTCTATTACATGGTCGAGGATTTTGCAGGGAACGTGGCCATCGCTAAGTTAGGAGATCACT
TACCACAAACA
TTAGGTAAAACACCAATTAAACTTAAGCTTACAGACGGTAATTATCAGACCAAAGAAACGCTTAAAGATAATC
TTGAAATGACACA
GTCTGACACAGGTCTAGTCACAAATCAAGCCCAGCTAGCAGTGGTGCACCGCAATCAGCCGCAAAGCCAGCTA
ACAAAGATGA
ATCAGGATTTCTTTATCTCACAAACGAAGATGGGAATAAAGACTTTGTGGCCTTTAAAGGCTTGAAAAATAA
CGTGTATAATGA
CTTAACGGTTAACGTATACGCTAAAGATGACCACCAAAAAACAAACCCCTATCTGGTCTAGTCAAGCAGGCGCT
AGTGTATCCGC
TATTGAAAGTACAGCCTGGTATGGCATAACAGCCCGAGGAAGCAAGGTGATGCCAGGTGATTATCAGTATGTT
GTGACCTATCG
TGACGAACATGGTAAAGAACATCAAAAGCAGTACACCATATCTGTGAATGACAAAAACCAATGATCACTCAG
GGACGTTTTGAT
ACCATTAATGGCGTTGACCACTTTACTCCTGACAAGACAAAAGCCCTTGACTCATCAGGCATTGTCCGCGAAG
AAGTCTTTTACT
TGGCCAAGAAAAATGGCCGTAAATTTGATGTGACAGAAGGTAAAGATGGTATCACAGTTAGTGACAATAAGGT
GTATATCCCTAA
AAATCCAGATGGTTCTTACACCATTTCAAAAAGAGATGGTGTCACTGTCAGATTATTACTACCTTGTCGAA
GATAGAGCTGGT
AATGTGTCTTTTGCTACCTTGCGTGACCTAAAAGCGGTGCGAAAAGACAAAGCAGTAGTCAATTTTGGATTAG

ACTTACCGGTC

CCTGAAGACAAACAAATAGTGAACCTTTACCTACCTTGTGCGGGATGCAGATGGTAAACCGATTGAAAACCTAG
AGTATTATAATA

ACTCAGGTAACAGTCTTATCTTGCCATACGGCAAATACACGGTCGAATTGTTGACCTATGACACCAATGCAGC
CAAACCTAGAGT

CAGATAAAATCGTTTCCTTTACCTTGTGAGCTGATAACAACCTCCAACAAGTTACCTTTAAGATAACGATGTT
AGCAACTTCTCAA

ATAACTGCCCACTTTGATCATCTTTTGCCAGAAGGCAGTCGCGTTAGCCTTAAAACAGCTCAAGATCAGCTAA
TCCCGCTTGAAC

AGTCCTTGTATGTGCCTAAAGCTTATGCGAAAAACCGTTCAAGAAGGCACTTACGAAGTTGTTGTCAGCCTGCC
TAAAGGCTACC

GTATCGAAGGCAACACAAAGGTGAATACCCTACCAAATGAAGTGCACGAACTATCATTACGCCTTGTCAAAGT
AGGAGATGCCT

CAGATTCAACTGGTGATCATAAGGTTATGTCAAAAAATAATTCACAGGCTTTGACAGCCTCTGCCACACCAAC
CAAGTCAACGAC

CTCAGCAACAGCAAAAGCCCTACCATCAACGGGTGAAAAAATGGGTCTCAAGTTGCGCATAGTAGGTCTTGTG
TTACTCGGACT

TACTTGCGTCTTTAGCCGAAAAAATCAACCAAAGATTGA

SPy0430

Seq ID 20

ATGAAATGGAGTGGTTTTATGAAAAACAAATCAAAACGCTTTTTAAACCTAGCAACCCCTTGCTTGGCCCTAC
TAGGAACAACCTT

GCTAATGGCACATCCCGTACAGGCGGAGGTGATATCAAAAAGAGACTATATGACTCGCTTCGGGTTAGGCGAT
TTAGAAGATGA

TTCAGCTAACTATCCTTCAAATTTAGAAGCTAGATATAAAGGATATTTAGAGGGATATGAAAAAGGCTTAAAA
GGAGATGATATAC

CCGAACGGCCCAAGATTCAGGTTCTGAGGATGTTTCAAGCATCTGACCATGGCGACTATAGAGATGGTTATGA
GGAAGGATTT

GGAGAAGGACAACATAAACGTGATCCATTAGAAACAGAAGCAGAAGATGATTCTCAAGGAGGACGTCAAGAAG
GACGTCAAGG

ACATCAAGAAGGAGCAGATTCTAGTGATTTGAACGTTGAAGAAAGCGACGGTTTGTCTGTTATTGATGAAGTA
GTTGGAGTAATT

TATCAAGCATTTAGTACTATTTGGACATACTTAAGCGGTTTGTCTAA

SPy0433

Seq ID 21

ATGAAAAAGACATTAACCTTTGCTACTGGCACTCTTGCCATCGGGGTAACCTAGTAGCGTCAGAGCGGAGGATG
AACAAAGTAGT

ACACAAAAGCCAGTAAAATTTGATTTGGATGGACCTCAACAAAAAATTAAAGATTATAGTGGCAACACAATCA
CTCTAGAAGACT

TATATGTTGGTAGTAAAGTAGTAAAAATATATATCCCTCAAGGATGGTGGGTATATCTTTACAGACAATGTGA
TCATAACAGTAA

GAACGAGGAATTTTAGCTAGTCCTATTCTCGAAAAAATATAACAAAAACAGATCCTTATCGTCAATATTATA
CAGGAGTACCTTA

TATTCTTAACTTAGGAGAAGATCCTTTGAAGAAAGGAGAAAAATTAACCTTCTCATTAAAGGAGAAGACGGA
TTTTATGTGCGTA

GCTATATCTATAGAGACTCTGATACTATAAAAAAGAAAAAGAAGCTGAAGAAGCACTTCAAAAAAGGAAGA
GGAAAAGCAACA

AAAACAGCTAGAAGAAAGCATGCTAAAGCAGATAAGAGAAGAAGACCATAAACCTTGGCATCAGCGTTAAGT
GAGAGCATCCA

AGATCAGTGGTGGAACTTTAAGGGACTGTTTCAGTGA

SPy0437

Seq ID 22

ATGAAAAAACATTAACCTTGCTACTGGCACTCTTTGCCATCGGGGTAAGTAGTAGCGTCAGAGCGGAGGATG
AACAGAATAAG

TTTATACTTGATGGATTACAGGAAAAAGTAAAAAGAAGTTAGTGTATCAGATTTTTCTGTTGGAGAATCTAAAA
TCAAAGTCTGGCT

TCCTCAAGCTTGGTCGGTCAAAATTTCTAGAGAACATTCACCAAAATCAAGCATTTCTAATTCTGGAGAACAA
AAACCTTTAAGCA

ATAGCTCAGAGAATAAAGAAGGTCAATTTCTAAAAGATTACCTTATGGTACCCAACATACTATTAAATTATC
ATCCCAACTTACA

AAAGGTGAGAGAGTCACTTTGACATTCAGAGATGAAGATTTTGGGGAGCAGGTTACTGCTTCTATAGAGATT
CACTATCCATAA

AAGAAGACAAACAATACGAAGAAGAAATTAAGAAAAATTGAGGATGACCTAGAGAGACAAGATCTTGAAAATGA
TGCACTAGAGA

TGTTTTAAAAACAAACCGAAAGAGAGGCTAATAAACCTTGGCATCAGCGGTTAAGCGAAAACATCCAAGATCA
GTGGTGGAACT

TTAAGGGACTGTTTCAGTGA

SPy0469

Seq ID 23

ATGATTATTACTAAAAAGAGCTTATTTGTGACAAGTGTGCTTTGTGCTTAGTACCTTTGGCGACAGCGCAGG
CACAAGAGTGG

ACACCACGATCGGTTACAGAAATCAAGTCTGAACTCGTCCTAGTTGATAATGTTTTACTTATACTGTAAAT
ACGGTGACACTTT

AAGCACAATTGCTGAAGCAATGGGGATTGATGTGCATGTCTTAGGAGATATTAATCATATTGCTAATATTGAC

CTAATTTTCCAG

ACACGATCCTAACAGCAAACTACAATCAACACGGTCAGGCAACGAATTTGACGGTTCAAGCACCTGCTTCTAG
TCCAGCTAGCG

TTAGTCATGTACCTAGCAGTGAGCCATTACCCCAAGCATCTGCCACCTCTCAACCGACTGTTCCATGGCACC
ACCTGCGACAC

CATCTGATGTCCCAACGACACCATTCGCATCTGCAAAGCCAGATAGTTCTGTGACAGCGTCATCTGAGCTCAC
ATCGTCAACGA

ATGATGTTTCGACTGAGTTGTCTAGCGAATCACAAAAGCAGCCAGAAGTACCACAAGAAGCAGTTCCAACCTCC
TAAAGCAGCTG

AAACGACTGAAGTCGAACCTAAGACAGACATCTCAGAAGCCCCAACTTCAGCTAATAGGCCTGTACCTAACGA
GAGTGCTTCAG

AAGAAGTTTCTTCTGCGGCCCCAGCACAAAGCCCCAGCAGAAAAAGAAGAAACCTCTGCGCCAGCAGCACAAAA
AGCTGTAGCT

GACACCACAAGTGTGCAACCTCAAATGGCCTTTCTTACGCTCCAAACCATGCCTACAATCCAATGAATGCAG
GGCTTCAACCA

CAAACAGCAGCCTTCAAAGAAGAAGTGGCTTCTGCCTTTGGTATTACGTCATTTAGTGGTTACCGTCCAGGTG
ATCCAGGAGAT

CATGGTAAAGGTTTGGCCATTGATTTTATGGTGCCTGAAAAATCTGCTCTTGGTGATCAAGTTGCTCAATATG
CCATTGACCATA

TGGCAGAGCGTGGTATTTCATACGTTATTTGGAAACAGCGATTCTATGCGCCATTTGCAAGTATTTACGGACC
AGCCTACACAT

GGAACCCCATGCCAGATCGCGGCAGTATTACAGAAAACCATTATGATCATGTTTCATGTCTCCTTTAATGCTTA
A

SPy0488

Seq ID 24

TTGCGGCAGATTTCAGTCCATTCGTCTGATAGACGTTTTGGAGTTGGCTTTTGGAGTTGGCTATAAGGAAGAAA
CAACCTCTCAG

TTTTCTTCGGATCAGCCCTCCCAAGTGGTTTTGTATCGAGGTGAGGCTAACACGGTTAGGTTTGCCTATACCA
ATCAGATGTCTC

TGATGAAAGATATTTCGATTGCTTTGGATGGTTCTGATAAGTCTTTGACCGCTCAGATTGTTCCCTGGTATGGG
TCATGTTTATGA

GGGCTTTCAAACCTTCTGCTAGAGGGATTTTACGATGTCAGGAGTTCCTGAAAGCACTGTTCCCGTTGCTAAC
CCTAATGTACAA

ACCAAATATATAAGGTATTTCAAAGTCATTGATGATATGCATAACACAATGTATAAAGGAACTGTTTTTCTTG
TTCAACCGCAAGC

TTGGAAATACACCATGAAATCTGTTGATCAGTTACCAGTAGATGACTTGAACCATATTGGCGTTGCTGGTATT
GAACGAATGACA

ACTCTCATTAATAAATGCGGGTGCCCTTTTAACCACAGGAGGTAGTGGGGCTTTCCCAGACAATATTAAAGTAT

CTATTAATCCAA

AGGGGAGGCAGGCCACGATTACTTATGGGGACGGCTCTACGGATATTATTCCTCCAGCAGTTTTATGGAAAAA
AGGCTCCGTA

AAAGAGCCTACTGAAGCCGATCAATCTGTCGGAACACCGACTCCTGGTATTCCTGGTAAATTCAAACGAGACC
AGAGCCTTAAC

GAGCATGAAGCTATGGTAAATGTCGAACCACTGTCTCATGTAGTAAAAGACAATATAAAGGTCATAGATGAAA
AATCAACAGGG

CGGTTTGAGCCTTTTAGACCTAATGAAGATGAGAAGGAGAAGCCTGCCAGCGATGTTAAGGTAAGACCAGCAG
AAGTTGGTAG

CTGGCTAGAACCAGCGACAGCTCTTCCTAGTGTGAAATGAGCGCTGAGGACAGGTTAAAAAGTTAG

SPy0515

Seq ID 25

ATGAAAGTCTTATTGTATTTAGAAGCAGAAAAATTATCTAAGAAAATCAGGAATTGGTCGAGCGATTAAGCATC
AGGCTAAAGCCT

TGTCACCTGTTGGTCAACATTTTACGACTAATCCAAGAGAACTTATGATTTGGTTCATCTCAATACCTATGG
TTTAAAAAGTTGG

CTGCTGATGATAAAAGCACAAAAAGCTGGTAAGAAGGTTATCATGCATGGGCATTCTACAGAAGAAGATTTTA
GAAATTCTTTTA

TTTTTTCAAATCTATTATCTCCTTGGTTTAAAAAATACCTTTGTCACTTTTACAATAAGGCAGATGCTATCAT
TACCCCTACCCTAT

ATTCTAAGTCTTTGATTGAGAGTTATGGAGTGAAGTCACCTATTTTTTGCAGTGTCAAATGGGATTGATTTGGA
GCAGTACGGAGC

AGATCCTAAAAAGGAAGCAGCTTTTCGTCGCTACTTTGACATTAAAGAGGGTGAAAAAGTGGTTATGGGAGCA
GGATTATTTTTT

CTGAGGAAAGGAATTGATGACTTTGTCAAAGTTGCCCAAGCTATGCCAGATGTTTCGTTTTATCTGGTTTGGCG
AGACCAACAAAT

GGGTCAATTCCTGCTCAAGTTCGCCAAATGGTCAATGGTAACCAACCCGAAAAATCTTATTTTCCCAGGATACAT
TAAAGGGGATGT

TTATGAAGGTGCCATGACTGGTGCAGATGCCTTTTTCTTTCCAAGTCGTGAAGAAACAGAAGGCATTGTTGTC
TTAGAAGCCTT G

GCCAGTCGCCAGCACCTTGTTTTACGTGATATACCAGTTTACTACGGATGGGTGATCAAAGTAGTGCGGAAT
TAGCAACCGAT

ATACCAGGTTTTATAGAAGCTCTGAAAAAAGTCTTTTCTGGTGCCAGCAACAAAGTTGAAGCTGGTTACAAGG
TTGCCCAGAGT

CGTCGCCTAGAAACGGTTGGCCATGCCTTAGTAGATGTCTATAAAAAAGTAATGGAGTTATAA

SPy0580

Seq ID 26

ATGGAATAACAATAATCACAACATTGCAGAGGCTTTGTCAGTTAGTCTCCATCAAATTGAACAGGTGCTTG
CCCTAACAGCAC
AAGGCAATACCATTCCCTTTATTGCACGTTACCGTAAAGAAGTAACAGGAAATCTAGATGAAGTGGTGATTAA
GTCCATTATTGA
TATGGATAAGTCTCTCACTACCCTGAATGAGCGCAAAGCAACCATTTCTCGCTAAAATTGAGGAACAGGGAAAA
CTGACAGATCA
GTTACGAACGAGTATTGAAGCAACCGAGAAACTCGCTGATTTGGAAGAGTTGTACCTGCCTTATAAGGAAAA
CGCCGTACAAA
AGCGACGATTGCGCGTGAAGCAGGCTTGTCCCATTAGCACGTTGATTTACAAAATGCGCAAAACCTTGAA
ACAGCAGCAGA
ACCCTTTGTCACAGAAGGATTTGCCAGCCACAAGAAGCTCTAGCAGGAGCTGTGGACATCCTTGTGGAGGCC
ATGTCAGAAG
ATGCTAAATTACGCTCTTGGACCTACAATGAGATTTGGCAGTACAGCCGCTTAGTATCAACGCTTAAAGATGA
GCAGTTGGATG
AGAAAAAGGTTTTCCAAATCTATTACGATTTTTCTGACCAAGTGTCTAACATGCAAGGATATCGTACTTTAGC
CCTTAACCGTGG
CGAAAAGTTAGGCATCTTAAAAGTGTCTTTTGAGCATAACTTGGAGAAAATGCAACGCTTTTTTCAGTGTGCGG
TTCAAAGAAACC
AACCCTTATATTGAAGAGGTCATCAATCAGACCATCAAAAAGAAAATTGTTCCAGCTATGGAAAGACGGGTTC
GTTCAGAACTCA
GTGATGCCGAGAAGATGGGGCTATCCACCTCTTTTCTGAAAATCTCCGTCATCTTTTGTAGTGTCTCCTTT
AAAAGGGAAAAT
GGTTCTTGGGTTTGACCCTGCCTTTGAAACAGGTGCAAAATTGGCTATCGTTGATCAGACTGGAAAATTGTTG
ACTACTCAGGTT
ATTTATCCCGTAGCACCAGCTAGCCAAACGAAGATTCAAGCAGCTAAAGAAACACTGACTCAGCTCATTGAGA
CTTACCAGATT
GATATTATTGCCATTGGAAACGGAACAGCCAGTCGGGAAAGTGAAGCCTTTGTAGCAGACGTTTTGAAAGATT
TCCCAAAATACG
TCCTATGTCATCGTTAACGAAAGCGGTGCTTCTGTTTATTCAGCGTCAGAGTTAGCCCGCCATGAGTTTCCAG
ACCTCACAGTG
GAGAAGCGCTCTGCTATTTCTATTGCTCGTCGCTTGCAAGACCCACTCGCAGAATTGGTTAAAATTGATCCTA
AATCTATTGGCG
TCGGTCAGTACCAGCACGATGTGAGTCAGAAAAAGTTGAGTGAGAATCTTGGTTTTGTTGTGGATACTGTGGT
TAACCAAGTCG
GTGTTAATGTGAACACGGCTAGTCCATCTTTATTAGCACATGTGTCAGGTTTGAACAAAACGATTTTCAGAAAA
TATTGTCAAATAC
CGCGAAGAAAATGGAGCACTGACATCACGCGCTGACATCAAAAAAGTGCCGCGTTTAGGAGCCAAGGCGTTTG
AGCAAGCAGC
AGGTTTTCTCAGAATTCCAGGAGCCAAAAATATTTTAGACAACACAGGAGTTCACCCCGAGTCCTACCCAGCA

GTCAAAGAACT

CTTTAAGGTACTTGGTATTCAGGACTTGGACGACGCTGCCAAGGCAACTTTAGCAGCAGTTCAAGTTCCCCAA
ATGGCAGAAAC

ATTGGCTATTGGGCAGGAAACCCCTTAAAGATATCATTGCTGATCTCCTTAAACCGGGTCGTGACCTTCGTGAT
GACTTTGAAGC

ACCAATCTTACGTCAAGATATCCTTGATTTGAAAGATTTGGAAATTGGCCAGAAGCTTGAAGGAACTGTGAGA
AACGTCGTTGAC

TTTGGTGCTTTTGTAGACATTGGTGTTTCATGAAGATGGGCTTATTCACATTTCTGAAATGAGCAAAACCTTTG
TCAACCACCCTA

GTCAAGTTGTTTCAGTGGGTGACTTGGTAACTGTTTGGGTCTCTAAGATTGACTTGGACCGCCACAAGGTCAA
TCTCAGCCTAT

TGCCACCGCGTGACACTCACTAA

SPy0621

Seq ID 27

ATGAATGAAAAAGTATTTTCGTGACCCTGTTCACTATATCCACATTGATAATCCATTAATTTATGACCTTA
TCAATACCAAAGAA

TTCCAACGCTTACGACGCATCAAAACAAGTCCCTACAACAGCTTTCACCTTCCACGGCGCAGAACATAGCCGTT
TTTCTCACTGC

CTAGGTGTTTACGAAATTGCCCGTCGGGTGACAGCTATTTTGAAGAGAAATACGCTGATATCTGGAATAAAG
ATGAAAGTCTG

GTCACTATGACGGCTGCCCTCTTACATGACATTGGTCATGGAGCCTATTCTCATACCTTTGAGGTTCTTTTTC
ATACGGATCACG

AGGCCTTTACTCAGGAGATAATTACCAATCCTGAGACAGAAATCAATGCCATTTTGGTGCGTCATGCTCCTGA
TTTTCCAGACAA

AGTTGCCAGCGTGATTAATCATACTTATCCTAACAAGCAAGTTGTGCAACTGATTTCCAGCCAGATTGACTGC
GACCGTATGGAT

TACCTTTTAAGAGATTCTTATTTTTCTGCTGCCAATTACGGTCAGTTTGATTTAATGCGCATCCTACGTGTGA
TTCGACCTGTTGA

AGATGGCATTGTCTTTGAACATAGTGGTATGCATGCTGTGAGGACTATATTGTCAGTCGTTTTCAAATGTAC
ATGCAGGTCTAT

TTTCATCCCGCCAGCCGTGCTGTTGAGCTCATCTTACAAAATCTCCTCAAAAGAGCCCAACACTTATATCCTG
AGCAACAGGCTT

ATTTCCAAAAAACAGCTCCTGGACTCATCCCCTTCTTTGAAAAGAAGGCTAATTTAGCAGATTATATAGCCCT
AGATGACGGAGT

CATGAATACTTACTTCCAAGTCTGGATGGCAAGCGAAGACCATATCTTATCAGATTTGGCCAGTCGCTTTATT
AATCGAAAAATT

CTGAAGTCTGTTACCTTTGACCAAGATTCCCAAGGAGAGTTAGAACGCTTGCGCCAATTGGTCGAATCAGTTG
GCTTTGATCCA

GATTATTACACCGGTATCCATATCAACTTCGACCTGCCTTATGATATTTATCGACCAGAGCTCGAAAATCCAA
GAACTCAGATTG

AAATGATGCAAAAAGACGGGAGCCTTGCTGAATTATCTCAGTTATCTCCTATTGTCAAAGCATTAACAGGGAC
AACCTACGGAG

ACCGTCGCTTTTATTTTCCAAAGGAAATGTTAGAGCTCGATGACCTCTTTGCTCCTAGCAAGGAACTTTTAT
GAGTTATATTTCC

AATGGGCATTTTCACTTTAGTCAGTAA

SPy0630

Seq ID 28

ATGGATATCAATTTGTTACAGGCACTTTTAATTGGTCTTTGGACAGCCTTTTGTTCAGCGGAATGTTACTTG
GCATCTACACCAA

CCGTTGTATTATTCTGTCTTTTGGTGTAGGGATTATCTTAGGTGACTTGCCAACTGCACTTAGCATGGGAGCT
ATCTCCGAATTA

GCTTATATGGGATTTGGAGTGGGTGCCGGTGGTACCGTTCCTCCTAATCCTATCGGCCCTGGTATCTTTGGTA
CCTTGATGGCT

ATTACCAGTGCTGGTAAAGTCACCCAGAAGCGGCACTAGCCTTATCAACACCAATCGCAGTTGCCATCCAAT
TCCTTCAAACA

TTGCCTATACAGCTTTTGTGCTGGCGCTCCCGAAACCGCTAAAAACAATTGCAAAAAGGCAATATTAGAGGCT
TCAAATTCGCTG

CCAACGGCACTATCTGGGCTTTGCTTTTATCGGATTAGGCCTTGGTTTATTAGGTGCCTTGTCAATGGATAC
TCTCCTTCACTT

GGTTGATTATATCCACCTGTCTTGCTCAATGGATTGACTGTCGCTGGTAAAATGTTACCTGCTATCGGATTT
GCCATGATCTTA

TCTGTTATGGCCAAGAAAGAATTGATTCCTTTTGTACTAATTGGTTATGTTTGTGCAGCCTACCTCCAAATTC
CAACCATTGGTAT

CGCCATTATTGGTATCATTTTCGCCTTGAATGAATTTTACAACAAACCTAAACAAGTCGATGCAACAACCTGTC
CAAGGAGGCCAA

CAAGATGACTGGATCTAA

SPy0681

Seq ID 29

TTGACCCCTCGTAGCGGAAAGACCACAGCTGGGCATTTTCGTTATGCTAGGTATCTGATTGAGTCAGAAGATG
AAAATCACCTT

GTGACTGCTTATAATCAAGAACAAGCTTATCGTTTGTATTCGACGGCGATGGTACGGGTTTGATGCATATAT
TTGACGGTAACT

GTGAAATAAAACACGACGAGCGTGGAGATCACTTGTTAATCACGACACCAAAAGGCAATAAGCGCGTTTATTA
TAAAGGCGGCG

GTAAAGTTAACAGTGTTGGTGCTATTACAGGTATGTCTTTAGGATCAGTAGTATTCTGCGAGATTAACCTACT

GCACATGGATTTT

ATCCAGGAGTGTTTTAGGCGTACTTGGGCGGCTAAGCTACGTTATCATCTAGCAGATTTAAATCCCCCAGCAC
CTCAACATCCA

GTAATTAAAGATGTCTTTGATGTTTCTGAGAACACGAGGTGGACTCATTGGACCATGGATGATAACCCAATACTAA
CCGCAGAGCGT

AAACAAAACATTATCAACAGTCTTAAAAAAAATCCATATCTATACAAACGAGATGTACTTGGACAGCGGGTCA
TGCCTCAGGGAG

TTATTTATGGCCTTTTTGACACGGAAAAAAATGTTTTGGATGCTTTGATTGGCGAACCAGTAGAGATGTATTT
CTGTGCAGATGG

AGGTCAATCAGATGCCACCTCTATGTCTTGTAATATCGTAACAAGAGTTAGAGATAACGGTAGGATAAGCTTC
AGACTTAATCGT

GTAGCTCACTACTACCACAGCGGAGCTGACACTGGCCAAGTAAAAGCTATGTCAACCTACGCTTTAGAGTTAA
AAGTTTTTATAG

ACTGGTGC GTTAAAAAGTATCAGATGCGCTATACAGAGGTATTTGTGGATCCTGCCTGTAAATCTTTGAGAGA
GGAGCTGCATA

AGTTAGGAGTATTTACTCTGGGAGCTCCGAACAATTCTAAAGATGTATCTAGCAAAGCAAAAGGTATTGAGGT
CGGTATCGAAC

GCGGCCAAAACATTATCTCAGATGGCGCTTTTTATCTTGTTAATCATAGCGAAGAAGAGTATGACCACTACCA
CTTTTTAAAAGA

GATAGGGCTGTACAGTCGTGACGACAATGGCAAACCTATTGATAAAGATAACCATGCCATGGACGAGTTTAGA
TACAGCGTCAA

CGTGTTTGTGCATCGGTATTACAATAA

SPy0683

Seq ID 30

ATGAAAAAGAAGCCTATTAAGTTAAATGACGAACAGCTTCTTTTGGAAGCTAGTCAGTTATCTGATATGTATC
ATCAGCTGACTCT

TGATTTATTTGATCAAGTGATTGAGAGGATAAAAGCCAGAGGCTCAGCGAGCTTAGCCGATAATCCTTATCTT
TGGCAAGCTAAT

AAGTTACATGACGTTGGACTGCTTAATGCAGATAACATCAAGCTTATTGCAAAGTATTCTGGCATTGCGGAAG
ATCAACTTCGCT

ATATTATCAAGAATGAAGGATTTAAAAATTTATAAAAAACAGTCTGAGCAGCTAGAAGAGGCTCTAGGTAGAGA
GTCTGGGGTAAA

CAGTACTATCCAAGACGACCTATCTAACTATGCTAGACAAGCTATTGATGATGTGCATAATTTGACTAACACC
ACCTTGCCATTTA

GTGTTATAGGAGCTTATCAAGGGATAATCCAAGACGCTGTTGCTGGTGTGGTGACAGGCTTAAAAACGCCTGA
CCAAGCTATCA

ATCAAAGTGATTAATGGTTTTAAAAAGGGTTTTATGGTTTTACAGATAAAGCTGGGAGAAAGTGGAGAGC
AGACTCTTATGC

TCGTACCGTTATCAATACTACGACTTGGCGAGTCTTTAACGAAGCCAAAGAAGCCCCTGCTAGGGAGTTTGGC
ATTGATACCTT

CTATTACTCAAAAAAAGCTACAGCTAGAGAGATGTGTGCACCTTTGCAACATCAAATTGTCACTACTGGCGAA
GCGAGAGAAGA

AGGAGGGATAAAAAATCTTAGCTTTATCTGATTACGGGCATGGTGAGCCTGATGGATGCTTGGAATCAACTGC
AAGCACACTAA

AACGCCGTTTGTGTCGGTGTGAATAGTAAGCCAGAAATTGCCAGAGCATCTAAAAATATCACTCCTGCACAA
GCTAAAGCTAA

TGCGAATGCGCAAGCTAAGCAGAGGGCAATCGAGAGATCAATACGTAAGAGTAAAGAGCTACTGCACGTTGCG
AAGCAATTGG

GTGATAAAGAGTTGATTAGGCAATATCAATCGGATGTTAGAAGTAAACAAGATGCACTCAATTATCTGATAAA
CAACAATGCCTTT

TTACATCGCAATCAAGCCAGAGAAAAAGCGTTACAATAATCCTTATACCAAAACTCAAAGTGAAGTCGAAGTTA
GAAAAGAAAAAG

CTAAATTAGATAAACGTAGGGATGTTGAAAGTGCTATAATAGGAGTAGAACTAGTGAAGGGATACCGCTAAA
AATAACAAAGCA

TTTAGCCGAAAGGGCGGTGCTGAGAAATATAGCACCTATTGATATTGTCGATTCTATAAAAGAACCGTTGAAG
ATAGCTCCTATT

AAGTACGATAACCTTGATAGACCTTCCCAGAAAATACATTGGTAAGTGTGTCTCGACAGTAATAAACCCGATAG
ACGGAAATATTG

TTACAGTTCATGCTACTAGCACGAGAATCCGCAAAAAATATGGAGGAAATTGA

SPy0702

Seq ID 31

ATGAGCAGAGACCCAACACTTATTTTAGACGAGTCAAACCTCGTTATTGGTAAGGATGGACGTGTGCATTACA
CATTTACCACAG

AGGACGACAACCCAAAAGTCAGACTAGCTAGCAAGTGTCTAGGCACAGCGCATTTTAATCAGCTCATGATTGA
GCGAGGAGAC

CAAGCTACTAGCTATGTTGCGCCAGTAGTAGTTGAGGGTACAGGTAATCCGACTGGACTATTTAAAGACCTCA
AAGAGATTAGC

TTAGAGCTGACAGATACTGCTAATTCCCAGCTTTGGTCAAAAATCAAGCTGACTAACCGTGGTATGTTGCAGG
AATACTACGAC

GGTAAGATCAAGACCGAGATAGTCAACTCCGCCAGAGGTGTCGCTACACGTATCAGCGAGGATACTGATAAAA
AGCTAGCGCT

CATCAATGACACCATTGATGGTATCAGGCGTGAGTATCGAGATGCTGATAGGAAGCTATCCGCAAGCTATCAG
GCAGGCATCG

AGGGGCTAAAAGCCACAATGGCCAATGATAAAATCGGTTTACAAGCTGAGATTAAAGCCTCAGCACAAGGGCT
ATCGCAAAAGT

ATGATGATGAGTTGCGCAAGCTATCGGCTAAGATCACAACAACCTCAAGCGGCACTACAGAGGCCTACGAGAG

TAAGCTTGCG

GGCTTACGTGCTGAGTTTACTCGCTCAAAATCAAGGCACGAGGACAGAGCTCGAGTCACAAATTAGCGGGCTAA
GAGCGGTACA

GCAGTCAACAGCTAGCCAAATCTCTCAAGAGATTAGAGACCGTGAAGGTGCTGTCAGTCGTGTGCAGCAGAGT
TTGGAGAGTT

ACCAAAGGCGGATGCAGGACGCAGAAGAAAACTATAGTAGCTTGACCCATACGGTTAGAGGGCTACAGAGCGA
CGTTGGATCT

CCGACTGGTAAATCCAATCGCGCCTTACTCAACTAGCAGGACAAATTGAGCAGCGGGTTACTAGAGATGGTG
TCATGAGTATT

ATTAGTGCGCTGGAGACAGCATTAAATTAGCTATCCAAAAGGCTGGCGGCATTAATGCCAAAATGTCTGGTA
ATGAGATTATC

TCAGCAATTAACCTCAACTCCTACGGAGTAACAATCGCAGGTAAACACATCGCTCTCGATGGGAATACGACGG
TTAATGGCACC

TTTACCACAAAAATAGCCGAGGCTATCAAGATTAGGGCTGATCAGATTATTGCAGGCACGATTGACGCTGCTA
GGATTAGAGTG

ATTAACCTTAACGCAAGTAGTATCGTTGGTTTAGACGCTAACTTTATCAAAGCTAAAATTGGCTATGCTATCA
CTGATTTGCTCGA

GGGTAAGGTCATTAAGGCTCGTAATGGAGCGATGCTTATCGACTTAAATACAGCTAAGATGGACTTTAATAGC
GATGCCACAAT

TAATTTTAATAGCAAAAAACAATGCCTTAGTACGTAAAGATGGCACACATACTGCCTTTGTACATTTTAGTAAT
GCGACGCCCAA

GGTTATACAGGGTCAGCGTTGTATGCATCGATCGGGATAACCTCATCTGGTGACGGTGTTAACTCGGCTTCTT
CCGGTCGTTTT

GCAGGGCTAAGGTCATTTAGGTACGCTACGGGATATAATCACACTGCGGCAGTCGACCAGACTGAAATTTACG
GTGATAATGTT

TTAGTTGTGGATGATTTTAATATTACTCGGGGATTTAAGTTTAGACCAGACAAGATGCAAAAAATGCTTGACA
TGAACGACTTGTA

TGCGGCTGTAGTAGCCTTAGGCCGCTGTTGGGGCACTTGGCTAACGTCGGCTGGAATACTGCTCATAGCAAT
TTTACAAGTG

CTGTGAATAGGGAATTGAATAACTACATCACAAAAATTTAA

SPy0710

Seq ID 32

ATGACCTTTTTAGATAAAATTAACAAGGCTGTTTAGATGGCTGGGCTAAGTACAAAATCTTGCCATCCTTGA
CCGCAGCACAAG

CTATCTTAGAGAGCGGGTGGGGCAAAACATGCCCCACACAACGCTCTGTTTGGTATTAAGGCAGATAGCTCTTG
GACTGGTAAAT

CATTTGATACCAAAACCCAAGAGGAATATCAAGCAGGTGTTGTCACGGATATTGTGGACCGATTTAGGGCGTA
TGATAGTTGGG

ATGAGTCGATAGCTGATCACGGACAATTTTTAGTTGATAATCCACGCTATGAGGCAGTTATTGGGGAGACTGACTATAAAAAAGG

CTTGTTACGCTATTAAAGCAGCTGGATACGCTACGGCAAGTAGCTATGTCGAACTTTTAATCCAAGTATTGGAAAACGACTT

ACAAAGTTGGGATAGAGAAGCTCTTAAAAATAATAAGGAGGAAACGATGACAACCGCAAACGAAATTGTACAACTACTGTGTAAAC

CTTGCTAATTCAGGCATGGGTGTTGACAAAGACGGTGCTCACGGGACGCAATGCTGTGACTTGCCTTGTTTTGTCGCTAAAAAT

TGGTTTGGTGTTGATCTTTGGGGCAATGCGATTGATTTATTAGACAGCGCAAGTGCGCAAGGCTGGGAAGTCCATCGTATGCCA

ACAGAGGCCAAACCCAAAAGCAGGCGCTACATTTGTCCAATCAGTGCCGTATCATCAATTTGGACATACGGGAATTGTTATCGAG

GATAGTGACGGTTACACCATGCGCACTGTCGAGCAAAACATTGATGGCAATCCTGATGCTTTGTATGTCGGTGACCCAGCTCGT

TTTAACACTCGTGACTTTACTGGCGTGATAGGTTGGTTTACCCACCATATCAAGGGGATACAGTCACGCAAC CAGTCAGCACC

GAGCCGCAAACCTTCTGACACTATCGTAGAGACAGCAAAAACAGGCACCTTTACCCTTGATGTTGCAGAGATCAATATCAGACGC

TGGCCAAGTCTAGCCAGCGAGGTTGTAGGTATCTACAAGCAAGGTGATACTGTCAGCTTTGATAGCGAGGGCTACGCTAATGG

CTATTATTGGATTAGCTATGTTGGAGGCTCAGGTATGCGTAACTACCTAGGTATTGGACAGACTGATAAAGATGGGAATCGCAT

CAGCCTTTGGGGTAAATTAAATTAG

SPy0711

Seq ID 33

ATGAAAAAGATTAACATCATCAAAATAGTTTTTCATAATTACAGTCATACTGATTTCTACTATTTACCTATCA TCAAAAGTGACTCT

AAGAAAGACATTTTGAATGTTAAAAGTGATTTACTTTATGCATACACTATAACTCCTTATGATTATAAAAAATTGCAGGGTAAATTTT

TCAACGACACACACATTAAACATTGATACTCAAAAATATAGAGGGAAAGACTATTATATTAGTTCCGAAATGTCTTATGAGGCCTC

TCAAAAATTTAAACGAGATGATCATGTAGATGTTTTTGGATTATTTTATATTCTTAATTCTCACACCGGTGAGTACATCTATGGAG

GAATTACGCCTGCTCAAAAATAATAAAGTAAATCATAAATTATTGGGAAATCTATTTATTTTCGGGAGAATCTCAACAGAACTTAAAT

AACAAAATTATTCTAGAAAAAGATATCGTAACTTTCCAGGAAATTGACTTTAAAATCAGAAAATACCTTATGGATAATTATAAAATT

TATGACGCTACTTCTCCTTATGTAAGCGGCAGAATCGAAATTGGCACAAAAGATGGGAAACATGAGCAAATAG

ACTTATTTGACT

CACCAAATGAAGGGACTAGATCAGATATTTTTGCAAAATATAAAGATAATAGAATTATCAATATGAAGAACTT
TAGTCATTTTCGAT
ATTTATCTTGAAAAATAA

SPy0720

Seq ID 34

ATGATAACAACCTTTTGAAACAATTTTAGATAAAATAAAAGCTCACCAAACCTATTATTATCCATCGCCATCAAA
ATCCTGACCCTGAT
GCTCTTGGTAGTCAGGCCGGCTTGAAAGAAATTATTGCACAAAATTTCCCAGACAAAAAGGTTTTGATGACTG
GTTTTGATGAGC
CTAGTTTAGCTTGGATTAGCCAAATGGATCAGGTGACTGACAAAGACTATAAAGAGGCTTTGGTCATCATTAC
AGATACAGCGAA
TCGACCAAGGATTGATGATGAGCGCTACACACTGGGGAAGTGCTTAATTAAGATTGATCACCATCCCAACGAT
GATGTGTATGG
TGACTTCTATTATGTGGACACAAGCGCTTCTAGCGCAAGTGAAATCATTGCAGACTTTGCCTTTAGTCAGAAT
CTTACTCTCTCT
GACAAGGCTGCTAAGCTCTTATACACCGGTATCGTTGGTGATACAGGGCGATTTCTTTATGCCTCAACCACTA
GTAAAACCCCTT
CCATTGCTAGCCAACTCAGACATTTTCGAGTTCGACTTTGCTGCGATTTCAAGGCAAATGGATTGTTTTCTTT
GAAAATAGCAAA
GCTGCAAAGCTACGTCTTTGAGCATTTAACAATTGATGAGAGTGGGGCTGCTTATGTCCTTGTCAGCCAAGAA
ACCTTAAACAT
TTTGACGTGACCCTAGCAGAAAGCTCTGCCATTGTCTGTGCTCCTGGTAAAATTGATAACGTTCAAGCTTGGG
CTATTTTTGTTG
AGTTAACTGACGGCAACTACCGTGTGCGTATGCGCAGTAAAGAAAAGATTATTAATGGCATTGCTAAGCGTCA
CGGTGGAGGG
GGGCATCCCCTTGCTAGCGGAGCCAACTCAGCTAATTTAGAAGAAAATCAAGCTATTTTCCGAGAACTCATCG
CTGTTTGCCAA
GAGATTTAG

SPy0727

Seq ID 35

ATGATTGAAGAAAATAAACATTTTGAAAAAAAAATGCAAGAATACGATGCCAGTCAAATTCAGGTTCTAGAAG
GGCTGGAGGCTG
TCCGCATGCGTCCAGGGATGTATATTGGCTCGACAGCTAAAGAGGGTTTGCATCATTTAGTCTGGGAAATTGT
TGACAACTCAA
TTGACGAAGCATTAGCAGGTTTTGCCTCTCATATTAAGTCTTTATTGAAGCAGATAATTCCATTACAGTAGT
AGATGATGGCCG

TGGAATTCCAGTTGATATCCAAGCCAAGACAGGACGTCCCGCCGTTGAAACAGTTTTTACAGTCTTACACGCA
GGTGGTAAATT
TGGTGGAGGCGGCTATAAGGTTTCTGGAGGATTACATGGTGTAGGGTCATCTGTTGTTAATGCTTTATCAACA
CAATTAGATGTA
CGTGTTTATAAAAAACGGCCAAATTCATTACCAAGAATTTAAACGCGGGGCTGTTGTAGCAGATCTTGAGGTCA
TTGGAACCACT
GATGTGACTGGCACGACCGTACACTTTACACCCGATCCAGAAATTTTACCGAAACGACTCAGTTTGATTACA
GTGTTTTAGCAA
AACGTATTCAAGAGTTAGCCTTTTTGAATCGTGGTTTAAAAATTTCCATTACAGATAAGCGCTCAGGTATGGA
ACAAGAAGAACA
TTTCCTTTATGAAGGTGGAATTGGTTCTTATGTTGAATTTTTAAATGATAAAAAAGATGTTATCTTTGAAACG
CCCATCTATACAGA
TGGTGAATTAGAAGGTATTGCAGTTGAAGTAGCCATGCAATACACGACTAGCTATCAAGAAACAGTCATGAGT
TTTGCTAATAAT
ATTCATACTCATGAAGGTGGAACGCATGAACAAGGCTTTAGAGCGGCTCTTACTCGGGTCATCAATGACTACG
CTAAGAAAAAT
AAAATTCTTAAAGAAAATGAGGACAATTTGACAGGAGAAGATGTTTCGTGAAGGTTTGACGGCGGTAATTTCTG
TTAAGCATCCAA
ATCCTCAATTTGAAGGTCAAACCAAAACAAAATTTGGGCAACTCAGAAGTGGTTAAGATCACTAATCGTCTCTT
TAGTGAGGCCTT
TCAACGTTTTCTTTTGGAAAAACCCACAAGTTGCTCGTAAGATTGTGGAAAAAGGGATTTTGGCTTCTAAAGCT
AGAATTGCAGCT
AAGCGAGCCCGCGAAGTCACCCGCAAAAAATCAGGCTTAGAAATTTCAAACCTACCTGGAAAATTAGCAGACT
GTTTCGTCAAAT
GACGCTAACCAAAACGAACTTTTTCATCGTCGAAGGAGATTACAGCGGGTGGGTCGGCCAAATCAGGTCGTAACC
GAGAGTTTCA
AGCTATCTTGCCTATTTCGCGGTAAAAATTTGAACGTGGAAAAAGCAACTATGGATAAGATTCTTGCCAACGAA
GAAATTAGAAGT
CTCTTTACCGCTATGGGTACAGGTTTTGGTGCAGATTTTGACGTGTCAAAGCTCGCTACCAAAGCTGGTTA
TCATGACCGAT
GCCGATGTGGATGGCGCTCATATTAGAACCTTACTTTTAACCTTGATTTACCGCTTTATGAGACCTGTTCTAG
AAGCTGGCTATG
TTTACATCGCCAGCCACCTATTTATGGTGTTAAGGTCGGTAGTGAGATTAAAGAGTATATTCAGCCAGGTAT
TGATCAAGAAGA
CCAATTA AAAACAGCTCTTGAAAAATATAGTATTGGTCGTTCAAACCAACTGTTCAACGTTATAAAGGTCTT
GGGGAATGGAT
GACCATCAACTTTGGGAAACTACTATGGATCCTGAAAAATCGTTTGATGGCGCGTGTGACAGTTGATGATGCCG
CAGAAGCAGAT
AAAGTATTTGATATGTTAATGGGAGATCGTGTGAACCAAGACGTGATTTTCATTGAGGAAAATGCGGTTTATA

GTACACTGGATA

TTTAG

SPy0737

Seq ID 36

ATGCGTAAGGTCAAAAAAGTCTTTGTTAGTTCATGTATGCTTTTAACAGTGGGCCTCGGAGTTGCCGTACCTA
CTGGATTACGCC

AATCTAATGGCGTGATGGTTGTAAAGGCTGCGGAAGTGCCGGCGACAGATTTATCACGTCAGGCGTCTGATTC
GGAGAGGGTA

GATGAATCGTCTTTATTGCAGAAAGAAAACTTATCAGTAGATTCATTTAAATTAGAAAATTTAAATGGATGGG
AAGCTGAAAATGA

TACAGCAGGTAATTTGGGGAAATTTAAAGATCCAGATAGTTCGGGCTATCAAAATATTTTGACATCATCTGGA
AAGAATATCAGT

GTAGCTGTTGCTCCCAAAGGTTTCAGGTAAAAATGAACATTAAAGTAACTAAAAGATCAAATTTTCAGGGTGGAT
ATTATGTAGGTG

GTCTTAGAACTCAAACTCCGGTATTGAAGTTAAATGATGTTTATCGATATTCTTTTACAATAAAAAATTATC
AGGAAATTCTTCA G

AGTTCAAAACGAGAGTTAAGCCCGTTGAATCTAATAATAAACTAGGGAAAGAGCTTGTTATTAGGGTGGATAA
TAAAAATGTATC

TACTAAGCATGATTGGCTTCCAGACATCTCTGATGGAATCATACTGTGGACTTCACTGGTCTTGATAAAAAA
TTATCTGTTGCTT

TCAGATTTTCTCCAAGACAAACTTCGAATGTTGTTTACGAATTTTCTAACATAAAATATAAAAAACATTAGTCC
TGCATCAGTGCCG

GCTATTCCTTCGAAAGTTTTAGAGGGAACCAGCGTCTTGTCGGGTACTGCAATATCTTCTGGAGATACATTAG
AAAAAAGAAAAT

CGTTTGATGGCGATATCCTAAGAGTTTATAAAGATAGCAAAATCATTGCTAGAACAGTAATAAAAGGCAATAA
GTGGGATGTTAA

ACTTTCAAAGCCTCTTATTGCAGGTGAAAAATTAGATTTTGAGATTTTGCATCCGAGATCTCAAAACGTTAGT
AAAAAAATTTCAA

AACAAGTCGAAGCTAAACCATTTGATCCAGCTTCCTATAAAGAAAAAGTTATAGCCAAATTAAAGCCGGTTTA
TGAAGCTACTAG

TGAAAAAATCACAAATGATGCTTGGTTGGATGAAATGCGAAGGATTTGCAAAAACAAAAATTAGAAGAACAA
TATATTTCTGGA

AAAGTAGCGATATCAGAGGCTGGAATAAACAAGAAGCTATAGATGCAGCATATAATAAATATTCAAGTCAAA
CAGATCCAGACT

CTCTTCCTAGTCAGTATAAACAAGGTAATAAAGAAAAATGAACAAGAAAAAGGGCGTCAAGATTTAATCCAGAC
TCGTGATCTGAC

GTTGAAAGCCATTCAAGAAGACAAATGGCTAACAGAGCAGGAGAAAACAATTCAAAAAGAAGAAGCATTAAAA
GCTTTTGAAACT

GGTATAGAAAGTGTTAATCAAACAGTATCATTAGAACAGTTGAAGCAACGGTTAATAGT
GTATAAAGCTTCTGAAAAAGATTCAG
AGAAAAAGAATATCCTGAGTCAATTCCTAATCAGCATATTCCAGGGAAAGAAAAAGAAGTTAAAGCTGCTAA
ACAAGAAGAACT
TAAAAAATTACATGACACAACCTCTTGAAAAAATCAATCAAGATAAATGGTTGACGCCAGACCAACAAGCTGAA
CAGTTAAACAA
GCGGAAGTTACTTTTAAAAAAGGCCAAGAAGCAATTAAAAGTGCTCAGACTTTAACTCAGCTTGAGACAGACT
TAGCTGATTATG
TTTCTGAGAATGAAGGTAAGGGAAATTCTATTCCCGATAAATACAAATCTGGCAATAAAGATGATTGGTAAA
TAAGGCTGAAGT
TAACTTAAGGAAGCTCACGAAGCTACTAAACAAGCAATTGAAAAAGATCCATGGTTGAGTCCGGAACAGAAA
AAAGCTCAAAA
AGAAAAAGCCAAAGCAAGACTAGATGAGGGCTTGAAAAGCTCTTAAAGCTGCAGATAGTTTAGAGATTCTTAAA
GTGACAGAAGA
AGCTTTCGTTGATAAAGAAAAAATCCAGATTCAATTCCAAATCAACATAAAGCTGGAAGCTGCTGATCAAGCT
AGAAAACAAGCT
TTAGATAGTTTAGATAAGGAGGTTCAAAAAGAGTTAGAGTCAATTGATAACGATAATACTCTAACAAGCTGATG
AGAAAGCAGCTG
CTAAGAAAAAGTCAATGACGCTTATGATGTAGCTAAGCAAACAGCTATGGAAGCCAATTCTTATGAAGATTT
GACTACTATTAAA
GATGAGTTCTTATCTAATTTACCTCATAAAACAAGGAACGCCGCTTAAAGATCAACAATCTGATGCTATTGCAG
AATTAGAGAAGA
AGCAGCAAGAAATTGAAAAAGCTATTGAGGGTGATAAAACATTACCAAGAGACGAAAAAGAGAAACAAATTGC
TGAAGCTAAGG
AACGCTTAAATCTGACACGCAAAAAGTTAAAGATGCTAAAAATGCTGATGCTATTAAAAAGCATTGAAGA
AGGGAAAGTGAA
TATTCCTCAAGCACATATCCCAGGTGATTTGAACAAGGATAAAGAAAACTTCTTGACAGATTGAAGCAAAAA
GCAGATGATACT
GAAAAAGCTATTGATGTTGATAAACTCTGACAGAAGATGAGAAAAAGAGCAAAAAGTCAAAACAAAAGCTG
AACTTGAAAAAG
CTAAAAGTATGTTAAAAATACTCAGACACGTGAAGAACTAGATAAAAAAGTTCCAGAACTTAAGAAAGCTAT
TGAAGACACTCA
CGTTAAAGGTAATCTTGAAGGTGTTAAGAATAAGGCTATTGAAGATCTTAAAAAGCTCATACTGAAACAGTT
GCTAAATAAATG
GTGATGATACCCTTGACAAAGCTACTAAAGAAGCTCAAGTGAAAGAAGCTGACAAAGCTTTGGCAGCAGGTAA
AGATGCGATCA
CTAAAGCAGATGATGCTGATAAAGTAAGTACAGCTGTTACAGAGCACACACCAAAAATTAAAGCAGCACATAA
AACTGGTGACC
TTAAAAAAGCTCAAGTAGATGCTAACACAGCTCTTGACAAAGCAGCTGAAAAAGAACGTGGAGAAATCAATAA

AGATGCTACACT

AACGACAGAAGATAAAGCAAAACAACCTGAAAGAAGTTGAGACAGCTCTTACTAAAGCTAAAGATAACGTGAAA
GCTGCTAAGAC

AGCAGACGCTATCAATGACGCACGTGATAAAGGCGTAGCAACAATTGATGCCGTCCATAAAGCAGGTCAAGAC
TTAGGTGCTC

GTAAGTCAGGTCAAGTCGCTAAACTTGAAGAAGCAGCTAAAGCAACGAAAGACAAGATTTGAGCTGATCCAAC
TTTGACAAGCA

AAGAAAAAGAAGAGCAATCTAAAGCTGTTGACGCTGAACTTAAGAAAGCGATAGAAGCTGTTAACGCAGCTGA
CACAGCTGACA

AGGTTGACGATGCTCTTGGTGAAGGTGTTACAGACATCAAGAACCAACACAAGTCAGGTGACTCTATCGACGC
TCGTCGTGAG

GCTCATGGTAAGGAACTTGATAGAGTCGCTCAAGAAACTAAAGGTGCGATTGAAAAAGACCCTACTTTGACGA
CTGAAGAAAAA

GCTAAACAAGTTAAAGACGTAGATGCCGCTAAAGAAAGAGGCATGGCTAAGCTTAATGAAGCTAAAGATGCTG
ATGCTTTAGAC

AAAGCTTACGGTGAAGGTGTTACAGACATCAAGAACCAACACAAGTCAGGTGACCCTGTGACGCTCGTCGTG
GATTACACAA

CAAGTCAATCGACGAAGTGGCGCAAGCAACTAAGGACGCTATCACAGCAGATACGACTTTGACTGAAGCTGAA
AAAGAAACAC

AACGTGGCAATGTTGATAAAGAAGCAACTAAAGCTAAAGAAGAAGCTTGCTAAGGCTAAAGATGCTGATGCTTT
AGACAAAGCGT

ACGGTGACGGTGTAACCAGCATCAAGAACCAACACAAGTCAGGTAAAGGTCTTGACGTTTCGTAAAGATGAGCA
CAAGAAAGCT

CTTGAAGCTGTTGCTAAACGTGTCACTGCTGAAATTGAGGCTGATCCAACCTTAACACCAGAAGTGAGAGAAC
AACAAAAAGCA

GAGGTTCAAAAAGAGCTTGAAGCTGCGACTGATAAGATTGCTGAAGCTAAAGATGCAGATGAAGCAGACAAAG
CTTACGGTGAC

GGTGTACAGCGATCGAAAATGCCACGTTATTGGTAAAGGTATCGAAGCTCGTAAAGACCTTGCTAAGAAAG
ACCTTGCTGAA

GCTGCTGCTAAGACAAAAGCTCTCATTATTGAAGACAAAACGCTTACTGATGATCAACGTAAAGAACAGTTAT
TAGGTGTTGATA

CAGAGTATGCTAAAGGTATCGAGAATATTGATGCAGCTAAGGATGCTGCAGGTGTTGATAAAGCATATAGTGA
CGGTGTTTCGTG

ACATCCTGGCACAGTACAAAGAAGGTCAAAACCTTAATGATCGTCGTAATGCTGCCAAAGAATTTCTTCTTAA
AGAAGCAGACAA

AGTGACGAACTAATCAATGATGATCCAACCTTGACTCATGACCAAAAAGTTGATCAAATCAACAAAGTTGAA
CAAGCTAAGTTA

GACGCAATCAAGTCTGTTGATGATGCTCAAAACAGCTGATGCTATCAATGATGCTCTTGCTAAGGGTATTGAAA
ACATCAACAACC

AATACCAACATGGCGATGGCGTTGATGTTTCGTAAAGCGACTGCCAAAGGCGATCTTGAAAAAGAAGCTGCTAA
AGTGAAAGCTC

TTATTGCTAAGGATCCGACCTTAACTCAAGCTGATAAAGACAAACAAACAGCAGCGTTGACGCAGCTAAGAA
TACAGCAATTG

CAGCGTTGATAAAGCGACAACAACCTGAAGGCATTAACCAAGAACTTGGTAAAGGCATCACAGCTATCAATAA
AGCTTACCGTC

CAGGTGAAGGTGTTAAAGCACGTAAAGAAGCCGCTAAAGCTGATCTTGAAAAAGAAGCTGCTAAAGTGAAAGC
TCTTATTACTA

ACGACCCAACCTTAACAAAAGCTGATAAAGCTAAACAAACAGAAGCTGTTGCTAAAGCCCTTAAAGCTGCTAT
CGCAGCGGTTG

ATAAAGCGACAACAGCTGAAGGCATTAACCAAGAACTTGGTAAAGGCATCACAGCTATCAATAAAGCTTACCG
TCCAGGTGAAG

GTGTTAAAGCACGTAAAGAAGCCGCTAAGGCTGATCTTGAAAGAGAAGCTGCTAAGGTTTCGTGAAGCTATCGC
TAACGACCCAA

CCTTAACAAAAGCTGATAAAGCTAAACAAACAGAAGCTGTTGCTAAAGCTCTTAAAGCTGCTATCGCAGCGGT
TGATAAAGCGA

CAACAGCTGAAGGCATTAACCAAGAACTTGGTAAAGGCATCACAGCTATCAACAAAGCTTACCGTCCAGGTGA
AGGTGTTGAAG

CACATAAAGAAGCTGCTAAAGCTAATCTTGAAAAAGTAGCTAAAGAACTAAAGCTCTTATTTAGGAGACCG
TTACTTGAGCGA

AACTGAAAAAGCAGTCCAAAAACAAGCTGTTGAGCAAGCTCTTGCGAAAGCACTTGGTCAAGTTGAGGCTGCT
AAGACAGTTGA

AGCTGTTAAGTTGGCAGAAAACCTTGGTACTGTAGCTATCCGTTTCAGCATATGTTGCTGGTTTAGCTAAAGAT
ACTGATCAAGCA

ACAGCTGCTCTTAACGAAGCGAAACAAGCTGCTATTGAAGCTCTTAAACAAGCTGCGGCAGAAACACTTGCTA
AGATTACAACT

GATGCTAAATTGACTGAAGCTCAAAAAGCTGAACAATCAGAAAATGTATCATTAGCGCTTAAGACGGCTATTG
CGACTGTTTCGTT

CAGCACAATCTATTGCGTCTGTGAAAGAAGCAAAAAGATAAAGGTATTACTGCTATCCGTGCAGCCTATGTGCC
TAATAAGGCAG

TCGCAAAATCATCGTCAGCGAACCATCTTCCAAAATCAGGTGATGCAAACCTCAATTGTTCTTGTGGCTTAGG
AGTTATGTCTCT

TCTTTTAGGTATGGTGCTTTATAGCAAGAAAAAAGAAAGTAAAGACTAA

SPy0747

Seq ID 37

ATGATTAACAAGAAATGTATAATACCTGTTTCATTGTTGACACTAGCTATTACGCTTACTAGTGTTGAAGAAG
TTACTTCAGCCA

AAATTTGACTTATGCCAATGAAATCGTAACACAAAAGGCCAAAAGAGAGAATCTGTTATTAGTGATAAATCGAAT

TTTCCCGTCATAT

CACCTTACCTAGCAAGTGTGGATTTTGGTGAGAGAAAAACACCTTTGCCAACACCTGATAAAGGAGTAAAAGT
AACTACTGAACA

GTCTATTGCTCAAGTAAGAAAGGGGCCTGAAGAAAGACCCTATACTGTTACTGGCAAGATTACGAGTGTGATC
AATGGCTGGGG

AGGCTATGGCTTTTATATTCAAGATAGTGAAGGTATTGGACTTTATGTTTATCCTCAAAAAGATTTAGGATAC
AGTAAGGGAGATA

TTGTTCAATTAACAGGTACACTTACTCGCTTTAAAGGTGATTTACAACCTCCAACAGGTGACTGCACACAAAAA
GTTAGAGTTATCT

TTTCCGACTTCTGTAAAGAAGCAGTAATATCAGAATTAGAAACAACAACACCCTCAACATTAGTTAAGTTAT
CTCACGTGACAGT

TGGAGAATTATCAACTGATCAATATAACAACACATCTTTCCTTGTAAGGGATGATAGTGGTAAAAGTATAGTT
GTTTCATATAGATC

ATCGTACAGGGGTAAAGGGGCTGATGTTGTTACTAAAAATAAGTCAGGGTGATTTGATTAACCTCACAGCCAT
ATTGTCTATTGT

TGATGGTCAATTACAATTAAGACCGTTTTCTCTTGAACAATTGGAAGTGGTTAAAAAGGTCACAAGCTCAAAT
AGTGATGCTTCAT

CTCGTAATATTGTGAAAATAGGCGAGATTCAAGGAGCTAGTCATACGTCGCCACTTCTCAAAAAAGCGGTAC
CGTAGAACAGG

TTGTTGTCACTTATTTAGACGATTCCACTCATTTTTATGTTCAAGATCTTAATGGTGATGGTGATTTAGCGAC
TTCAGATGGTATT

CGTGTTTTTGCTAAAAACGCTAAGGTTCAAGTCGGCGATGTTTTGACCATTTTCAGGTGAAGTGAAGAATTCT
TTGGTCGTGGTT

ATGAGGAACGTAAGCAGACTGACCTTACCATCACCCAAATTGTGGCTAAAGCAGTGACCAAAACAGGGACAGC
TCAAGTTCCAT

CACCGCTTGTTTTAGGGAAAGATCGTATCGCGCCAGCCAATATTATTGATAATGATGGCTTGCCTGTGTTTGA
TCCAGAAGAAG

ACGCTATTGATTATTGGGAATCAATGGAAGGCATGTTAGTGGCGGTTGATGATGCTAAAATCCTTGGTCCAAT
GAAAAATAAAGA

AATTTATGTCTTACCTGGCTCTAGTACAAGACCGTTAAATAATTCAGGTGGAGTATTACTTCCAGCTAATTCT
TATAACACAGATG

TGATTCCTGTTCTTTTCAAAAAAGGCAAAACAAATTATTAAAGCAGGAGACTCTTACAAAGGAAGATTAGCTGG
GCCAGTATCTTA

TAGCTATGGTAATTACAAGGTCTTTGTTGATGACAGCAAAAAACATGCCAAGTTTAATGGATGGTCATCTAAAA
CCTGAAAAAACA

AACTTGCAAAAAGACCTTAGCAAGTTAAGCATTGCTTCTTACAATATTGAAAACCTCTCAGCCAATCCTTCTT
CAACTAAAGATGA

GAAGGTCAAACGGATTGCCGAATCCTTTATTATCATGATCTGAATGCTCCAGACATTATTGGATTAATTGAAGTC
CAAGATAATAAT

GGGCCGACTGATGATGGGACAACGGATGCGACACAAAGCGCGCAACGCCTCATTGATGCTATTAAAAAAGTAG
GTGGCCCAAC

TTATCGTTATGTTGATATTGCTCCAGAAAAATAATGTTGACGGAGGTCAACCAGGTGGTAATATTCGAACAGGA
TTCCTTTATCAA

CCAGAGCGCGTCAGCCTTTCTGATAAGCCAAAAGGCGGTGCTCGTGATGCTCTAACTTGGGTAAATGGAGAAT
TAAACCTTAGT

GTTGGTCGAATTGATCCAATAACGCCGCTTGGAAGATGTTTCGTAAATCACTAGCAGCAGAATTTATCTTCC
AAGGTCGTAAAG

TCGTTGTTGTTGCAAATCATTGAACTCTAAGCGTGCGGATATGCTCTTTATGGTTGTGTGCAACCAGTCAC
TTTTAAATCTGA

GCAAAGACGTCACGTCTTGGCTAATATGCTAGCACAAATTTGCGAAAGAAGGGGCAAAACACCAAGCTAATATT
GTGATGCTAGG

TGACTTTAATGATTTTGAATTCACAAAGACGATTCAATTAATCGAAGAAGGTGACATGGTTAACTTGGTGAGC
CGACATGATATTT

CAGATCGGTATTCTTATTTTACCAAGGCAATAATCAGACCCTTGATAATATATTAGTTTCACGCCATTTACT
TGATCACTACGAA

TTTGACATGGTTCATGTGAATTCCTTATGGAAGCTCACGGACGCGCATCAGATCATGATCCATTGTTAC
TTCAATTATCATT

TTCCAAAGAAAATGATAAGGCAGAGTCTTCTAAACAAAGTGTAAGCTAAAAAACTTCAAAAGGAAAACTG
TTGCCAAAAACA

GGAGATAGTCTTGTTTATGTGATAACGCTACTAGGAACGGCTAGTTTATTAGTGCCTATTTTATTATTGACTA
AAGGCAAAAAGG

AATCATAG

SPy0777

Seq ID 38

GTGATTTCTTTTGCCCCATTTTAAAGCCCCGAAGCTATTAAACATTTGCAAGAAAACGAAAGGTGCAGAGATC
AGTCTCAAAAAC

GCACAGCTCAACAAATTGAAGCAATTTATACTAGTGGCCAAAATATACTTGTATCAGCTTCTGCTGGTTCAGG
AAAAACCTTTGT

AATGGTCGAACGCATACTTGATAAAATTTGAGAGGTGTTTCAATTGATCGGCTTTTATCTCAACCTTTACT
GTTAAAGCAGCTA

CAGAACTGCGTGAGCGGATTGAAAACAAATTATACTCACAAATTGCTCAAACCTACAGATTTTCAAATGAAAGT
TTATTTAACAGAA

CAATTGCAATCTCTTTGTCAAGCTGATATTGGTACTATGGATGCTTTTGACAGAAAGTAGTAAGTCGCTATG
GTTATAGCATTG

GCATTTTCATCCCAATTTGATCATGCAGGATAAAGCAGAACAAAGATGTTTTAAAGCAAGAGGTGTTTAGCAA
ACTCTTTAATGA

GTTTATGAATCAAAAAGAGGCACCGGTGTTTAGGGCTCTGTGAAAAATTTTCTGGTAACTGTAAAGACACT

TCAGCTTTTAGA

GAGTTAGTTTATACTTGTTATTCTTTTAGCCAATCGACAGAAAACCCAAAAATATGGTTGCAAGAAAATTTTC
TAAGCGCTGCTAA

AACTTACCAAAGACTTGAAGATATCCCGGATCATGATATTGAACTCTTACTTTTGGCAATGCAAGACACTGCA
AATCAGCTAAGA

GATGTGACTGATATGGAAGATTATGGGCAGCTGACTAAGGCAGGTAGCCGATCTGCTAAATACACTAAACACT
TAACGATCATA

GAAAAGTTGTCTGATTGGGTGCGTGATTTTAAATGTTTGTATGGAAAAGCCGGATTGGATCGGTTGATCAGAG
ATGTGACAGGC

CTTATACCATCTGGGAATGATGTTACAGTCTCGAAGGTAAAAATACCCTGTTTTTAAGACCTTGCATCAAAAAT
TAAACAATTTAG

GCATTTAGAAACAATTTTAATGTATCAGAAAGACTGTTTTTCCTTATTGGAACAGTTACAAGATTTTGTGCTT
GCGTTTTTCAGAAG

CTTATTTAGCTGTCAAAATACAAGAAAGTGCTTTTGAATTTTCAGATATTGCACACTTTGCAATCAAAATTTT
AGAGGAAAATACG

GATATTCGCCAATCCTATCAGCAACACTATCATGAGGTGATGGTTGATGAATATCAAGATAACAATCATATGC
AAGAGCGACTCC

TGACCTTACTATCGAACGGTCATAATCGCTTTATGGTAGGAGATATCAAACAATCGATCTATCGATTTTCGGCA
AGCCGATCCTCA

GATTTTTAATCAAAAGTTTAGAGACTATCAAAAAAACCTGAGCAGGGGAAAGTGATTTTACTCAAAGAAAAC
TTTCGTAGCCAAT

CAGAGGTGTTAAATGTCAGCAATGCTGTTTTTAGTCACTTAATGGACGAATCAGTAGGAGACGTCTTATACGA
TGAGCAACATCA

GTTAATAGCAGGTAGTCATGCTCAAAACAGTCCCCTATCTAGACCGTCGTGCTCAGTTATTGCTATATAATAGC
GATAAAGATGAT

GGCAACGCCCCCTTCAGATAGTGAGGGTATTTCAATTTAGTGAGGTTACAATTGTTGCCAAAGAAATTATTAAGC
TTCACAATGATA

AGGGTGTCCCTTTTGAAGACATTACGTTACTCGTTTCTTCAAGAACAAGAAATGATATCATTTCTCATACATT
CAATCAATATGGT

ATTCCTATAGCAACAGATGGTGGGCAGCAAACTATCTTAAATCTGTTGAAGTGATGGTTATGTTAGATACAT
TACGCACCATTA

ATAACCCAAGAAATGATTATGCCCTTGTGGCTTTACTGCGCTCACCGATGTTTGCCTTTGATGAGGATGATTT
AGCAAGAATAGC

ACTTCAAAAAGACAATGAGCTAGATAAAGATTGCCTATATGACAAGATACAAAGGGCTGTGATTGGAAGAGGT
GCTCATCCTGA

ATTGATTCACGATACCTTGCTTGGCAAGTTAAATGTTTTTTTAAAGACGTTGAAAAGCTGGCGTCGATACGCT
AAGCTAGGGTCG

TTGTATGACTTGATTTGGAAAATTTTAAATGATCGTTTTTATTTTGATTTTGTAGCTAGTCAAGCAAAAGCAG
AACAAGCACAAGC

TAATCTATACGCATTAGCTCTACGTGCTAATCAGTTTGAAAAATCGGGCTATAAAGGGCTATACCGTTTTATT
AAAATGATTGATA
AGGTA CTTGAGACGCAAAATGACTTAGCTGATGTGGAAGTGGCTACTCCTAAACAAGCTGTTAATTTAATGAC
CATTCACAAGTC
TAAAGGTTTACAATTTCCGTATGTATTTATCCTTAATTGTGACAAGCGCTTCTCAATGACAGATATTCATAAA
TCATTTATTCTGAA
TCGGCAGCACGGTATCGGTATCAAGTACCTTGCAGATATCAAAGGTTTACTTGGTGAAACAACACTCAATTCT
GTTAAAGTAAGC
ATGGAACCTTACCTTATCAATTGAACAAACAAGAGTTGCGCTTAGCAACTTTATCAGAAGAAATGCGCTTAC
TGTATGTTGCTAT
GACACGAGCTGAAAAAAAAAGTTTATTTTATTGGTAAAGCTAGTAAGAGCAAAAGTCAAGAAATCACAGATCCT
AAAAAGTTAGGC
AAACTTTTGCCGCTGGCTTTACGAGAACAGTTATTGACATTCCAAGATTGGCTATTAGCAATAGCAGATATAT
TTTCAACTGAAGA
TCTTTATTTTGATGTTTCGCTTTATTGAAGATAGTGATTTGACACAAGAGTCAGTCGGACGACTTCAAACACCA
CAGTTATTAAATC
CAGATGATCTTAAAGATAATCGTCAATCAGAAACAATTGCACGGGCTTTAGATATGTTAGAAGCAGTGTCTCA
ATTGAATGCCAA
TTATGAAGCAGCTATTCAATTTGCCAACAGTTCGAACGCCTAGCCAACTTAAGGCAACTTACGAGCCTTTATTA
GAACCCATTGGT
GTAGATATTATAGAGAAATCTTCTCGATCGCTATCTGATTTTACTTTGCCACATTTTTCAAAAAAGCAAAAG
TTGAAGCAAGTCA
TATTGGATCAGCTCTTCATCAGTTGATGCAGGTGCTCCCTTTGTCAAAACCGATAAATCAACAAACGCTTTTA
GACGCTTTAAGA
GGAATTGATAGTAACGAAGAGGTAAAAACAGCTCTTGATCTCAAAAAAATAGAGTCGTTCTTTTGTGATACAA
GCCTAGGCCAAT
TTTTTCAGACTTACCAAAAAACACTTGTATCGAGAAGCGCCATTTGCTATTTTAAACTTGACCCTATCAGTCA
AGAAGAGTATGTC
CTACGTGGTATTATAGATGCCTACTTTTTGTTTGATGATCATATTGTATTAGTGGACTATAAAACAGATAAAT
ACAAGCAGCCCAT
TGAGTTAAAAAAGCGTTACCAACAACAGTTGGAGTTATATGCAGAAGCTCTCACTCAAACGTATAAACTTCCT
GTGACTAAGCGC
TATCTTGTTTTAATGGGAGGTGGAAAGCCAGAAATTGTGCAAGTTTAA

SPy0789

Seq ID 39

ATGGTAAAAACGGATTTTAAATTACGCTATCAAGGGAGCGCAATTGGCTATCTATGGTCTATCTTAAAACCGC
TTATGATGTTTAC
GATTATGTACTTGGTATTTATTCGTTTTCTTCGCCTGGGTGGAAATGTACCTCATTTTCCAGTAGCGCTTTTA

TTGGCAAATGTCA

TCTGGTCTTTTTTTCAGAAGCAACTAGCATGGGAATGGTATCTATTGTATCTCGGGGAGACTTGTTCGAAA
ATTGAACTTTTCT

AAGCACATCATTGTTTTTTCGGCAGTGTTAGGAGCTTTAATTAATTTTCTTATTAATTTGGTTGTTGTTTTAA
TTTTTGC GTTGATAA

ATGGTGTGACTATATCAGGGTATGCTTATCTCTCTCTTTTTCTTTTTATAGAATTAGTTGTTTTAGTGCTTGG
AATTGCATTGTTAT

TGTCGAATGTCTTTGTTTATTATCGTGACTTAGCTCAAGTCTGGGAAGTACTATTACAAGCAGGTATGTATGC
CACTCCAATCATT

TATCCGATCACTTTTGTTT TAGATAGCCACCCTTTGGCGGCAAAGTTGTTGATGCTAAATCCAGTAGCACAAA
TGATTCAAGATTT

TCGTTATTTATTGATTGACAGGGCCAACGTAACGATTTGGCAGATGTCAACCAATTGGTTTTACATTGTTATT
CCATATTTAGTAC

CATTTGTTATATTATTTATTGGCATCTTTGTCTTTAAGAAAAATGCCGATAGATTTGCGGAGATTATTTAA

SPy0839

Seq ID 40

ATGACATTTTATCTGATTTGATATCATTAATGACAAAAATCAGATTATCTTGGGTAATAAAGGCGGGTATTT
TTCAATTATTATTT

GTAACGATTGCTAATATTGTCTTATCAGAATTTTCTATTTTATATTAGACGTTACTGGTCAATATCATTTAG
ATAAAGACAATGTT

GTGACTTTTTTAAAAAATCCTATAGCACTTGCTTTATTAGGTGCCTATTTATTTTTATTAGCTGCTTTTATTC
ACCTTGAGTTTTTTTG

CTCTATATCGAATTATTGCGGATCAAGAAATTAGTTTCTATCTTTTTAGAAAACAGTTTCTTATTACCTAAG
GGGGCTTTGAAA

ACATTTTCTGGTTACCAATTATTACTTTTTTTGCTTTATATCCTATTGACTATTCCAGTCTTACATATTGGTT
TATCTTCTGTGATTA

CTCAAAAGCTTTATCTTCCAGAATTTATTGTTGGGGAATTATCAAAGATAACTAGCACAAAGTACTTGCTTTA
TGGCAGTCTTATT

CTTGTGTTTTACCTTAACCTAAGATTAGTATATTTTACCATTGATAGCAATCAACCATCGTACGTTGCTC
AAGCATGGAGAGA

GAGTTGGCAAAAGACTAAAAAGAAACATGTATTGTTATGGATGAACTTTTGAATCAATGGTCTTACGATT
GTAGTCTTATCGC

TAGCTATTTCCATGATTCTTATTTTTGTTGATATGTTTAATCCTAAGGGGAATAATATTATTGTTTCAGCTGGG
AGCTTTGACCTTTA

CATGGGAACCTATTTTTTTTACTACTATTTTTTTTAACTCTGTTCAGCAATGATTTTAAAGAGGCAATTGA
ACCACAAAAGCAAT

ATGATGAGCCAAGAAGAAGTAATAAGGCATATGTTGTAATCTTTATCGTGGTTACAGTAGGTTTTGCTTATCA
ATCTCTTGAACGT

TTAACTTTTTTTGACACATCTCACTCTAAGACAGTTATCGCGCATAGAGGACTTGTATCAGCAGGTGTAGAAA
ATTCTCTGGAAG

CCCTTGGAAGGTGCTAAGAAAGCAGGAAGTGATTATGTAGAACTGGATCTAATCTTGACTAAGGATAATCACTT
TGTGGTGTCTCA

TGATAATCGATTGAAGCGTTTAGCTGGAGTAAATAAGACGATTCGCAACTTAACCTTAAAAGAAGTTGAACAT
CTAACGAGTCAT

CAAGGACATTTTTTCAGGGCGTTTTGTTTCTTTTGACACTTTTTATCAAAAGGCTAAGAAGTTGAATATGCCAT
TACTTATTGAACT

CAAGCCAATTGGTACAGAACCTGGAAATTATGTCGATTTGTTTTTAGAACTTATCATCGACTTGGTATAAGC
AAAGATAATAAAG

TCATGTCTTTAGATTTAGAAGTAATAGAAGCTATCAAGAAAAAAATCCATCAATTACGACTGGTTATATCAT
ACCAATTCAATTTG

GATTTTTTTGGAGATGAATTTGTTGATTTCTATGTCATTGAAGACTTTTCTTATCGGTCTTATTTGTCGTCCCA
AGCTTTTTTGAATA

ATAAAGAAATTTACGTTTGGACTATTAATGATCCCAAGCGCATAGAGCATTATCTCCTAAAGCCTATTCAGGG
AATTATTACAGAC

CAACCAGCTTTAACTAATCAATTGATTAAAGACTTAAAAACAAGATAATTCTTATTTTAGTCGATTAGTCAGAA
TTATTAGTAGTCTT

TATTAA

SPy0843

Seq ID 41

ATGAAGAAACATCTTAAACAGTTGCCTTGACCCTCACTACAGTATCGGTAGTCACCCACAATCAGGAAGTTT
TTAGTTTAGTCA

AAGAGCCAATTCTTAAACAAACTCAAGCTTCTTCATCGATTTCTGGCGCTGACTACGCAGAAAGTAGCGGTAA
AAGCAAGTTAAA

GATTAATGAAACTTCTGGCCCTGTTGATGATACAGTCACTGACTTATTTTCGGATAAACGTACTACTCCTGAA
AAAATAAAAGATA

ATCTTGCTAAAGGTCCGAGAGAACAAGAGTTAAAGGCAGTAACAGAGAATACAGAATCAGAAAAGCAGATCAC
TTCTGGATCTC

AACTAGAACAATCAAAAGAGTCTCTTTCTTTAAATAAAACAGTGCCATCAACGTCTAATTGGGAGATTTGTGA
TTTTATTACTAAG

GGGAATACCCTTGTTGGTCTTTCAAAATCAGGTGTTGAAAAGTTATCTCAAAGTATCATCTCGTATTGCCTA
GTCAAGCAGCAG

ATGGAAGTCAATTGATACAAGTAGCTAGTTTTGCTTTTACTCCAGATAAAAAGACGGCAATTGCAGAATATAC
CAGTAGGGCTGG

AGAAAATGGGGAAATAAGCCAAGTAGATGTGGATGGAAAAAGAAATTATTAACGAAGGTGAGGTTTTTAATTCT
TATCTACTAAAG

AAGGTAACAATCCCAACTGGTTATAAACATATTGGTCAAGATGCTTTTGTGGACAATAAGAATATTGCTGAGG

TTAATCTTCCTGA

AAGCCTCGAGACTATTTCTGACTATGCTTTTGCTCACCTAGCTTTGAAACAGATCGATTTGCCAGATAATTTA
AAAGCGATTGGA

GAATTAGCTTTTTTTGATAATCAAATTACAGGTAACTTTCTTTGCCACGTCAGTTAATGCGATTAGCAGAAC
GTGCTTTTAAATC

AAACCATATCAAAACAATTGAGTTTAGAGGAAATAGTCTAAAAGTGATAGGGGAAGCTAGTTTTCAAGATAAT
GATCTGAGTCAA

CTAATGCTACCTGACGGTCTTGAAAAAATAGAATCAGAAGCTTTTACAGGAAATCCAGGAGATGATCACTACA
ATAACCGTGTG

TTTTGTGGACAAAATCTGGAAAAATCCTTCTGGTCTTGCTACTGAAAATACCTATGTTAATCCTGATAAGTC
ACTATGGCAGGAA

AGTCCTGAGATTGATTATACTAAATGGTTAGAGGAAGATTTTACCTATCAAAAAATAGTGTTACAGGTTTTT
CAAATAAAGGCTT

ACAAAAAGTAAACGTAATAAAAACTTAGAAAATCCAAAACAGCACAAATGGTGTTACTATTACTGAAATTGGT
GATAATGCTTTTC

GCAATGTTGATTTTCAAAAATAAACTTTACGTAAATATGATTTGGAAGAAGTAAAGCTTCCCTCAACTATTTCG
GAAAATAGGTGCT

TTTGCTTTTCAATCTAATAACTTGAAATCTTTTGAAGCAAGTGACGATTTAGAAGAGATTAAAGAGGGAGCCT
TTATGAATAATCG

TATTGAAACCTTGAATTAAAAAGATAAAATTAGTTACTATTGGTGATGCGGCTTCCATATTAATCATATTTAT
GCCATTGTTCTTCC

AGAATCTGTACAAGAAATAGGGCGTTCAGCATTTCCGGCAAAATGGTGCAAATAATCTTATTTTTATGGGAAGT
AAGGTTAAGACC

TTAGGTGAGATGGCATTTTTATCAAAATAGACTTGAACATCTGGATCTTTCTGAGCAAAAACAGTTAACAGAGA
TTCCTGTTCAAG

CCTTTTCAGACAATGCCTTGAAAGAAGTATTATTACCAGCATCACTGAAAACGATTCGAGAAGAAGCCTTCAA
AAAGAATCATTT

AAAACAACCTGGAAGTGGCATCTGCCTTGTCCCATATTGCTTTTAATGCTTTAGATGATAATGATGGTGATGAA
CAATTTGATAATA

AAGTGGTTGTTAAACGCATCATAATTCCTACGCACTAGCAGATGGTGAGCATTTTATCGTTGATCCAGATAA
GTTATCTTCTACA

ATAGTAGACCTTGAAAAGATTTTAAAACTAATCGAAGGTTTAGATTATTCTACATTACGTCAGACTACTCAAA
CTCAGTTTAGAGA

CATGACTACTGCAGGTAAAGCGTTGTTGTCAAAATCTAACCTCCGACAAGGAGAAAAACAAAAATTCCTTCAA
GAAGCACAATTT

TTCCTTGGCCGCGTTGATTTGGATAAAGCCATAGCTAAAGCTGAGAAGGCTTTAGTGACCAAGAAGGCAACAA
AGAATGGTCAG

TTGCTTGAAAGAAGTATTAACAAAGCGGTATTAGCTTATAATAATAGCGCTATTAaaaaagctaATGTTAAGC
GCTTGAAAAAGA

GTTAGACTTGCTAACAGGATTAGTTGAGGGAAAAAGGACCATTAGCGCAAGCTACAATGGTACAAGGAGTTTAT
TTATTAAAGAC

GCCTTTGCCATTGCCAGAATATTATATCGGATTGAACGTTTATTTTGACAAGTCTGGAAAATTGATTTATGCA
CTTGATATGAGTG

ATACTATTGGCGAGGGACAAAAAGACGCTTATGGTAATCCTATATTAAATGTTGACGAGGATAATGAAGGTTA
TCATGCCTTGGC

AGTTGCCACTTTAGCTGATTATGAGGGGCTCGACATCAAAACAATTTTAAATAGTAAGCTTAGTCAATTAACA
TCTATTCGTCAGG

TACCGACTGCAGCCTATCATAGAGCCGGTATTTTCCAAGCTATCCAAAATGCAGCGGCAGAAGCAGAGCAGTT
ATTGCCTAAAC

CAGGTACGCACTCTGAGAAGTCAAGCTCAAGTGAATCTGCTAACTCTAAAGATAGAGGATTGCAATCAAACCC
AAAAACGAATA

GAGGACGACACTCTGCAATATTGCCTAGGACAGGGTCAAAAGGCAGCTTTGTCTATGGAATCTTAGGTTACAC
TAGCGTTGCTT

TACTGTCACTAATAACTGCTATAAAAAAGAAAAATATTAA

SPy0872

Seq ID 42

ATGAAAAATATTTTATTTTAAAAAGTAGTGTATTGAGTATCCTGACTAGTTTTACTCTATTAGTTACAGATG
TTCAAGCAGATCAA

GTTGATGTGCAATTCCTTGGCGTCAATGATTTTCACGGCGCTCTTGATAATACCGGAACAGCTTACACACCAA
GTGGTAAATAC

CAAATGCTGGGACGGCTGCTCAATTAGGTGCTTATATGGATGACGCTGAGATAGACTTCAAGCAAGCAAATCA
AGACGGAACAA

GTATACGTGTTCAAGCTGGAGATATGGTCGGAGCCAGTCCTGCTAACTCTGCACTTTTACAAGATGAGCCTAC
TGTCAAAGTCT

TTAACAAAATGAAATTTGAATATGGCACTCTTGGTAATCATGAATTTGACGAAGGACTAGATGAATTTAACCG
TATCATGACAGGT

CAAGCGCCTGATCCTGAATCAACAATTAATGATATCACCAAACAATATGAGCACGAAGCTTCGCATCAAACCA
TCGTCATTGCTA

ATGTTATTGATAAAAAAACCAAGGATATCCCCTATGGTTGGAAACCTTATGCTATAAAAGACATAGCCATTAA
TGACAAAATCGTT

AAGATTGGCTTCATTGGTGTGTGACTACAGAGATTCCAAATCTCGTTTTAAAGCAAACTATGAACACTATC
AATTTTTAGATGT

AGCTGAAACCATTGCCAAATATGCTAAAGAACTACAAGAACAACATGTTTCATGCTATT
GTGGTTTTAGCTCATGTTCTGCAACA

AGTAAAGATGGTGTGTTGATCATGAAATGGCTACGGTTATGGAAAAAGTGAACCAAATCTATCCCGAACATA
GCATTGATATTA

TTTTTGCAGGACATAATCATCAATACACTAATGGAACCTATCGGTAAAACACGTATCGTTCAAGCCCTCTCTCA

AGGAAAAGCTTA

TGCAGATGTCCGTGGTACGCTAGATACTGATACCAATGATTTTATTAAAACTCCATCAGCAAATGTTGTTGCT
GTAGCACCAGGT

ATCAAAACAGAAAATTCAGATATCAAAGCTATAATAAATCATGCTAATGATATTGTTAAAACAGTTACTGAAC
GAAAAATCGGAAC

TGCAACTAATTCTTCAACTATTTCTAAAAACAGAAAATATTGATAAAGAATCTCCTGTCGGTAACTTAGCAACA
ACGGCTCAGCTTA

CTATTGCTAAGAAAACCTTTTCCAACGTGTGACTTTGCTATGACCAATAATGGTGGTATTCTGAAGTGACCTAGT
TGTCAAAAATGAC

CGGACCATCACCTGGGGAGCTGCACAGGCTGTACAACCATTTGGTAATATCCTTCAAGTCATTCAAATGACTG
GTCAACACATT

TACGATGTCCTAAATCAGCAATACGATGAAAACCAGACCTATTTTCTTCAAATGTCAGGTTTAACATACACTT
ATACAGATAATGA

TCCTAAGAACTCTGATACCCCTTCAAGATAGTTAAGGTTTATAAAGACAATGGTGAAGAAATTAAGTTAACA
ACTACTTACACCG

TTGTTGTCAACGACTTTCTTTATGGTGGTGGTGATGGCTTTTCAGCATTTAAAAAAGCTAAATTAATCGGAGC
TATTAACACAGAT

ACTGAAGCTTTCATCACATATATCACAAATTTAGAAGCATCAGGTAAAACCTGTTAATGCTACTATAAAAGGGG
TTAAAAATTATGT

AACTTCAAACCTTGAAAAGTTCGACAAAAGTTAATAGTGCTGGTAAACACAGTATCATTAGTAAGGTTTTTAGA
AATCGTGATGGC

AATACAGTGTCTAGTGAAGTCATTTTCAGACCTTTTGACTTCTACTGAAAACACTAATAACAGCCTTGGCAAAA
AAGAAACAACAA

CAAACAAAAATACTATCTCTAGTTCCACTCTTCCAATAACAGGGGACAATTATAAAATGTCTCCTATTATGAC
AATCCTTGCCTTG

ATAAGCTTAGGTGGACTAAACGCTTTTATTAAAAAAGGAAATCCTAG

SPy0895

Seq ID 43

ATGACTAATAATCAAACACTAGACATCCTTTTGGATGTCTATGCTTATAATCACGCCTTTAGAATTGCTAAAG
CCTTGCCAAATAT

CCCTAAAACCTGCCCTCTATTTACTAGAGATGTTAAAAGAGCGCAGAGAATTGAACCTTGCCTTTCTAGCGGAA
CATGCAGCAGA

GAATCGGACCATTGAAGACCAGTATCACTGTTTATTATGGCTTAACCAATCGCTTGAAGATGAGCAGATTGCC
AATTACATTTTG

GATTTAGAAGTTAAAGTAAAAAACGGTGCTATTATTGATTTTCGTCAGGTCAGTGTGCCTATTCTTTACCGAC
TTTTTCTCAGACT

AATCACGTCAGAAATTCAAAACCTTCAAGGCTTATATTTTGTATACAAAGAATGACCAATATGATACCTGGCAT
TTTCAGGCCATGT

TGGAATCTGATCACGAGGTTTTCAAGGCTTACCTGTCTCAAAAGCAGTCTCGCAATGTGACGACCAAAAGCTT
AGCAGACATGT

TGACGTTGACCTCCTTACCTCAGGAAATCAAGGACTTGGTTTTTTTTGTTACGACATTTTGAAAAGGCTGTCCG
TAATCCTCTGGC

TCATTTGATTAAGCCTTTTGATGAAGAGGAACTGCATCGCACCCTCATTTTTCTTCTCAGGCTTTTTTGGA
AACATTATCACCT

TGGCGACTTTTTCTGGTGTAATCTACCGACGTGAGCCTTTTTACTTTGATGACATGAATGCCATTATTAAAA
GGAGTTGAGCCT

TTGGAGACAATCTATTGTCTGA

SPy0972

Seq ID 44

ATGAAGACAACATCCCTGATTAAAGTAGATTTGCCATCAACAATCGGTATAGGTTATGGCGCTTTTTGGCGGT
CTAGAAATTTTT

ATCGAGTAGTTAAAGGCAGCCGTGGATCTAAAAATCTAAAACGACTGCTTTAAATTTATCGTCAGACTGCT
GAAGTACCCTTG

GGCTAACTTATTGGTCATCCGTAGATACTCAAACTAACAACAATCTACTTATACCGATTTTAAATGGGCG
TGTAATCAATTAA

AGGTTACACACCTTTTTAAGTTTAATGAGAGTTTGCCAGAAATAACTGTAAAGGCAACGGGCCAAAAGATACT
GTTCCGTGGACT

TGATGATGAGTTAAAAATCACATCTATTACTGTCGATGTTGGCGCTTTGTGCTGGGCTTGGTTTGAAGAGGCT
TATCAAATTGAG

ACCGAAGATAAGTTTTCAACAGTTGTCGAATCAATCCGCGGTAGTTTAGATGCTCCTGATTTTTTTAAACAGA
TAACAGTCACGTT

TAACCCGTGGTCAGAAAGACATTGGCTTAAACGTGTCTTTTTTGATGAAGAACTAAACGGGCTGATACATTT
TCTGGGACTACA

ACATTTAGAGTAAACGAATGGCTTGATGATGTCGATAAAAGACGCTACGAAGATTTGTACAAGACTAATCCAA
GGCGGGCTAGA

ATCGTGTCGATGGTGAATGGGGCGTTGCTGAAGGTCTTGTTTTTGATAACTTTGAAGTCGTAGATTTTGATG
TTGAAAAACAA

TTCAACGCGTTAAAGAGACCTCGGCCGGTATGGACTTTGGGTTTACTCAAGACCCTACAACCTTTATATGTGT
TGCAGTTGACCT

CGCAAACAAAGAGTTATGGCTTTACAACGAACATTATCAAAAGGCTATGTTAACAGATCATATTGTCAAAATG
ATAAGAGATAAAA

ACTTGCATAGGTCTTACATCGCAGGGGATAGCGCCGAAAAACGCCTCATTCAGAAAATAAAAAGTAAAGGGGT
GTCTGGAATTG

TCCCGAGTATTAAAGGTAAAGGGTCAATCATGCAAGGGATTCAATTCATGCAGGGGTTAAGATATATATTCA
CCCATCTTGCGA

ACACACAATAGAAGAGTTTAATACTTACACTTTTAAGCAAGACAAAGAAGGTAATTGGTTAAACGAACCGATA

GATAAGAATAAC

CACGTTATTGATGCGATTAGATATGCGCTTGAAAAATACCATATCAGAAGCAACGAGTCAAATCAGTTTGAAG
TTCTTAGGGCTG
GTTTTGGTTACTAG

SPy0981

Seq ID 45

ATGGCAGAAGAAACACAAACAGTTGAAACGGTTGAAGAGCAAGTGGTACCAGAAGCAAAACAACCGCAAGACG
AAAAAAAGTA
CACAGATGCAGATGTGGACGCTATCATCGACAAAAAGTTTGCGAAGTGGAAAGTCAGAACAAGAAGCGGAGAAA
TCGGAAGCTA
AAAAAATGGCTAAGATGAATGAAAAAGAGAAAGCAGACTACGAAAAGCAGAAGCTGTTAGACGAATTGCAAGA
GCTAAAAAACG
ATAAGACACGCAATGAGTTAACAGCAGTAGCTCGTCAAATGTTTGCAGAATCTGAAATCAACGTCAACGATGA
CGTACTTGGTTT
AGTTGTGACTTTGGACGCAGAACAAACAAAAGCAAATGTAACAACGCTAGCAAACGCATTTGCTAAAGTTATC
GCTGATGACCG
CAAGGCTCTTGACGCCAGACTACTCCGTCAACAGGTGGTGGATTGAGCAAACAAACCAATTACGGTGCTAAC
TTGGCTAGTAA
GGCAGCACAACAAAGCACCAAACCTTTTTTAG

SPy1008

Seq ID 46

ATGAGATATAATTGTCGCTACTCACATATTGATAAGAAAATCTACAGCATGATTATATGTTTGTCAATTTCTTT
TATATTCCAATGTT
GTTCAAGCAAATTCTTATAATACAACCAATAGACATAATCTAGAATCGCTTTATAAGCATGATTCTAACTTGA
TTGAAGCCGATAG
TATAAAAAATTCTCCAGATATTGTAACAAGCCATATGTTGAAATATAGTGTCAAGGATAAAAAATTTGTCAGTT
TTTTTTGAGAAAGA
TTGGATATCACAGGAATTCAAAGATAAAGAAGTAGATATTTATGCTCTATCTGCACAAGAGGTTTGTGAATGT
CCAGGGAAAAGG
TATGAAGCGTTTGGTGAATTACATTAATAATTCAGAAAAAAAAGAAATTAAAGTTCCTGTAAACGTGTGGG
ATAAAAGTAAACA
ACAGCCGCCTATGTTTATTACAGTCAATAAACCGAAAAGTAACCGCTCAGGAAGTGGATATAAAAGTTAGAAAG
TTATTGATTAAG
AAATACGATATCTATAATAACCGGGAACAAAAATACTCTAAAGGAACTGTTACCTTAGATTTAAATTCAGGTA
AAGATATTGTTTTT
GATTTGTATTATTTTGGCAATGGAGACTTTAATAGCATGCTAAAAATATATTCCAATAACGAGAGAATAGACT
CAACTCAATTTCA

TGTAGATGTGTCAATCAGCTAA

SPy1032

Seq ID 47

GTGAATACTTATTTTTGCACACACCATAAAACAATTACTACTTTATTCAAACCTATTCCTTAGCTTTGCTATGA
TGGGCCAAGGAAC

TGCCATTTATGCCGATACACTGACTTCAAATTCAGAACCTAATAATACTTACTTTCAAACGCAAACGCTCACT
ACTACAGATAGCG

AAAAAAGGTTAGTACAGCCACAACAAAAAGACTACTATACTGAATTGTTAGACCAATGGAACAGTATTATCGC
AGGCAACGATGC

TTATGATAAAACCAATCCTGACATGGTCACTTTTCATAATAAAGCTGAAAAGGATGCTCAAACATTATTAAA
AGCTATCAAGGGC

CTGACCACGAAAAATAGAACTTACCTTTGGGAACATGCAAAGGATTATTCGGCTTCTGCTAATATCACGAAAAC
TTACCGCAATAT

TGAAAAAATAGCAAAACAGATCACTAATCCTGAATCATGCTATTATCAAGATAGTAAAGCTATTGCTATTGTA
AAAGACGGTATGG

CCTTCATGTATGAACACGCTTATAATCTAGATCGTGAAAATCATCAAACAACCTGGAAAAGAAAACAAAGAAAA
TTGGTGGGTTTA

TGAAATTGGAACCTCCTCGTGCTATTAATAATACCTTATCCTTGATGTATCCTTATTTTACTCAAGAAGAAATT
CTTAAATACACAGC

TCCAATCGAAAAATTTGTGCCTGACCCTACTCGTTTTAGGGTTCGCGCTGCCAATTTTTACCTTTTGAAGCC
AATAGCGGAAAT

TTAATTGATATGGGGCGTGTTAAACTCATTTCCGGTATTCTTCGTAAAGATGATCTCGAAATTAGTGATACAA
TCAAAGCAATTGA

GAAAGTTTTACGCTAGTTGATGAAGGAAATGGTTTTTACCAAGACGGTTCTTTAATTGATCACGTGGTTACT
AACGCTCAAAGT

CCACTTTATAAAAAAGGCATTGCTTACACTGGAGCTTACGGTAATGTGCTTATAGATGGCTTATCGCAATTAA
TTCCTATTATTCA

AAAAACAAAGTCTCCTATAAAAGCGGATAAAATGGCTACTATCTATCATTTGGATTAACCATTCCTTTTTTCCCT
ATCATCGTTCGTG

GAGAAATGATGGATATGACTCGAGGGCGTTCTATCAGTCGTTTTAATGCCCAATCTCATGTTGCTGGCATTGA
AGCACTTCGTG

CTATTTTACGTATTGCTGACATGTCTGAAGAGCCTCACCGTTTGGCACTTAAAACACGTATAAAAACACTCGT
CACACAAGGGAA

TGCTTTTTACAATGTCTATGATAATTTGAAAACCTATCACGATATCAAACCTTATGAAAGAACTACTAAGTGAT
ACTTCTGTTCCAGT

CCAAAACTTGATAGTTACGTAGCTAGTTTCAATAGTATGGATAAATTGGCACTATATAATAATAAACACGAT
TTTGCTTTTGGCC

TATCAATGTTTTCGAATCGAACTCAAAATTATGAAGCTATGAATAATGAAAATCTTCATGGCTGGTTTACTTC

TGATGGAATGTTTT

ACCTATACAATAACGATTTAGGACACTACAGTGAAAACTATTGGGCAACGGTAAATCCCTACCGCTTACCTGG
AACCACAGAAA

CTGAGCAAAAACCACTAGAGGGAACTCCTGAGAATATTTAAACGAACTATCAACAAGTTGGCATGACTGGTCT
CTCTGATGACG

CTTTTGTTGCAAGTAAAAAACTTAATAACACAAGTGCTCTAGCTGCTATGACCTTCACTAATTGGAATAAAAG
CCTCACCCCTCAAT

AAAGGGTGGTTTATCTTAGGAAACAAAATAATCTTTGTTGGTAGCAATATCAAAAACCAATCATCTCACAAGG
CGTATACAACCTAT

TGAACAACGAAAAGAAAATCAAAAGTACCCTTACTGTTCTTATGTTAACAATCAACCCGTTGACTTGAATAAT
CAGCTAGTTGATT

TTACAAACACTAAAAGTATTTTCCTTGAAAGTGATGATCCCGCTCAAAATATTGGTTACTACTTCTTCAAGCC
AACAACACTTAGC

ATAAGTAAGGCACTTCAAACAGGGAAATGGCAAAACATAAAAGCTGATGACAAATCACCAGAAGCCATCAAAG
AAGTTTCAAATA

CCTTTATCACTATCATGCAAAACCATACTCAAGATGGCGATCGTTATGCCTATATGATGCTTCCAAATATGAC
TCGTCAAGAATTT

GAAACCTATATTAGCAAGCTTGATATCGACTTATTAGAAAACAATGACAAACTGGCCGCTGTCTACGATCATG
ATAGTCAACAGA

TGCACGTCAATCACTATGGAAAAAAGCAACGATGTTTTCAAATCATAATCTTTCTCATCAAGGCTTTTATAG
TTTTCTCATCCT

GTCAGGCAAAATCAACAATAA

SPy1054

Seq ID 48

TTGCTGACCTTTGGAGGTGCAAGTGCGGTTAAGGCGGAAGAAAATGAAAAAGTAAGAGAGCAAGAAAAGCTCA
TACAGCAACTT

TCTGAAAAGCTAGTGGAATTAATGACTTACAACTTTAAATGGTGATAAAGAGAGTATACAGTCTCTCGTAG
ATTATCTGACTCG

AAGAGGAAAACCTGAAGAAGAATGGATGGAATATTTGAATTCTGGTATTCAACGCAAACTTTTGTTGGTCCA
AAAGGACCTGCA

GGTGAAAAAGGAGAACAAGGTCTACTGGAAAAACAAGGCGAGCGTGGTGAGACCGGCCCTGCAGGTCCACGTG
GTGACAAGG

GCGAACTGGTGACAAAGGAGCCCAGGGTCCAGTAGGTCCCGCTGGCAAGGACGGCCAAAACGGTAAAGATGG
TCTTCCAGG

TAAAGACGGCAAGGACGGCCAAAACGGTAAAGATGGTCTTCCAGGTAAAGACGGCAAGGACGGCCAAGACGGT
AAAGATGGC

CTCCCAGGTAAAGACGGTAAGGATGGCCAAAATGGCAAAGATGGTCTTCCAGGTAAAGACGGTCAACCAGGTA
AACCAGCTCC

TAAACACCAGAGGTCCCTCAAAACCCAGATACTGCACCACATACTCCAAAAACCCCTCGGATCCCTGGTCAA
TCAAAAGACGT

GACACCTGCTCCTCAAAACCCCTTCTAATAGAGGTCTAAACAAACCACAAACACAAGGTGGTAATCAGCTCGCA
AAAACACCGGC

AGCTCACGACACACAGACAATTGCCAGCAACAGGCGAAACAACCAATCCATTCTTTACAGCAGCTGCTGTA
GCTATCATGAC

GACAGCTGGAGTTGTAGCTGTTGCAAAACGTCAAGAAAACAACTAA

SPy1063

Seq ID 49

ATGTATATATTCTCATCGTCAAAAAAAGATAGTGCTAAAGAATTAGTTATCTTGACTCCTAATAGCCAAACTA
TTTTAACAGGGAC

TATTCCAGCCTTTGAGGAAAAGTATGGGGTTAAAGTAAGATTAATCCAAGGTGGGACGGGCCAACTTATTGAT
CAATTAGGTCG

AAAAGATAAAACCATTAACGCTGATATTTCTTTGGTGGCAATTACACTCAATTTGAAAGCCATAAAGATTTA
TTTGAATCTTATGT

TTCTCCGAGGTTTCTACTGTCATTTAGATTACCAATTGCCTAGTCATCGCGCAACCCCATATACGATCAAT
GGCAGTGTACTG

ATTGTTAATAACGAATTAGCAAGAGGACTTCATATTACCAGTTATGAGGATTTGCTACAACCAGCTTTAAAAG
GCAAAATTGCTTT

TGCTGATCCCAACAGTTCATCAAGTGCCTTCTCACAGCTGACTAATATATTGTTAGCTAAGGGGGGGTACACA
AACGCTGACGC

TTGGGCTTACATGAAGCGCTTGTTGGTCAATATGAATTCTATTAGGGCTACGAGTTCTTCAGAAGTCTATCAA
TCTGTCGCTGAG

GGTAAGATGATTGTTGGGCTAACCTACGAAGATCCTTGTATCAACCTGCAAAAAAGTGGTGCCAATGTTTCCA
TTGTTTATCCAA

AAGAAGGAACGGTGTTTGTGCCCTCCTCTGTTGCTATTATCAAACATGCGCCAAACATGACAGAGGCTAAGCT
CTTTATTAATTT

TATGTTATCACGTGATGTGCAAAATGCCTTTGGCCAATCAACCAGTAACCGACCCATTCGTCAAGATGCCCAA
ACCAGTCACGA

CATGAAAGCCTTAGAAACGATAGCTACTTTGAAAGAGGATTATGCTTATGTTACCAAGCACAAGAAAAAATA
GTGGCTACGTAC

AACCAGTTGCGCCAACGGTTGGAAAAAGCTAAGTAG

SPy1162

Seq ID 50

ATGCCGACTAGTATTAAAGCTATTAAAGAAAGCTTAGAGGCCGTTACTAGCCTCTTGGACCCCTCTTTCAAG
AATTGGCAACC

GACACTAGGTGAGGCGTCCAAAAAGCTCTAAAAAGCCGACAAAAGGTTATTCAGGCCGAGTTAGCAGAAGAAG

AACGATTAGA

AGCCATGCTTTCTTATGAAAAAGCTCTTTATAAAAAAGGTTATAAAGCCATTGCAGGTATTGATGAGGTGGGA
CGTGGTCCCTTA

GCAGGTCCC GTTGTGGCAGCTTGTGTGATTTTACCTAAGTATTGTAAAATTAAAGGCCTTAATGATTCTAAAA
AAATCCCTAAAG

CTAAGCATGAGACCATTTATCAGGCAGTGAAAGAAAAGGCTTTGGCTATCGGTATCGGTATTATTGACAATCA
GCTTATTGATGA

GGTCAATATTTATGAAGCAACCAAACTGGCCATGCTAGAAGCCATTAAACAGTTGGAGGGCCAACTCACACAA
CCAGATTATCT

CTTGATTGATGCCATGACATTGGATATTGCTATTTGCGCAGCAGTCTATTCTTAAAGGCGATGCCAATTCCTTG
TCTATTGCAGCA

GCATCAATTGTAGCTAAGGTCACCAGAGATCAGATGATGGCTAACTATGATCGCATTTTTCTGTTTATGACT
TTGCTAAAAATG

CAGGCTATGGCACCAAGAACATTTACAGGGATTAAAAGCTTACGGCATAACGCCTATCCATCGTAAAAGTTT
TGAACCTGTTAA

ATCCATGTGCTGCGATTCAACTAATCCTTAA

SPy1206

Seq ID 51

ATGACAGTTAAGGAAGAAACGATGAGTATTTAGAAAGTTAAGCAGTTGAGTCACGGTTTTGGGGATCGGGCTA
TTTTTGAAAATG

TGTCATTTGCGCTCTTAAAAGGCGAACATATTGGACTAGTTGGGGCAAATGGTGAAGGAAAATCAACCTTTAT
GAGTATAGTCAC

AGGACATTTACAGCCTGACGAAGGAAAAGTAGAGTGGTCGAAGTATGTCAGTGCAGGTTACCTGGATCAACAT
ACAGTGTTGGA

ATCAGGACAAACCGTTCGTGATGTCTTGCGAACTGCTTTTGATGAGTTATTTAAGACCGAGAATCGTATTAAT
GAGATTTACGCG

TCAATGGCAGATGATAAAGCTGATATTGCTGTTTTGATGGAAGAAGTAGGTGAGCTTCAAGATCGTTTAGAAA
GTCGTGATTTCT

ATACTTTGGATGCTAAGATTGATGAAGTAGCGCGTGCGCTTGGTGTTATGGATTTTGAATGGAGTCAGATGT
TACATCCTTATC

AGGTGGGCAACGAACAAAGGTTTTATTAGCCAAATTACTATTAGAAAAACCTGATATTCTGCTATTAGATGAA
CCAACTAACCATT

TGGATGCTGAGCATATTGAATGGTTAAAACGCTATTTACAACATTATGAAAATGCTTTTGTGTTGATTTGCA
TGATATTTCTTTCT

TAAATGATGTGATTAATATTGTTTATCATGTTGAAAAATCAAAGTTTAGTTCGCTATACTGGGGATTATTACCA
ATTTCAAGCTGTTT

ATGAGATGAAACAATCTCAACTTGAAGCAGCCTATGAACGTCAACAAAAAGAGATTGCTAACTTGCAGGATTT
TGTCAACCGAAA

TAAGGCTCGGGTAGCGACACGTAACATGGCAATGTCTCGCCAAAAAAAACCTTGATAAGATGGATATAATTGAA

CTTCAAGCTGA

GAAACCAAAACCAATTTTGAATTTAAGCAAGCTAGAACTCCCAGTCGATTCATTTTTCAAACAAAAAATCTT
GTGATTGGTTATG

ATTACCCATTGACCAAGAACCCTTAAATATAACGTTTGAAAGAAATCAAAAAATTGCTATTGTTGGGGCCAA
CGGTATTGGA

ATCTACTTTGCTAAAAAGTTTATTAGGTGTTATTGAGCCTTTAGAAGGTCATATTGTCACAGGGGATTTTTTA
GAAGTTGGCTACT

TTGAACAAGAAGTGACAGGTGTTAACCGACAAACTCCGCTAGAAGTAGTTTGGGATGCTTTTCCTGCCTTAA
TCAGGCAGAAG

TTCGAGCGGCACTAGCTCGTTGCGGACTAACATCAAAACATATCGAAAGTCAAATTCAGTACTTTCGGGTGG
TGAACAAGCAA

AAGTTCGTTTTTGTGTTGATGAATCGTGAAAATAACGTGCTTATTTTAGACGAACCAACAAATCATCTTGA
TATTGATGCTAAA

AATGAGCTCAAACGTGCTTTAAAAGCATATAAGGGTCTATTTTAATGGTTTGTGATGAACCTGATTTCTACA
ATGGGTGGGTAA

CCGATACTTGGGATTTTAGTAAGTTAACCTAA

SPy1228

Seq ID 52

ATGAACAAGAAATTTATTGGTCTTGGTTTAGCGTCAGTGGCTGTGCTGAGTTTAGCTGCTGTGGTAATCGTG
GTGCTTCTAAAG

GTGGGGCATCAGGAAAACTGATTTAAAAGTTGCAATGGTTACCGATACTGGTGGTGTAGATGACAAATCATT
CAACCAATCAG

CATGGGAAGGCCTGCAATCTTGGGGTAAAGAAATGGGCCTTCAAAAAGGAACAGGTTTCGATTATTTTCAATC
TACAAGTGAAT

CTGAGTATGCAACTAATCTCGATACAGCAGTTTCAGGAGGGTATCAACTGATTTATGGTATCGGCTTTGCATT
GAAAGATGCTAT

TGCTAAAGCAGCTGGAGATAATGAAGGAGTTAAGTTTGTATTATCGATGATATTATCGAAGGAAAAGATAAT
GTAGCCAGTGTT

ACCTTTGCCGACCATGAAGCTGCTTATCTTGCAAGGAATTGCAGCTGCAAAAACAACAAAAACAAAAACAGTTG
GTTTCGTGGGC

GGTATGGAAGGAACTGTCATAACTCGATTTGAAAAAGGTTTGAAGCAGGAGTTAAGTCTGTTGACGATACAA
TCCAAGTTAAAG

TTGATTATGCTGGATCATTTGGTGACGCTGCAAAAAGGAAAAACAATCGCAGCAGCTCAGTATGCAGCAGGTGC
TGATGTTATTT

ACCAGGCAGCAGGAGGCACTGGAGCAGGTGTATTTAATGAAGCAAAAGCTATTAATGAAAAACGTAGTGAAGC
TGATAAAGTTT

GGGTTATTGGTGTGACCGTGATCAAAAAGACGAAGGAAAAATACACTTCTAAAGATGGCAAAGAAGCAAACTT
TGTA

ATCAATCAAAGAAGTCGGTAAAGCTGTTTCAGTTAATCAACAAGCAAGTAGCAGATAAAAAATTCCCTGGAGGA
AAAACAACGTGC

TATGGTCTAAAAGATGGCGGTGTTGAAATCGCAACTACAAATGTTTCAAAAGAAGCTGTAAAGCTATTAAAG
AAGCGAAAGCAA

AAATTAAATCTGGTGACATTAAAGTTCCTGAAAAATAG

SPy1245

Seq ID 53

ATGAAAATGAAAAAAAATTCTTTTTGTTAAGTCTTTGGCCCTATCAACTTTCTTTTTATCCGCATGTTCTA
GCTGGATTGATAAA

GGTGAGTCAATAACCGCTGTAGGATCAACAGCACTACAACCCTTAGTAGAAGCAGTAGCTGATGAATTTGGAA
GCAGTAATCTA

GGCAAGACTGTCAATGTTCAAGGTGGTGGTTCAGGTACAGGGTTGTCTCAAGTTCAATCAGGAGCTGTCCAAA
TTGGAAATAGT

GATGTCTTTGCGGAAGAAAAAGATGGTATTGATGCTTCTAAATTAGTTGATCATCAAGTAGCTGTTGCAGGAC
TTGCAGTTATTG

CCAATCCTAAAGTCAAGGTTTCCAATCTCAGTAGTCAGCAGTTGCAAAAGATTTTTTCAGGAGAATATACCAA
TTGGAAACAAGT

TGGAGGAGAAGATCTTGCGATTTTCAGTGATCAACCGAGCAGCAAGTTCTGGCTCACGAGCAACCTTTGACAGT
GTTATCATGAA

AGGGGTCAACGCTAAACAAAGTCAAGAGCAAGACTCCAATGGGA TGGTTAAATCGATTGTTTCACAAACACCA
GGTGCCATTTT

TTACCTTTTCTTTGCCTACGTTGATTTCATCTGTAAATCTTTGCAATTAAATGGGTTTAAGGCAAATGCTAAG
AACGTGGCTACAA

ATGATTGGCCAATCTGGTCCTACGAACACATGTATACCAAAGATAAACCAACAGGGTTGACCAAGGAATTTCT
TGATTATATGTT

TTCAGATGAAGTACAACAGAACATTGTTACACATATGGGATATATTTTCGATAAATGATATGGAAGTGGTCAAA
TCTCATGATGGA

AAAGTAACAAAAAGGTAA

SPy1315

Seq ID 54

ATGACGCACAAAATAAAAGTATTGCTGCTTGCGATAATGTCTATTTTTTTGACATGCAATATTGCAAGTGCTG
AAACTATTGCTAT

TGTTTTAGATACAGCTTATGCCCCATTTGAATTTAAAGACTCAGATCAAATTTACAAAGGAATTGACGTTGAT
ATTATTAATGAAG

TAGCCAAACGTCAATCTTGGGATTTTCAGTATGAGTTTCCCGGGTTTTGATGCAGCTGTAAATGCTGTTCAATC
TGGTCAAGCGA

GTGCTCTAATGGCCGGTACAACCATTACGAATGCTCGTAAGAAAGTCTTTCATTTCTCAGAGCCATATTACGA

TACCAAAATTGT

CATTGCGACACGTAAAGCCAATGCCATCAAAAAATACAGTGACTTAAAAGGAAAAACGGTCGGTGTTAAAAAT
GGAACAGCGGC

TCAAGCCTTTTTGAATAACTATAAAAAAAGTATGATTATACTGTAAAAACATTTGACACAGGTGATCTTATG
TATAATAGTTTATC

TGCTGGTTCTATTGCCGCTGTTATGGATGATGAGGCGGTTATCCAATACGCAATCAGCCAAAACCAAGATATT
GCTATTAACATG

AAAGGAGAGCCCATTTGGAAGCTTTGGGTTTGCTGTCAAAAAGGGAAGCGGATATGATTATCTAGTTAATGATT
TCAATACAGCTC

TTAAAGCTATGAAAGCTGATGGTACCTACCAAGCTATCATGACCAAGTGGTTAGGCACAGATGATAAAGCTAC
CACCAGTCAGG

CAACGGGAAATCCATCTGCCAAAGCTACACCTACAAAGGACAGTTATAAAATTGTCTCTGATTCGTCTTTTGC
ACCGTTTGAATT

TCAAAATGGTAAGGGCAAATACGTTGGTATTGACATAGAATTAATCAAAGCTATTGCTAAACAACAAGGTTTC
AAAATTGAAATCG

CTAATCCAGGTTTCGATGCTGCCTTAAATGCTGTGCAATCTAGCCAAGCAGATGGGGTCATTGCTGGTGCAAC
TATTACTGACG

CTCGTAAAGCTATCTTTGATTTTTCTGATCCTTATTATACTTCTAATATCATTTTAGCTGTAAAGCTGGAAA
AAACATCAAGAACT

ATGAAGACTTAGACAGAAAAACAGTCGGTGCTAAAAACGGCACTTCATCTTACTCTTGGTTAAAAGAAAACGC
TCCTAAATATGG

TTATAATGTCAAGGCATTTGATGATGGTTCTAGCATGTATGATAGCTTAAATTCAGGTTCTGTAGATGCTATC
ATGGATGATGAG

GCGGTTCTTAAATACGCTATCTCTCAAGGTCGTCGCTTTGAAACACCTCTTGAGGGCATTCTACTGGTGAAG
TTGGTTTTGCTG

TCAAGAAAGGAACTAATCCAGAATTAATCGAAATGTTCAACAAATGGCTTAGCTGCTCTCAAAAAATCTGGTCA
GTATGATGACAT

TATAGATAAATACCTTGACTCTAAGAAAGCTGCAACTCCTTCTGAAAAAGGTGCTGATGAGTCTACTATTTCA
GGCCTATTATCAA

ATAACTACAAACAACTATTGGCAGGACTTGGAACCACGCTCAGTTTAACCCTTATTTCAATTTGCTATTGCTAT
AATTATCGGGATC

ATCTTTGGGATGATGGCCGTGTCACCAACTAAATCACTTCGACTTATTTCAACGGTCTTTGTGGACGTTGTTC
GAGGGATTCCTT

TGATGATTGTGGCTGCCTTCATTTTCTGGGGAGTACCAAACCTTATCGAGAGTATGACCGGCCACCAGTCACC
GATTAATGATT

TCTTAGCTGCTACAATTGCACTGTCACTTAATGGCGGAGCCTATATTGCTGAAATTGTTTCGCGGTGGTATCGA
AGCTGTTCCAG

CAGGGCAAATGGAAGCTAGTCGAAGTCTTGGTTTGTCTTACGGAACCACGATGAGAAAAGTAATTCTCCCACA
AGCTGTGAAAC

TAATGTTACCTAACTTTATCAATCAGTTTGTATTTCATTGAAGGATACAACAATCGTCTCAGCAATTGGTTT
AGTGGAACCTCTC

CAAACAGGTAAATCATTATTGCTAGAAATTACCAGTCGTTCCGTATGTATGCTATTTTAGCAATTATTTACC
TTATCATGATTATA

CTCTTAACAAGACTTGCAAAACGTTTAGAAAAGAGGCTTAACTAA

SPy1357

Seq ID 55

ATGGGAAAAGAAATAAAAGTGAAATGCTTTTTGCGTAGATCAGCTTTTGGATTAGTTGCGGTGTCAGCATCAG
TATTAGTCGGTT

CAACAGTATCTGCTGTTGACTCACCTATCGAACAGCCTCGAATTATTCCAAATGGCGGAACCTTAACTAATCT
TCTTGGAATGC

TCCAGAAAACTGGCATTACGTAATGAAGAAAGAGCCATTGATGAATTAACCAACAAGCTATTGAGGATAAA
GAAGCTACGACA

GCTATAGAAGCAGCAAGTTCAGATGCCTTAGAAGCATTAGCGGATCAAACAGACGCTTTACAATCAGAAGAAG
CTGCGGTTGTT

AAAGCGGATAACGCTGCTAGTGACGCCTTAGAAGCATTGGCGGATCAAACAGACGCTTTACAATCAGAAGAAG
CTGAAGTAGTT

CAATCAGATAACGCTGCTAGTGACGCCTGGGAAAAAGCAGCAACTCCAATCGCTTTAGATGTTAAGAAAACTA
AAGATACAAAA

CCTGTAGTTAAAAAGAAGAAAGACAAAACGTTAATACCCTTCCTACAACCTGGTGAAGAGTCTAACCCATTCT
TTACAGCTGCTG

CGCTTGCAATAATGGTAAGTACAGGTGTGTTAGTTGTAAGTTCAAAGTGCAAAGAAAATTAG

SPy1361

Seq ID 56

ATGAAAACGAAAAAGTTATTATTTTAGTTGGTCTATTGTTATCATCTCAGTTGACTTTGATAGCTTGTCAAT
CACGAGGTAATGG

TACATATCCCATTAACGAAACAATCACGTAAGGGAATGACGTCAAACAAAATTAAACCGATTAAAAAAGC
AAAAAGACAAAC

AAGACTCACAAAGGTGTGGCGGGTGTGATTTTCTACAGATGATGGGTTTATTTTAACCAAAGACTCAAAAA
TCTTATCAAAAA

CAGATCAGGGAATCGTTGTTGACCATGATGGTCATTTCGATTTTATTTTTATGCCGATTAAAGGGAAGTCC
ATTTGAATACCTT

ATTCCAAAAGGAGCAAGTTTAGCTAAGCCAGCTGTTGCTCAGCGAGCAGCTAGTCAAGGGACTTCTAAAGTAG
CAGATCCTCAT

CACCATTATGAATTTAACCAGCGGATATTGTGGCTGAAGATGCTTTAGGCTACACGGTTCGCCACGATGATC
ACTTCCATTATA

TTTTGAAGTCAAGCTTATCAGGTGAGACACAGGCACAAGCTAAACAGGTTGCTACTCGCTTGCCACAAACCAG

TAGCCTTGTTT

CAACAGCTACAGCTAATGGTATTCCAGGCTTGCATTTCCCAACCTCAGATGGTTTTCAATTTAACGGTCAAGG
TATTGTTGGGGT

AACAAAAGACAGTATTTTAGTGACCACGATGGTCACTTACATCCTATTTCTTTTGC GGACCTTCGTCAGGGT
GGCTGGGCACA

TGTGGCAGATCAATACGATCCCGCTAAAAAAGCAGAAAAGCCAGCAGAAACCCATCAGACACCAGAGCTATCT
GAACGTGAAAA

GGAATACCAAGAAAAATTAGCTTATTTGGCAGAAAAATTGGGGATTGATCCATCAACTATTAAACGTGTGGAA
ACACAAGACGGT

AAACTTGTTTTGGAATACCCTCACCATGACCACGCACACGTATTGATGTTATCTGATATTGAAATCGGAAAAG
ACATTCCAGATC

CACATGCTATTGAGCATGCCCGTGAATTGGAAAAACATAAGGTTGGAATGGATACCTTGCGTGCCTTAGGGTT
TGATGAAGAAG

TGATTTTGGATATCGTTCGCACTCACGATGCTCCAACCCCATTCCCATCAAATGAAAAAGATCCGAATATGAT
GAAAGAATGGTT

AGCAACGGTTATCAAACCTTGACTTGGGCAGCCGTAAAGATCCTTTGCAACGTAAAGGACTTTCAGTGTACCC
AACTTAGAAACT

TTAGGAATTGGCTTTACACCAATCAAAGATATCTCACCTGTTTTGCAATTTAAAAAATTGAAACAGTTGTAA
TGACAAAAACAGG

GGTGACTGATTATAGATTTTTTGGATAATATGCCACAGTTAGAAGGCATTGATATTCACAAAACAATCTCAAA
GATATTAGTTTTCT

TGAGCAAATATAAAAACTTAACTCTAGTAGCGGCTGCTGATAATGGTATTGAAGATATTAGGCCGCTTGGTCA
ATTACCAAATCT

CAAATTCCTCGTATTGAGTAACAATAAGATTTCTGATTTAAGCCCACTGGCATCGTTACATCAATTGCAAGAA
TTGCACATTGATA

ATAATCAGATTACAGATTTAAGCCCTGTTTCTCATAAAGAATCATTGACGGTTGTTGATTTATCAAGAAATGC
TGATGTTGACTTA

GCAACACTTCAAGCACCCAAATTAGAAACGTTAATGGTCAATGATACCAAGGTTTCTCATTGGATTTCTTGA
AAAATAATCCTAA

TCTATCTAGCCTATCTATTAACCGTGCGCAATTGCAATCTCTTGAAGGTATTGAAGCAAGTAGCGTCATTGTC
AGAGTAGAAGCA

GAAGGTAACCAAATTAATCGCTTGTGCTTAAAGACAAGCAAGGGTCACTTACTTTCTTGGATGTGACAGGCA
ACCAGTTGACTT

CTCTAGAAGGTGTTAATAATTTTACAGCACTTGACATTTTAAGCGTGTCTAAAAACCAATTAACAAATGTCAA
CCTATCTAAACCC

AATAAGACAGTTACTAACATTGATATTAGTCATAACAATATCTCATTAGCAGACCTTAAATTGAACGAGCAAC
ATATTCCAGAAGC

CATTGCGAAAACTTCCCAGCGGTTTACGAAGGTTCTATGGTAGGTAATGGAACAGCTGAAGAAAAAGCAGCT
ATGGCTACTAA

GGCGAAAGAAAGTGCTCAAGAAGCATCGGAATCACATGACTACAACCATAATCATACCTATGAAGATGAAGAA
GGTCATGCTCA

CGAGCACAGAGACAAAGATGATCACGACCATGAACATGAGGATGAAAATGAAGCTAAAGATGAGCAAAACCAT
GCTGACTAA

SPy1371

Seq ID 57

TTGGCAAAACAATATAAAAAATTTAGTGAACGGTGAATGGAACTATCAGAAAACGAGATTACCATTTACGCAC
CAGCAACAGGTG

AAGAGTTAGGATCAGTTCCAGCGATGACGCAGGCAGAGGTAGATGCTGTTTACGCTTCAGCTAAAAAGGCTCT
ATCAGATTGGC

GCGCTTTGTCTTATGTGGAACGTGCAGCTTACCTTCATAAAGCGGCTGATATTTTAGTACGTGATGCTGAAAA
GATCGGCGCGA

TTCTTTCAAAAGAAGTAGCCAAAGGTCACAAGGCAGCTGTCAGTGAAGTTATTCGTACCGCTGAAATCATTAA
TTATGCAGCAGA

AGAAGGGCTTCGTATGGAAGGTGAAGTTCTTGAAGGTGGTAGCTTCGAAGCTGCAAGTAAGAAGAAGATTGCT
ATTGTTCTGTCG

TGAACCAGTTGGTTTAGTTCTTGCCATCTCACCTTTTAATTATCCCGTTAACTTGGCAGGTTCTAAAATTGCT
CCAGCTCTTATTG

CAGGAAATGTTGTTGCTCTTAAACCACCAACACAAGGCTCTATTTCTGGTTTGTACTAGCAGAAGCTTTTGC
AGAAGCTGGTAT

TCCAGCAGGTGTCTTTAATACCATTACAGGGCGAGGTTCTGTTATCGGTGATTATATCGTTGAGCACGAAGCG
GTTAGCTTTATC

AACTTTACAGGTTCTACTCCAATTGGGGAAGGAATCGGTAAATTAGCGGGTATGCGACCAATTATGCTTGAGC
TTGGCGGTAAG

GATTCTGCTATCGTTTTGGAAGATGCAGATCTTGCTTTAGCAGCGAAAAATATTGTAGCCGGTGCTTTTGTT
ACTCAGGCCAAC

GTTGTACAGCGTTAAACGTGTTCTTGTGATGGACAAGGTGGCGGATCAATTGGCGGCTGAGATTAACAACT
TGTTGAAAAAC

TAAGTGTGGAATGCCTGAAGACGATGCTGATATTACCCATTAATTGATACATCAGCTGCTGATTTTGTTGA
AGGGTTGATTAA

AGATGCAACTGATAAGGGAGCTACTGCTTTGACAGCCTTTAATCGTGAAGGCAATCTTATTTACCCGTTCTC
TTTGATCATGTG

ACAACTGACATGCGTTTGGCATGGGAAGAGCCGTTTCGGCCCAGTATTACCAATTATTCGTGTAACCACTGTAG
AAGAAGCCATC

AAGATTTCTAATGAGTCTGAATATGGTTTGCAAGCTTCTATTTTACAACATAATTTCCCAAAAGCTTTTGCA
TTGCTGAGCAATT

AGAAGTTGGAAGTGTTCACCTTAACAATAAAACACAACGTGGAACAGATAATTTCCCATTCCTTAGGCGCTAAA
AAATCAGGTGCA

GGGGTACAAGGAGTTAAATATTCTATCGAAGCTATGACAACTGTTAAATCTGTTGTATTTGATATCCAGTAA

SPy1375
Seq ID 58

ATGAGTCTCAAAGATCTTGGCGATATTTTCATATTTTCGCCTAAATAATGAAATTAACCGTCCTGTTAATGGTA
AAATTCCACTTCA

TAAAGACAAAGAAGCTTTAAAAGCTTTTTCCGCTGAAAAATGTGCTGCCAAACACCATGTCTTTTACTTCCATT
ACGGAAAAAATTG

AGTATTTAATCTCAAATGATTACATTGAATCAGCTTTTATTCAGAAATACCGCCCTGAATTTATTACTGAATT
AGATAGCATAATCA

AATCAGAAAATTTTCGCTTTAAATCATTTATGGCAGCCTACAAGTTCTACCAGCAATACGCCTTAAAAACAAA
TGATGGAGAGCAT

TATTTAGAAAACCTTGAAGACCGTGCTTGTTTAATGCTTTGTATTTTGCAGATGGTCAAGAAGACTTAGCAA
AAGATTTAGCCGT

TGAAATGATTAACCAACGTTACCAACCGGCTACTCCTTCCTTTTTAAATGCTGGTCGAAGCCGTCGTGGTGAA
TTGGTCTCTGT

TTCTTGATTCAAGTAACTGATGACATGAACTCTATCGGACGTTCTATCAACTCTGCTTTGCAATTATCCCGTA
TTGGTGGAGGAG

TTGGGATTACCTTGTCTAACCTCCGTGAAGCTGGCGCACCAATCAAAGGCTATGCTGGTGCAGCCTCAGGAGT
TGTTCCCTGTTA

TGAAATTATTTGAAGATAGTTTTTCTTATTCAAATCAACTTGGGCAACGTCAAGGAGCTGGTGTGTTTACCT
AAATGTTTTTCAT

CCTGATATCATTGCTTTCTTATCTACTAAAAAAGAAAAATGCCGATGAAAAGGTGCGTGTTAAACCTTGTCAC
TAGGGATTACCG

TTCTTGATAAATTCTACGAATTAGCTCGTAAAAACGAGGACATGTATCTCTTTAGTCCTTACAATGTTGAAAA
AGAATATGGCATT

CCCTTTAACTATCTCGACATTACCAATATGTACGATGAGTTAGTGGCGAACCTAAAATTACTAAGACTAAAA
TTAAAGCTCGTGA

TCTTGAAACAGAGATTTCAAAATTACAACAAGAATCTGGTTACCCTTATATCATCAATATTGATACAGCTAAT
AAAGCTAATCCTAT

CGATGGAAAAATCATCATGAGCAACTTGTGTTCTGAAATTTTACAAGTTCAAACACCTAGCCTTATCAATGAT
GCGCAAGAGTTT

GTAGAAATGGGAACTGATATTTTCATGTAACCTTAGGTTCCACTAATATCCTGAACATGATGACCTCACCAGACT
TTGGCCGTTCTA

TTAAGACCATGACACGTGCCCTAACTTTTGTTACTGATTCATCAAGCATTTGAAGCTGTTCCAACCATTAAACA
TGGCAATAGCCA

AGCTCATACTTTTGGCCTTGGAGCTATGGGACTACATTCTTACCTTGCTCAACATCATATTGAATATGGCAGT
CCAGAATCCATC

GAGTTTACTGATATTTACTTTATGCTCCTGAATTATTGGACCTTGGTCGAATCCAATAACATCGCTCGTGAGC

GCCAAACTACCT

TTGTTGGCTTTGAGAACTCTAAGTACGCTAATGGTAGTTACTTTGATAAATACGTTACAGGACACTTTGTTCC
AAAATCTGATTTG

GTGAAAGATCTGTTCAAAGACCATTTTATTCCGCAAGCTTCAGATTGGGAGGCTCTTCGCGACGCCGTTCAAA
AAGATGGTCTT

TATCATCAAAACCGACTAGCAGTTGCTCCAAATGGCTCTATTTCTTATATCAATGACTGCTCTGCTTCTATTC
ACCCAATCACACA

ACGCATCGAAGAGCGTCAAGAAAAGAAAATTGGTAAAATCTACTATCCTGCAAATGGTTTGTCTACGGATACC
ATTCCTTACTAT

ACATCTGCTTACGATATGGACATGCGCAAAGTTATTGATGTCTATGCCGCTGCGACCGAACATGTGGACCAAG
GCTTGTCTATTA

ACTCTATTCCTTCGTAGTGAGTTGCCTATGGAGCTTTATGAGTGGAAAACACAAAGCAAACAAACCACTCGTG
ATTTATCCATCT

TACGAAACTACGCTTTCAATAAAGGCATTAAATCTATCTACTATATCCGTACCTTTACGGATGATGGGGAAGA
AGTGGGCGCAAA

CCAATGTGAATCTTGTGTCATTTAA

SPy1389

Seq ID 59

ATGAAAGAATTATCGTCTGCACAAATCCGCCAAATGTGGTTGGATTTCTGGAAATCTAAAGGACATTGCGTTG
AGCCTTCAGCTA

ACTTGGTTCCTGTGAACGACCCAACGCTTCTTTGGATCAACTCAGGTGTTGCAACCTTGAAAAAATATTTTGA
TGGTTCAGTGAT

TCCAGAAAATCCACGTATTACCAATGCACAAAAATCAATTCGTACTAATGATATTGAAAATGTTGGTAAAACA
GCACGTCACCATA

CTATGTTTGAAATGCTTGGTAACTTCTCAATTGGAGACTATTTCCGTGATGAAGCTATTGAGTGGGGATTGTA
ACTCTTGACAAG

TCCGACTGGTTTGATTTCCCTAAAGACAAGCTCTACATGACTTATTACCCAGATGACAAGGATTCGTATAAC
CGTTGGATTGCT

TGTGGCGTTGAACCAAGTCACTTGGTGCCGATCGAGGATAACTTCTGGGAAATCGGTGCTGGTCCTTCAGGTC
CAGATACGGA

GATTTTCTTCGACCGTGGTGAAGATTTGATCCAGAAAAATATCGGACTTCGCCTCTTGGCTGAAGATATCGAA
AACGATCGTTAC

ATCGAAATCTGGAACATCGTTCTCTCACAATTCAATGCTGACCCAGCCGTACCACGTTTCAAGATACAAAGAAT
TACCAACAAAA

ACATTGATACAGGTGCTGGTCTTGAACGTCTTGCAGCTGTTATGCAAGGGGCAAAAACAACTTTGAACTGA
CCTCTTCATGC

CAATCATCCGTGAAGTAGAGAAGTTGTCAGGTAAAACTTTACGATCCAGATGGCGACAACATGAGTTTCAAGG
TTATCGCTGACC

ACATCCGTGCGCTTTCATTTGCTATCGGTGATGGTGCGCTTCCTGGAAATGAAGGTCGTGGTTACGTTCTTCG
TCGTCTTCTCC

GTCGTGCGGTTATGCACGGTCGCCGCTTGGCATCAACGAACTTTCCTTTACAAATTGGTTCCGACTGTTGG
ACAAATCATGG

AAAGCTACTACCCAGAAGTGCTTGAAAAACGTGATTTTATCGAGAAAATCGTTAAACGTGAGGAAGAAACATT
TGCTCGTACTAT

CGATGCAGGTAGCGGTCACTTAGATTCAATTGCTT GCGCAGCTTAAGGCTGAAGGTAAGGATACTCTTGAAGG
TAAAGATATCTT

CAAACCTTTATGATACTTATGGATTCCCGGTTGAATTGACAGAGGAATTGGCAGAAGATGCAGGCTACAAGATT
GACCACGAAGG

CTTTAAGTCAGCCATGAAAGAACAACAAGACCGTGCGCGTGACGCTGTTGTTAAGGGTGGTTCAATGGGGATG
CAAAATGAAA

CCCTAGCTGGTATTGTT GAAGAATCACGATTCGAATACGACACATATAGTCTTGAATCAAGTCTTTCAGTCA
TCATCGCTGATAA

TGAACGTACCGAAGCTGTTTCAGAAGGTCAAGCCCTTCTTGTCTTTGCTCAAACACCATTCTATGCTGAAATG
GGTGGACAGGT

TGCTGACACAGGTAGAATCAAAAATGATAAGGGTGACACAGTTGCTGAGGTTGTTGATGTTCAAAAAGCACCA
AATGGTCAACC

TCTACACACTGTAAACGTTT TAGCATCACTTTCAGTTGGAACAACTACACACTTGAAATCAACAAAGAGCGT
CGTTTGGCTGTT

GAGAAAAAC CACACAGCTACTCACTTGCTCCATGCAGCTCTTCACAATGTTATCGGTGAACAC
GCAACTCAGGCTGGTTCATTG

AACGAAGAAGAATTCTTGCGCTTTGATTTTACTCACTTGAAGCAGTAAGCAATGAGGAACTTCGTCACATTG
AACAAGAAGTTA

ATGAGCAAATTTGGAACGCTCTTACAATCACAACGACTGAAACTGACGTTGAAACCGCAAAAGAGATGGGAGC
AATGGCGCTTT

TTGGTGAGAAATATGGTAAAGTGGTTCGTGTGGTTCAAAATTGGTAATTATTCTGTTGAACTTTGTGGTGGAAAC
TCACTTAAATAAT

TCTTCAGAAATCGGTCTCTTCAAGATTGTCAAAGAAGAAGGTATTGGTTCAGGCACTCGTCGTATTATTGCAG
TTACTGGTAGAC

AAGCTTTTGAAGCTTATCGTAACCAAGAGGATGCCCTAAAAGAGATCGCTGCTACTGTAAAAGCTCCGCAATT
GAAAGATGCAG

CAGCTAAAGTACAAGCTCTTAGCGACTCGCTTCGTGATCTTCAAAAAGAAAATGCAGAACTTAAAGAAAAAGC
AGCAGCTGCAG

CAGCTGGTGATGTCTTTAAAGATGTTCAAGAAGCTAAGGGCGTGCGCTTCATTGCTAGTCAAGTTGATGTTGC
AGATGCAGGGG

CACTTCGTACATTTGCTGATAACTGGAAACAAAAAGACTACTCTGATGTGCTTGTTCTCGTAGCAGCTATTGG
TGAGAAGGTTAA

TGTCCTTGTTGCAAGCAAAACCAAGATGTCCACGCTGGTAACATGATCAAAGAATTGGCACCAATTGTAGCA

GGTCGTGGTGG

AGGTAAACCAGACATGGCTATGGCAGGTGGTAGCGATGCAAGTAAAATTGCAGAGCTGCTAGCAGCAGTTGCT
GAAATAGTGT
AA

SPy1390

Seq ID 60

ATGAAAACTCAAATAAACTCATTGCTAGTGTTGTGACATTGGCCTCAGTGATGGCTTTAGCAGCTTGTCAAT
CAACTAATGACA
ATACTAAGGTTATTT CGATGAAAGGTGATACAATTAGCGTTAGTGATTTTTACAATGAAACAAAAACACAG
AAGTATCGCAAAAA
GCGATGCTAAATCTGGTAATTAGTCGTGTTTTTGAAGCTCAATATGGT GATAAGGTTTCAAAAAAGAAGTT
GAAAAGGCGTATC
ATAAACAGCTGAACAGTATGGCGCTTCATTCTCTGCTGCTTTGGCACAATCAAGCTTGACACCTGAGACTTT
TAAGCGTCAGAT
CCGCTCTTCAAAATTAGTAGAATATGCGGTTAAAGAAGCAGCTAAAAAGAATTGACAACACAAGAATATAAG
AAAGCATATGAA
TCTTATACTCCAACAATGGCAGTCGAAATGATTACTTTAGATAATGAAGAGACAGCTAAATCAGTCTTA
GAGGAACTAAAAGCCG
AAGGCGCAGACTTTACAGCTATTGCTAAAGAAAAACAACAACACCTGAGAAAAAGTGACCTATAAATTTGA
TTCAGGTGCGAC
AAATGTACCGACTGATGTCGTAAGGCGGCTTCAAGTTTGAATGA GGGTGGCATATCAGACGTTATCTCGGT
TTTAGATCCAAT
TCTTATCAAAAGAAGTTTTACATTGTTAAGGTGACTAAAAAGCAGAAAAAAATCAGATTGGCAAGAATATA
AGAAACGTTTGAA
AGCTATCATTTATAGCTGAAAAATCAAAAGATATGAATTTCCAAAACAAGGTTATT
GCAATGCATTGGATAAAGCTAATGTAAAAA
TTAAAGACAAAGCTTTTGCTAATATTTTGGCGCAATATGCAATCTTGGTCAAAAACTAAAGCTGCAAGTGA
AAGTTCAACAACC
AGCGAATCATCAAAAGCTGCAGAAGAGAACCCATCAGAATCAGAGCAAACACAGACATCATCAGCTGAAGAAC
CAACTGAGACT
GAGGCTCAGACGCAAGAGCCAGCTGCACAATAA

SPy1422

Seq ID 61

GTGCTTTATCCAACACCCATTGCAAAGTTAATTGACAGTTACTCTAACTTCCAGGAATTGGTATCAAGACGG
CGACGAGATTAG
CCTTTTATACTATTGGAATGTCAAAATGAAGATGTCAATGATTTTGCTAAAACTTATTAGCAGCTAAAAGAGA
ACTGACCTATTGT

TCGATTTGTGGAAACCTTACCGATGACGATCCTTGTACATTTGCACAGACACGAGTCGTGATCAGACGACCA
TTCTGGTAGTA

GAAGATGCTAAAGATGTTTCTGCCATGGAAAAAATCCAAGAGTATCATGGCTATTATCATGTGCTTCACGGCT
TGATTTGCGCTA

TGAATGGTGTGGGGCCAGATGACATCAACCTTAAAAGTTTAATTACCCGTCTAATGGATGGTAAGGTGAGCGA
AGTTATCGTAG

CTACCAATGCCACAGCAGATGGGGAAGCAACGTCCATGTATATTTACGTCCTTGAAACCAGCCGGGATTAA
GGTAACTCGTT

TGGCAAGAGGTCTCGCAGTAGGTTTCAGATATTGAGTATGCTGATGAAGTAACATTATTGAGAGCTATTGAGAA
TCGTACTGAACT

TTAA

SPy1436

Seq ID 62

ATGGATATGTCTAAATCAAATCGTCGTACTTGGCAAGGTTTAGTTGTTATTTAATAGCTATTCTCACCCTT
TTACCACAAGTAC

TGTTACGGCAGCCAGAAAAATTAGAAATTTCCCTGATACCACGGAAATTTTGTTAGGAACGAAGGCGACTGAG
ACACCAGGAAT

CTTACCATTCACTGGTAGCTACCAATTAGTTTTGGGCGATCTTGACAATCTGCAAAGGCCAACCTTCGCACAC
ATCCAGCTAAAA

GATCAAGATGAGCCTAATATTAAACGAAAAGGACTTAAATTCATCCTCCTGGCTGGCATAATTACAAATTGA
CTGACGCTAATG

GAAAAACAACCTTGGTTAATGGACCGTGGCCATTTAGTTGGTTACCAATTTAGCGGCTTAAATGACGAGCCTAA
AAACCTAGTTAC

AATGACAAAATATCTTAATACTGGCTTTAGTGACAAAAATCCTTTAGGAATGCTCTATTATGAAAATAGATTA
GATAGCTGGTTAG

CTCTACACCCTAACCTTCTGGCTAGACTATAAAGTTACTCCTGTTTATCATAAAAATGAGTTAGTTCCTCGCCA
AGTAGTTCTACAG

TATGTTGGAATTGATGAAAATGGAGATCTACTTCAAATTAAGTTAGGTAGTGAAAAAGAAAGTGTAGACAAC
TTGGAGTAACAT

CAGTTACATTAGATAACGTATCTCCTTTAGCTGAATTGGATTACCAAACAGGAATGATGCTAGATTCAACTCA
AAACGAAGAAGA

TAGTAATTTAGAAACCGAAGAGTTTGAAGAAGCGGCTTAA

SPy1494

Seq ID 63

ATGACTAGTAAAAAAGCGTGTTTATCAAGCATCATTGTGTTAGCAAGTTTAACGTGTGGAAATGATACTGTTA
GTGCCAATCATCT

CTCAGCAACTGGAGATAAGTTTGATGATTGCTCAACACTTGTGAAAAAGATGTGGCCCCCTAAAGATGAACTT

GAGATGTTAGCA

TGGTCCTCGTCTCAAACAACTGATGATGCTGACAGAGACTATGAAGATTTTCTCGATGATGATTCTTTTATTT
CTCAAAATGAAAC

TGATAAGATGTTTGAGAATTTAACTGATGATAGGTTATTAAATGAATTAGATGAATTAGATGAAGAAAATGAA
GAAGATGAAGAAG

ATACAATTGAGCCAGAGCAAAATGTAATAATGCCTAGTGACGATGAGCTATTTGATTAACTGATGCTGTTGA
GACACGCCTTAC

TGTTTCTAGTGCTCCCCATTTAGAGGCTGAATTGCCGAAACCACATTTGAGGAGCCTATCAGATACAGCACTG
CGGTCTGGTGA

AATTAGAGGACATTTAGATAACAACTGGACGCTTTGTCTGTAACAGCTACAAAGTTAGCATTAACGATGGCT
CAAAAATTTGATT

TGACAACGCATGTCTATTCTATAGGTGAAAGCTTTAGTGAAGTATTAGCTGCTCATTATGAAGACAGAAAAGC
AGAATCAGCTTT

TTCTAAGAAAAAGAGATTTACCTTCCTATTGCTACTCCAGATGTTGTTATAGAGGAGTTAAGGCGCCTAGTC
TCTTCTATTGGA

AGTTCAAAAGAAGATGTTTCAGTTCCTTATAGTCGGAAGCTAGGTATGGCAGTTGCAAAAAGAAAAATAGCCC
TGCCACAAACG

GGAGAGAGGTTCTCTTATTATCCAGTTTTACTTGGTTTAATGATATTAGGATTAACGCCGATTATGATACCAA
AGAAGATAAATAA

TTAG

SPy1523

Seq ID 64

ATGGCAAAAGATAAAGAGAAAACAAAGTGATGACAAGCTCGTTTTGACAGAGTGGCAAAAGCGTAACATTGAAT
TTTTAAAGAAAA

AGAAGCAGCAAGCTGAGGAAGAAAAAACTCAAAGAGAAATTATTGAGTGATAAAAAAGCGCAGCAGCAAGC
TCAAAATGCTT

CTGAAGCGGTTGAGCTTAAACTGATGAGAAAACTGATAGTCAGGAAATTGAGTCAGAAACGACGTCAAAACC
TAAAAAAACCA

AAAAAGTTAGACAACCCAAGGAAAAAAGCGCGACACAAATCGCTTTTCAAAAATCCTTGCCTGTTCTTTTGGG
GGCGCTCTTACT

CATGGCGGTGTCTATTTTTATGATCACTCCTTATAGCAAAAAGAAAGAGTTTTCTGTAAGAGGAAACCATCAA
ACGAACCTTGAC

GAATTAATCAAAGCTAGCAAAAGTCAAAGCATCTGACTATTGGTTAACGCTGTAACTTCGCCTGGTCAGTATG
AACGACCGATTG

TTCGTACTATTCCATGGGTGAAATCTGTACATCTCTTACCAATTCCTAATCACTTTCTATTTAACGTTAT
TGAATTTGAAATCA

TCGCTTATGCACAAGTTGAAAACGGTTTTTCAGCCTATTTTGGAGAATGGAAAACGTGTGGACAAGGTCAGGGC
ATCAGAACTAC

CGAAATCTTTCTTGATTCTTAATTTAAAAGATGAGAAAAGCGATCCAACAGTTAGTTAAGCAATTAACGACATT
ACCTAAAAAATTA

GTCAAGAATATCAAGTCAGTGTCTCTTGCAAATTCAAAAACGACAGCGGATTTACTACTTATTGAAATGCATG
ACGGTAATGTAG

TTAGAGTACCGCAGTCACAACTCACATTGAAACTTCCCTATTATCAAAAAATTGAAAAAAACCTTGAAAATGA
TAGTATAGTGGAT

ATGGAAGTGGGAATTTATACTACAACACAGGAGATTGAAAAACAACCTGAAGTTCCTCTTACGCCTGAACAAA
ACGCAGCTGATA

AAGAAGGAGATAAGCCTGGTGAACATCAGGAACAGACAGACAATGATTCAGAAACGCCAGCAAATCAGAGTAG
TCCTCAGCAA

ACACCACCATCCCCAGAAACGGTCCTCGAACAGGCCCATGGCTAG

SPy1536

Seq ID 65

ATGAAAAGACTTAAAAAAATCAAATGGTGGTTAGTGGGTCTGCTAGCTTTAATCTCTTTGTTGCTAGCGTTAT
TTTTTCCGCTACC

TTATTATATTGAAATGCCTGGAGGCGCTTACGATATTCGGACTGTCTTACAAGTCAATGGCAAAGAAGACAAA
CGAAAAGGAGC

TTACCAGTTTGTTGCAGTGGGCATTAGTCGTGCCAGCCTCGCTCAGCTATTATATGCTTGGCTGACACCGTTT
ACTGAAATTAGT

ACAGCAGAAGATACAACAGGCGGATACAGCGATGCTGATTCCTTCGAATTAATCAATTTTACATGGAAACAT
CACAAAATGCAG

CTATTTATCAAGCTTTATCCTTAGCTGGAAAAACCAGTTACATTAGATTATAAAGGCGTATATGTTTTAGACGT
AAACAACGAATCT

ACTTTTAAAGGAACGCTACACTTAGCAGATACTGTAACAGGTGTAAATGGTAAACAGTTTACTAGTTCAGCAG
AACTTATTGACT

ATGTTTCTCACCTAAAACTAGGGGATGAAGTTACGGTTCAGTTTACGAGTGATAATAAGCCTAAAAAAGGAGT
TGGCCGTATTAT

CAAACCTGAAAAATGGGAAAAATGGGATTGGCATTGCCTTGACTGATCATACAAGTGTCATTCAGAAGACACA
GTGATCTTTAGT

ACTAAAGGAGTAGGAGGACCTAGTGCTGGTCTAATGTTTACTCTTGATATATATGATCAAATAACTAAAGAAG
ATTTACGCAAGG

GCCGTACAATTGCAGGTACAGGAACTATTGGCAAGGATGGCGAAGTAGGAGATATTGGTGGTGCAGGTCTTAA
AGTAGTTGCA

GCAGCTGAAGCTGGTGCAGATATATTTTTTGTCCGAATAATCCTGTTGATAAGGAAATTAAAAAAGTTAATC
CAAATGCTATAAG

TAATTACGAAGAAGCCAAACGGGCAGCCAAACGACTAAAGACCAAAATGAAGATTGTTCCCTGTTACGACTGTT
CAAGAGGCACT

GGTTTATCTTCGCAAATAA

SPy1564

Seq ID 66

ATGTTGGAACACAAAATTGATTTTATGGTAACTCTTGAAGTGAAAGAAGCAAATGCAAATGGTGATCCCTTAA
ATGGAAACATGC
CTCGTACAGATGCCAAAGGATATGGTGTGATGAGTGATGTCTCCATTAAACGTAAGATTCGTAATCGTTTGCA
AGATATGGGGA
AGTCTATTTTTGTGCAAGCTAATGAGCGTATTGAAGATGATTTTCGTTCACTGGAAAAACGCTTTTCGCAACA
TTTTACAGCTAAG
ACACCTGACAAAGAAATTGAAGAAAAAGCAAATGCATTATGGTTTGATGTTTCGTGCTTTTGGACAAGTTTTTA
CTTATCTGAAAAA
ATCAATTGGGGTGC GTGGACCAGTTTCCATCAGTATGGCTAAGTCCTTGGAGCCAATTGTCATTTCCAGCCTT
CAAATTACGCG
TAGTACCAATGGTATGGAAGCTAAGAATAATAGTGGCCGCTCTTCTGATACGATGGGGACAAAACATTTTGTA
GATTATGGTGTG
TATGTACTTAAAGTTCTATCAATGCTTATTTTGCTGAAAAGACTGGTTTTTCTCAGGAAGATGCTGAGGCTA
TTAAAGAAGTTTT
GGTTAGCTTGTTTGAAAATGATGCGTCGTCTGCACGTCCGGAAGGCTCTATGCGAGTTTGTGAAGTCTTTTGG
TTTACGCATTC
AAGCAAATTGGGAAATGTTTCAAGTGCGCGTGTCTTTGACTTGTTAGAGTATCATCAATCAATAGAAGAAAAA
AGCACTTATGAC
GCTTATCAGATTCATCTAAATCAAGAAAAATTGGCTAAATATGAAGCGAAAGGGTTAACGCTTGAAATCCTAG
AAGGACTCTAG

SPy1604

Seq ID 67

ATGGCAACTAAAAAGTACATATTATTTACACAGTCACTGGGATCGCGAGTGGTACATGGCTTACGAACAAC
ACCACATGCGT
CTGATTAACCTTAATAGATGACCTGTTAGAAGTTTTTCAAACGGATCCTGATTTTCATAGTTTTCATTTGGATG
GGCAAACCATTAT
CCTAGATGATTATTTAAAAGTACGCCCCGAACGAGAACCTGAGATTAGACAAGCCATTGCTTCGGGAAAACTC
CGTATCGGACC
TTTCTATATCTTACAGGACGATTTTTTGACCAGCAGTGAATCCAATGTGCGCAATATGCTGATTGGTAAGGAA
GATTGTGACAGA
TGGGGCGCTAGTGTGCCACTTGGTTATTTTCTGATACCTTTGGAAATATGGGACAAACACCACAGCTGATGT
TAAAAGCCGGC
CTACAAGCTGCTGCCTTTGGTCGTGGCATTTCGTCCAACTGGATTTAACAATCAGGTGGATACCAGTGAAAAAT
ACAGCTCCCAA
TTCTCTGAAATCAGTTGGCAAGGCCAGATAACAGTCGTATTCTTGGA CTCTTCGCCAACTGGTACAGCA

ATGGCAATGAG

ATCCCACAACAGAAGCTGAGGCGCGTCTTTTTGGGATAAAAACTTGCTGATGCCGAACGCTTCGCCTCAA
CCAAGCACCTT

CTGATGATGAACGGGTGTGATCATCAACCCGTACAACCTTGATGTCACCAAGGCAATCGCCTTAGCCAACCAAC
TCTATCCTGAC

TACGAATTTGTGCATTCCTGCTTTGAAGATTACTTGGCTGATCTCGCAGATGATTTACCAGAGAACCTTTCAA
CCGTCCAAGGAG

AGATTACCAGTCAAGAAACCGATGGCTGGTATACCCTAGCTAACACGGCTTCTGCTCGTATTTACCTCAAACA
GGCTAATACCA

GAGTCTCTCGCCAACTCGAAAACATCACCGAACCCCTTAGCAGCAATGGCTTATGAGGTAACAAGTACCTACCC
TCACGACCAAC

TGCGTTACGCTTGAAAAACCTCATGCAAAATCACCCCTCATGATTCTATCTGTGGTTGTAGTGTGATAGCGT
TCATCGGGAAAT

GATGACGCGCTTTGAAAAAGCCTATGAAGTCGGACACTATTTAGCAAAAGAAGCTGCTAAGCAAATTGCTGAC
GCCATTGATAC

CAGGGATTTTCCAATGGATAGCCAACCCCTTCGTCTTATTTAATACCAGCGGCCATTCCAAAACAAGTGTGCT
GAGCTCAGCCT

GACCTGGAAAAATATCATTTTTGGCCAACGCTTTCCTAAAGAGGTTTACCAAGAAGCTCAAGAATATTTGGCA
AGACTCTCCCA

TCTTTCCAAATTATTGACACTAGTGGACAAGTGAGACCCGAAGCAGAAATTTAGGCACAAGCATCGCTTTTG
ACTACGATTTGC

CCAAGAGATCCTTCCGCGAACCTTATTTCGCCATCAAAGTGAGATTACGGCTACCAATAACTCTCCCAGCCAT
GTCTTGAAAA

CCTTAGCATTAAAGCTAGGAAATGAAACAACCTCCTTCAGAAACCGTTTCCCTCTACGATGACAGTAATCAGTG
CCTTGAAATGG

GTTTCTAAAAGTTATGATACAAACCGATGGTCGTCTAACCATCACCGATAAACAATCTGGACTAATCTATCAA
GACCTGTTGCGG

TTTGAAGATTGTGGCGATATTGGAAATGAATATATTTCTCGCCAGCCAAATCATGACCAACCTTTCTATGCGG
ATCAAGGGACCA

TCAAGCTTAACATCATTAGCAACACCGCTCAAGTTGCTGAACTTGAAATCCAGCAAACCTTTGCCATTCTAT
CTCCGCAGATAA

GCTCTTACAGGCTGAGATGGAGGCTGTCATTGACATCACAGAACGCCAAGCAAGACGTTTACAAGAAAAGGCT
GAGCTAACCT

TAACAACCTTATCCGCATGGAGAAAAATAATCCTCGCCTCCAATTCACCACACGTTTTGATAACCAAATGAC
TAATCATCGCTT

GCGCGTCCTATTCCCAACGCACCTTAAAAACAGACCATCATCTAGCTGACAGTATTTTTGAAACTGTCAAACGT
CCAAATCATCCA

GATGCCACCTTTTGGAAGAATCCAAGTAACCCACAGCACCAAGAATGCTTTGTGAGTCTCTTTGATGGTGAAA
ATGGAGTCACT

ATTGGTAACTATGGCCTCAACGAATATGAGATCTTACCAGATACCAACACCATTGCCATCACTCTCTTACGTT
CTGTTGGCGAAA

TGGGCGACTGGGGTTACTTCCCAACACCTGAAGCCCAGTGTCTTGGCAAACACAGCCTTTCTTATAGTTTTGA
AAGCATCACTA

AGCAAACACAATTTGCCAGCTACTGGCGAGCTCAAGAAGGCCAAGTCCCTGTTATTACCACACAAACAAACCA
ACACGAGGGAA

CATTAGCCGCGAGAATATAGCTATTTGACGGGTACAAACGACCAAGTTGCCCTCACAGCTTTCAAACGTCGCTT
AGCAGACAATG

CCCTTATCACGCGCAGCTATAATCTCTCAAACGATAAAACTTGTGACTTTAGCCTAAGCCTGCCAAACTACAA
TGCCAAGGTCAC

TAATTTGTTAGAAAAAGACAGCAAGCAAAGCACACCCAGCCAACTTGGCAAAGCGGAAATTTTAACTCTAGCT
TGGAAGAAACA

ATAA

SPy1607

Seq ID 68

ATGAAAATCACTAAAAATTGAAAAGAAAAAACGCCTCTACCTTATCGAATTGGATAATGACGAATCCCTTTATG
TAACAGAAGATAC

TATTGTTTCGGTTTATGTTGAGTAAAGATAAAGTCCTTGACAATGATCAGCTTGAAGACATGAAACATTTTGCC
CAACTGTCCTAC

GGCAAAAATTTAGCCCTTTATTTTCTTTTCCTTTCAACAACGCAGCAACAAGCAAGTTGCTGATTACCTGCGCA
AGCATGAGATTG

AAGAACACATTATTGCTGACATCATCACTCAACTCCAAGAAGAACAATGGATAGACGACACCAAATTGGCTGA
TACCTACATTG

CCAAAATCAGTTAAATGGTGATAAAGGTCCCCAAGTCTTAAACAAAAATTATTACAAAAGGCATTGCAAGT
CATGACATTGATC

CTATCTTATCTCAAACCTGACTTTAGCCAACTCGCTCAAAAAGTAAGCCAAAACTCTTTGACAAATATCAAGA
AAAATTGCCACCA

AAAGCCTTGAAAGATAAAATCACCCAAGCATTACTGACCAAAGGCTTTTCATACGATCTAGCTAAACATAGCC
TCAATCACCTTA

ATTTTGACCAAGATAATCAAGAAATAGAAGATCTTCTTGACAAAGAATTAGACAAACAATATCGTAAACTCAG
TCGCAAATATGAT

GGTTATACCTTAAAGCAAAAAGCTCTATCAGGCTCTCTACCGAAAAGGCTACAACAGCGACGACATTAATTGCA
AGTTAAGAAATT

ATTTATAG

SPy1615

Seq ID 69

ATGATCTGTCTACTATGTCAACAAATTAGTCAAACACCAATAAGTATTACAGAAATCATCTTTTTAAGACGTA

TCTCTTCACCGATT

TGTCAACAATGTCAAAAAAGCTTTCAAAAGATAGGAAAAAGTGTGTCGACATGTTGTGCAAACTCAGATA
TAATAGCTTGTC

GAGATTGTCTAAAATGGGAAAAACAAAGGATACAATGTAAATCATAGAAGCTTATATTGTTATAATGCTGCTAT
GAAAGCATACTTC

AGTCAATATAAGTTTCAAGGAGACTATTTATTAAGAAAAGTTTTGTCAGTAGAACTTGCCGATGTTATCACCA
AGTACTATAAAGG

CTATATCCCAGTCCCGGTTCTGTAAAGTCCCGGTTGTTTTCGAGAAAAGACAATTTAATCAAGTGAGCGCTATT
CTTGAGGCAGCT

AATGTTAGCTACCTTTCTCTTTTGAAAAAGCTAGATAATACTCACCAATCTTCCAGAACAAAAAAGAGAGAT
TATTAGTAGAAAA

ATCTTATCGACTACTAAAAGTATCAAAACATTCCTGATAAAATCCTTATAGTAGATGATATTTATACTACTGGT
AGTACAATTATCGC

TCTTAGAAAACAATTGGCTAAAGTAGCAAATAGTGACATTAAGTTGTCAATTGCACGTAA

SPy1666

Seq ID 70

ATGAAATCCTTTTCTCTTACTTTTTTCATTTCTAAACCTTTTGAAGTATGGTACAATTAAAGTTATGACAAAAG
AATTTTCATCACGTG

ACCGTACTCCTTCACGAAACAGTGGACATGCTTGACATAAAGCCTGATGGGATTTATGTTGATGCGACGCTAG
GTGGCTCAGGC

CACTCAGCTTATTTGTTGTCCAAACTTGGTGAAGAAGGGCACCTCTATTGTTTTGACCAAGACCAAAAGGCTA
TTGACAATGCAC

AAGTTACCCTCAAATCTTATATTGACAAAGGACAGGTAACTTTTATTAAGATAATTTTAGACACCTCAAAGC
ACGTTTAACAGCG

CTTGAGTTGATGAAATTGATGGTATCTTATATGACCTTGGTGTTCAGCCCGCAATTGGATGAAAGAGAAC
GAGGGTTTTCTT

ATAACAAGATGCTCCATT GGATATGCGCATGGATCGTCAGTCGCTCTTAACAGCTTACGAAGTGGTGAATA
CCTATCCATTCAA

TGATTTGGTTAAGATTTTTTTCAAATATGGTGAAGATAAATTCTCCAAGCAGATCGCTCGAAAAATTGAACAA
GCAAGAGCTATTA

AGCCTATTGAGACAACAACAGAGTTGGCAGAATTGATTAAGGCAGCAAAGCCAGCTAAAGAGTTGAAGAAAAA
AGGCCACCCT

GCTAAACAGATTTTTCAAGCTATTCGCATTGAAGTCAATGATGAATTGGGAGCGGCCGATGAATCTATTACAGG
ACGCTATGGAAT

TATTAGCCCTTGATGGTCGTATCTCAGTTATTACCTTCCATTCTCTGGAAGATCGCCTAACCAAGCAGTTGTT
TAAAGAAGCTAG

TACGGTGGATGTGCCAAAAGGGCTTCCTCTAATTCCTGAAGATATGAAACCTAAGTTTGAAGTTGTTTCACGT
AAGCCGATCTTA

CCTAGTCATT CAGAGTTAACAGCTAATAAAAGGGCACACTCAGCCAAGCTACGTGTTGCCAAAAAATTCGGA
AATAA

SPy1727

Seq ID 71

GTGACAACGACGGAACAAGAACTTACCTTGACTCCCTTACGTGGGAAAAGTGGCAAAGCTTATAAAGGCACTT
ATCCAAATGGG

GAATGTGTCTTTATAAAATTAAATACGACCCCTATTCTACCTGCCTTAGCAAAAGAACAGATTGCGCCACAGT
TACTTTGGGCCA

AACGCATGGGCAATGGTGATATGATGAGTGCCCAAGAATGGCTTAACGGCCGTACATTGACCAAAGAAGATAT
GAACAGTAAG

CAAATCATTCATATTCTATTGCGCCTTCACAAATCTAAAAAATTAGTCAATCAACTGCTTCAGCTCAATTATA
AGATTGAAAACCCA

TACGATTTATTGGTTGATTTTGAGCAAAATGCACCCTTGCAAATTCAGCAAAATTCATACTTACAAGCTATCG
TTAAAGAATTAAA

ACGGAGCTTACCAGAGTTCAAATCAGAAGTAGCAACGATTGTGCATGGAGATATTAAACATAGCAATTGGGTG
ATTACTACTAGT

GGTATGATTTTTTTAGTAGATTGGGATTCTGTTCGTCTAACTGATCGGATGTATGATGTTGCTTACCTGTTGA
GCCACTATATTCC

ACGGTCTCGTTGGTCAGAATGGCTGTCTTATTATGGCTATAAAAAATAATGACAAGGTTATGCAAAAAATTATT
TGGTATGGTCAAT

TTTCTCACCTGACACAAATTCTCAAGTGTTTTGACAAGCGTGACATGGAGCATGTGAATCAGGAGATTTATGC
CCTCAGAAAATT

TAGAGAAATATTTAGAAAGAAATAA

SPy1785

Seq ID 72

ATGATATTAACAGCTCCTATGTCCAACTTAAAGGGATTGGACCAAAATCAGCAGAAAAATTCAGAAATTAG
ATATTTATACAGT

AGAAGATTTACTGCTTTATTATCCGTTTCGCTATGAAGATTTTAAATCAAAATCTGTTTTTGATTTAGTGGAT
GGTGAAAAAGCAG

TCATTACAGGCTTAGTCGTTACTCCAGCTAATGTACAATATTATGGTTTTAAACGTAACCGTTTAAGTTTCAA
ATTGCGTCAAGGG

GAAGCTGTCTTAAATGTTAGTTTTTTTAAATCAACCCTATTTAGCTGATAAAATAGAACTTGGTCAAGAGGTAG
CTGTTTTTGGTAA

ATGGGATGCCACTAAATCGGCTATTACTGGGATGAAGGTTTTAGCTCAAGTTGAAGATGACATGCAACCTGTT
TATCGCGTAGC

TCAGGGAATTTACAGTCTACTTTGATTAAAGCTATTAAGTCAGCTTTTGAAATCGATGCGCATTTGGAATTG
AAGGAAAATTTAC

CAGCTACTTTATTGGAAAAATACCGATTGATGGGTCGTAGTCAGGCTTGTTTAGCTATGCATTTCCCAAAGA
TATCACAGAGTA
TAAGCAAGCGCTCCGTCGGATTAAATTTGAAGAATTATTTTACTTTCAAATGAACCTTCAAGTTTGAAGCC
GAAAATAAATCTG
AAACAAATGGTTTGCCTATTCTTTATAGTAAACGTGCTATGGAGACAAAGATTCCTCTTTACCTTTTATTCT
AACGAATGCTCAA
AAGCGCTCTTTAGATGACATATTATCTGATATGTCATCGGGAGCTCATATGAATCGTTTATTGCAAGGAGATG
TAGGATCAGGAA
AGACAGTCATTGCTGGTCTATCAATGTATGCAGCTTATACAGCAGGTTTTCAATCGGCTTTGATGGTTCCAAC
GGAAATCCTAGC
TGAACAACACTACATTAGTCTGCAAGAGTTATTTCCAGATTTATCAATCGCTATATTAAGTTTCGGGTATGAAA
GCAGCTGTCAAG
CGTACGGTTTTAGCAGCTATTGCAAATGGCTCGGTTGATATGATTGTAGGAAGTCATGCTCTTATCCAAGACT
CGGTACAGTACC
ATAAACTGGGGCTTGTCATTACAGACGAGCAACATCGTTTTGGTGTTAAACAGCGTAGAATTTCCGTGAAAA
GGGAGAAAATC
CTGATGTTTTAATGATGACAGCCACCCCAATTCCCCGAAGTCTAGCAATCACAGCTTTTGGAGAAATGGATGT
TTCTATTATTGA
TGAATTACCTGCCGGTCGTAAACCTATTATGACACGCTGGGTGAAACACGAGCAGCTAGGTACTGTGTTGGAA
TGGGTAAAG
GTGAATTGAAAAAGATGCTCAAGTGTATGTCATTTACCGTTGATTGAAGAATCAGAAGCTTTAGATTTAAA
GAATGCAGTAGC
ATTGCATGCTGAATTATCTACTTATTTTGAAGGAATTGCTAAGGTTGCTCTTGTACATGGACGTATGAAAAAT
GATGAAAAAGATG
CTATAATGCAAGATTTCAAGGATAAAAAAAGTCATATTTTAGTATCCACAACAGTTATTGAAGTAGGGGTAAA
TGTCCCAATGCA
ACGATCATGATTATTATGGATGCCGATCGTTTTGGATTAAGTCAGTTACATCAACTTCGTGGGCGTGTGGTC
GTGGATATAAAC
AATCATACGCTGTTTTAGTGGCTAATCCCAAACTGATTCGGGGAAAAAACGAATGACAATCATGACAGAAAC
GACAGATGGTTT
CGTTTTAGCTGAGTCGGATTTAAAAATGCGTGGTCTGGTGAAATCTTTGGTACTCGTCAGTCTGGAATTCCA
GAATTTCAAGTA
GCTGATATCGTTGAGGATTATCCTATTTTAGAAGAAGCACGCAAAGTTCTGCAGCGATTGTTTCTGATCCTA
ACTGGATATATG
AAAAACAGTGGCAATTAGTGGCACAAAATATTAGAAAAAAGAAGTTTATGATTAA

SPy1798

Seq ID 73

ATGAAAAAATCAGCAAATGTGCGTTTGTGGCAATATCTGCCCTTGTTCTCATTCAGGCTACTCAAAGTGTAA

AATCACAAGAGC

CTTTAGTTCAGTCACAACCTCGTGACAACAGTAGCTTTAACTCAGGATAATCGACTTTTAGTTGAAGAGATAGG
CCCTTACGCTAG

TCAATCAGCTGGAAAAGAGTATTATAAACATATTGAAAAGATTATTGTTGATAATGATGTCTATGAAAAAAGC
CTGGAGGGCGAG

CGAACCTTTGATATTAACCTACCAAGGGATTAAGATCAATGCTGACCTATTAAAGACGGTAAGCATGAATTGA
CTATTGTTAATAA

AAAAGATGGTGATATCCTAATTACCTTTATTAAAAAGGGCGATAAAGTGACCTTTATTTTCAGCTCAAAAATTA
GGAACAACAGATC

ATCAGGATTCATTAAAAAAGATGTGCTCAGTGATAAAACAGTGCCACAAAACCAAGGCACACAAAAAGTTGT
TAAATCTGGGAA

AAATACTGCTAACTTGTCTAATAACAAAAATTGAGTCAAGAAGATGGTGCAATTTTATTTCCAGAAATTGAT
CGTTATTCTGATAA

CAAACAGATAAAAGCATTGACTCAGCAAAATCACAAAGGTTACAGTCAATGGTACAGTTTATAAAGATCTTATT
TCAGATTCTGTAA

AAGATACTAATGGCTGGGTCTCGAATATGACAGGGCTTCATCTTGGAACAAAAGCTTTCAAAGATGGAGAAAA
TACAATCGTGAT

ATCCTCAAAGGATTTGAAGACGTTACTATTACCGTTACCAAGAAAGATGGTCAAATCCATTTTGTATCTGCC
AAACAAAAACAAC

ATGTGACTGCTGAAGACAGACAATCAACAAAGTTGGATGTCACCACTTTGGAAAAAGCTATCAAAGAAGCGGA
TGCGATTATTG

CTAAAGAAAGCAACAAAGACGCGGTCAAAGATCTGGCTGAGAACTTCAAGTCATCAAGGATTCTTACAAAGA
AATCAAAGATA

GTAAGCTACTCGCCGATACTCATCGACTGTTAAAAGATACCATCGAGTCTTATCAAGCAGGTGAGGTTTCTAT
TAACAATCTCAC

AGAAGGAACCTATACGCTAAACTTTAAAGCTAATAAAGAAAACTCAGAAGAGTCCTCCATGCTTCAAGGTGCT
TTTGATAAAAGA

GCCAAATTAGTGGTTAAAGCAGATGGTACAATGGAAATTTCCATGCTTAATACTGCTTTGGGACAATTTTGA
TTGACTTTTCTAT

TGAAAGCAAAGGGACCTACCCAGCAGCAGTGCGTAAACAAAGTTGGCCAAAAAGATATCAATGGTAGCTATATT
CGAAGCGAATT

TACCATGCCTATTGATGATTTGGATAAAATTACACAAAGGTGCTGTTTTGGTATCAGCCATGGGAGGTCAAGAA
AGTGATTTAAAC

CACTATGACAAATACACCAAACTTGACATGACCTTTAGTAAGACCGTTACCAAAGGCTGGAGTGGTTATCAGG
TAGAAACTGAT

GATAAAGAAAAAGGGTTGGGACTGAACGTCTTGAAAAAGTTTATGTTAACTTGGCAAAGATTTAGACGGCG
ATGGTAAATTAT

CAAAAACGGAATTAGAACAGATTCGAGGCGAGTTGCGTCTAGACCATTACGAGTTAACTGATATTTCTTTATT
GAAACATGCTAA

AAATATTACAGAACTACATCTGGATGGAAACCAAAATTACGGAAATTCCTAAAAGAGTTATTTAGTCAAATGAAG
CAACTTCGATTTTC
TTAACTTAAGAAGTAATCATTTAACTTATCTAGACAAAGATACATTTAAAAGCAATGCTCAATTAAGAGAAGT
CTACTTATCAAGTA
ACTTTATTCACTCTCTTGAAGGAGGACTATTCCAGTCGCTTCATCACCTGGAGCAACTTGATCTTTCCAAGAA
TCGTATTGGCCG
ACTTTGTGATAACCCATTTGAAGGATTGTCTCGTCTGACTTCATTAGGTTTCGCAGAAAATAGTCTTGAGGAG
ATACCTGAAAAA
GCGCTAGAGCCTCTAACATCACTTAATTTTATCGACTTATCTCAAAAATAATTTAGCACTACTGCCAAAAACAA
TAGAAAAATTGCG
CGCCTTAAGCACTATTGTGGCAAGTAGAAATCATATTACTCGTATTGATAATATTTCAATTTAAAAATCTTCCT
AAATTATCTGTACT
CGATTTATCAACTAATGAAATTTCAAATCTTCCAAATGGTATATTTAAACAGAATAACCAATTAACAAAAGTT
GATTTTTTCAATAAC
TTGCTTACTCAGGTTGAAGAATCAGTATTTCCAGATGTTGAAACGCTTAATTTAGATGTGAAGTTCAATCAGA
TAAAAAGTGTGAG
TCTAAAAGTAAGAGCTCTTATCGGACAACACAACTGACTCCACAAAAACATATTGCAAACTTGAAGCTTCC
TTAGATGGCGAA
AAAATAAAATATCATCAAGCTTTTCACTCTTTTAGATTTGTATTATTGGGAGCAAAAAACAAATTCTGCCATTG
ATAAAGAAGTAGT
GTCTGTTGAAGAATATCAACAATTGTTACAAGAAAAAGGTTTCAAGATACGGTTTCTTTACTTAATGATATGCAA
GTCGATTGGAGTA
TTGTGATTCAAGTTGCAAAAAAAGCTTCCAATGGACAGTATGTGACGGTTGACGAAAAGCTTCTCTCAAATGA
TCCGAAAGATGA
CTTAACGGGAGAGTTTTCTTTAAAAGATCCAGGTACATATCGGATTCGCAAAGCTTTAATAACAAAGAAATTT
GCTACTCAAAAAG
AACATATCTATTTGACATCTAATGATATCCTTGTGGCGAAAGGACCACATTCACATCAGAAAGATTTAGTTGA
GAACGGCCTTAG
AGCATTAATCAAAAAACAATTGCGTGATGGTATTTACTATTTAAATGCCAGCATGTTAAAAACTGACTTAGCA
TCTGAGTCCATGT
CAAACAAGGCTATTAATCATCGAGTGACTTTGGTAGTTAAAAAAGGTGTTTCCTATTTAGAAGTTGAGTTTAG
AGGTATAAAGGTT
GGTAAATGTTAGGCTACCTTGGTGAATTGAGCTATTTTCGTAGATGGTTACCAAAGAGATTTAGCTGGTAAAC
CAGTTGGTCGAA
CAAAAAAGGCAGAGGTTGTGTCTTATTTACGGATGTAAGTGGCCTACCATTTGGCAGATCGTTATGGAAAAAA
CTATCCAAAAGT
GCTGCGTATGAAATTGATTGAACAAGCGAAGAAAGACGGACTTGTGCCATTACAGGTCTTTGTGCCTATCATG
GATGCCATTTTC
AAAAGGGTCTGGCCTTCAAACCGTTTTTATGCGTTTAGACTGGGCAAGCCTTACAACAGAGAAGGCAAAGGTT

GTCAAAGAAAC

TAATAATCCACAAGAAAAATAGCCATCTAACTTCAACAGATCAGTTGAAAGGACCTCAAAATCGTCAACAAGAA
AAAACACCTACA

AGTCCTCCTTCAGCAGCTACTGGTATTGCTAACTTAACTGATCTCCTGGCTAAAAAAGCAACCGGCAATCAA
CTCAAGAACTT

CTAAGACAGATGATACTGATAAGGCAGAGAAAATTGAAGCAGTTAGTGCGTGACCATCAAACATCAATTGAAGG
TAAACAGCAA

AAGATACTAAGACTAAAAAATCTGATAAGAAACATCGTTCCAATCAACAATCAAATGGTGAAGAAAGTAGCTC
TCGTTATCACTTA

ATTGCAGGTCTATCTAGCTTTATGATCGTAGCTCTGGGATTCATTATTGGTCGAAAGACATTATTTAAATAA

SPy1801

Seq ID 74

ATGAATAAAAACAAACTATTAAGAGTTGCCATGCTACTAAGTCTCTTAGCCCCGACAGCAGAAAGCATGACAG
TGCTGGCTCAA

GATGTAATGCTTGAGACGCATAAAGCAACTACAAATGAAACCAGTGATTCTTCTTCAAAAGAGGAAAAATAATA
AAAATGCAGCAC

CTACAACATCAGATAAAACTGACCAAGGTCCCCTTGATGCTTCTGCAGAAACAACTCTAATAGTCTTGTTAA
CGCGGATGATAA

AAAAAGAAGCGATTCTAGTCAGTCTGCTATAGGCTCTTCGGACAACAAGGCAGAAGCAGAAAACCAGGTAGAT
GATAAATCAAC

TGATCATTCGAAATCAACTGATCATTCGAAACCAACTGACCAGCCCCAAACCATCACCATCTAAAGTTGATACG
GCACCTGCTTCT

TCATTGTCGAAACAACTGCCAGAGGCAAGAACTCCTATTCAGTCGTTGTCCCCTTACGTATCAGATTTAGATT
TGAGTGAGATAG

ATATCCCTTCTGTCAACACATACGCGGCATATGTAGAGCATTGGAGTGGTAAAAATGCCTATACCCACCATCT
TTTATCTCGCCG

TTATGGTATTAAAGCTGACCAGATTGATAGTTACTTAAAAATCAACAGGCATTGCCTATGACAGCACACGTATT
AATGGTGAGAAG

CTATTGCAATGGGAAAAGAAAAGTGGGCTGGATGTTGAGCTATCGTAGCTATTGCGATGTCTGAGAGTTCTT
TAGGAACTCAA

GGGATTGCAACTTTGCTTGGAGCTAATATGTTTGGCTATGCAGCTTTTGATCTAGATCCGACTCAAGCAAGTA
AGTTTAATGATG

ATAGTGCTATTGTCAAAATGACACAAGACACCATTATTA AAAACAAAAATAGCAATTTTGCACTTCAAGATTT
AAAAGCGGCTAAG

TTTTACGAGGTCAATTAACTTTGCAAGTGACGGGGGTGTTATTTTACTGATACTACTGGTAGTGGTAAAC
GTCGCGCACAAA

TTATGGAAGACCTGGATAAGTGGATTGATGACCATGGTGGCACACCAGCCATTCCAGCCGAATTGAAAGTGCA
GTCATCAGCTA

GTTTTGCATCTGTGCCAGCAGGTTATAAGCTCTCTAAGAGTTATGATGTCTTGGGTTATCAAGCTTCGAGTTA
TGCTTGGGGACA

ATGCACTTGGTATGTGTATAATCGCGCCAAAAGAATTGGGTTACCAATTTGATCCTTTTATGGGAAATGGTGG
GATTGGAAGTAT

AAAGTAGGGTATGCCCTTTCAAAGACTCCAAAAGTAGGTTATGCTATTTCAATTTGCACCAGGGCAAGCGGGCG
CTGATGGCACT

TATGGCCACGTATCAATTGTAGAAGATGTTAGAAAAGATGGGTCTATTCTTATTTAGAGTCTAACTGTATCG
GCTTAGGTAAGA

TTTCTTATCGTACCTTTACAGCTCAGCAGGCTGAACAGCTAACATATGTTATTGGCAAGAGTAAAACTAA

SPy1813

Seq ID 75

ATGGATAAACATTTGTTGGTAAAAAGAACACTAGGGTGTGTTTGTGCTGCAACGTTGATGGGAGCTGCCTTAG
CGACCCACCAT

GATTCACTCAATACTGTAAAAGCGGAGGAGAAGACTGTTTCAGGTTTCAGAAAGGATTACCTTCTATCGATAGCT
TGCATTATCTGT

CAGAGAATAGCAAAAAAGAATTTAAAGAAGAACTCTCAAAAGCGGGGCAAGAATCTCAAAAGGTCAAAGAGAT
ATTAGCAAAAG

CTCAGCAGGCAGATAAAACAAGCTCAAGAAGCTGCCAAAATGAAAATTCCTGAGAAAATACCGATGAAACCGTT
ACATGGTTCTCT

CTACGGTGGTTACTTTAGAACTTGGCATGACAAAACATCAGATCCAACAGAAAAAGACAAAGTAACTCGATG
GGAGAGCTTCC

TAAAGAAGTAGATCTAGCCTTTATTTTCCACGATTGGACAAAAGATTATAGCCTTTTTTGGAAAGAATTGGCC
ACCAAACATGTG

CCAAAGTTAAACAAGCAAGGGACACGTGTCATTTCGTACCATTCATGGCGTTTCCTAGCTGGGGGTGATAACA
GTGGTATTGCA

GAAGATACCAGTAAATACCCAAATACACCAGAGGGAAATAAAGCTTTAGCCAAAGCTATTGTTGATGAATATG
TTTATAAATACAA

CCTTGATGGCTTAGATGTGGATGTTGAACATGATAGTATTCCAAAAGTTGACAAAAAGAAGATACAGCAGGC
GTAGAACGCTC

TATTCAAGTGTTTGAAGAAATTGGGAAATTAATTGGACCAAAGGTGTTGATAAATCGCGTTATTTATTATG
GATAGCACCTACA

TGGCTGATAAAAACCCATTGATTGAGCGAGGAGCTCCTTATATTAATTTATTACTGGTACAGGTCTATGGTTC
ACAAGGAGAGAA

AGGTGGTTGGGAGCCTGTTTCTAATCGACCTGAAAAAACAATGGAAGAACGATGGCAAGGTTATAGCAAGTAT
ATTCGTCCTGA

ACAATACATGATTGGTTTTTCTTTCTATGAGGAAAAATGCTCAAGAAGGGAATCTTTGGTATGATATTAATTCT
CGCAAGGACGAG

GACAAAGCAAATGGAATTAACACTGACATAACTGGAACGCGTGCCGAACGGTATGCAAGGTGGCAACCTAAGA

CAGGTGGGGT

TAAGGGAGGTATCTTCTCCTACGCTATTGACCGAGATGGTGTAGCTCATCAACCTAAAAAATATGCTAAACAG
AAAGAGTTTAAG

GACGCAACTGATAACATCTTCCACTCAGATTATAGTGTCTCCAAGGCATTAAAGACAGTTATGCTAAAAGATA
AGTCGTATGATC

TGATTGATGAGAAAGATTTCCAGATAAGGCTTTGCGAGAAGCTGTGATGGCGCAGGTTGGAACCAGAAAAGG
TGATTTGGA

CGTTTCAATGGCACATTACGATTGGATAATCCAGCGATTCAAAGTTTAGAAGGTCTAAATAAATTTAAAAAAT
TAGCTCAATTAGA

CTTGATTGGCTTATCTCGCATTACAAAGCTCGACCGTTCTGTTTTACCCGCTAATATGAAGCCAGGCAAAGAT
ACCTTGGAACA

GTTCTTGAAACCTATAAAAAAGGATAACAAAGAAGAACCTGCTACTATCCCACCAGTATCTTTGAAGGTTTCTG
GTTTAACTGGTC

TGAAAGAATTAGATTTGTCAGGTTTTGACCGTGAAACCTTGGCTGGTCTTGATGCCGCTACTCTAACGTCTTT
AGAAAAAGTTGA

TATTTCTGGCAACAACTTGATTGGCTCCAGGAACAGAAAATCGACAAATTTTTGATACTATGCTATCAACT
ATCAGCAATCATG

TTGGAAGCAATGAACAAACAGTGAAATTTGACAAGCAAAAACCAACTGGGCATTACCCAGATACCTATGGGAA
AACTAGTCTGC

GCTTACCAGTGGCAAATGAAAAAGTTGATTTGCAAAGCCAGCTTTTGTTTGGGACTGTGACAAATCAAGGAAC
CCTAATCAATAG

CGAAGCAGACTATAAGGCTTACCAAAATCATAAAATTGCTGGACGTAGCTTTGTTGATTCAAATATCATTAC
AATAACTTTAAAG

TTTCTTATGAGAACTATAACCGTTAAAGTAACTGATTCCACATTGGGAACCACTACTGACAAAACGCTAGCAAC
TGATAAAGAAGA

GACCTATAAGGTTGACTTCTTTAGCCCAGCAGATAAGACAAAAGCTGTTCATACTGCTAAAGTGATTGTTGGT
GACGAAAAAACC

ATGATGGTTAATTTGGCAGAAGGCGCAACAGTTATTGGAGGAAGTGCTGATCCTGTAAATGCAAGAAAGGTAT
TTGATGGGCAA

CTGGGCAGTGAGACTGATAATATCTCTTTAGGATGGGATTCTAAGCAAAGTATTATATTTAAATTGAAAGAAG
ATGGATTAATAAA

GCATTGGCGTTTCTTCAATGATTTCAGCCCGAAATCCTGAGACAACCAATAAACCTATTCAGGAAGCAAGTCTA
CAAATTTTAAAT

ATCAAAGATTATAATCTAGATAATTTGTTGGAAAAATCCCAATAAATTTGATGATGAAAAATATTGGATTACTG
TAGATACTTACGT

GCACAAGGAGAGAGAGCTACTGCATTTCAGTAATACATTAAATAATATTACTAGTAAATATTGGCGAGTTGTCT
TTGATACTAAAG

GAGATAGATATAGTTCGCCAGTAGTCCCTGAACTCCAAATTTTAGGTTATCCGTTACCTAACGCCGACACTAT
CATGAAAACAGT

AACTACTGCTAAAGAGTTATCTCAACAAAAAGATAAGTTTTCTCAAAAGATGCTTGATGAGTTAAAAATAAAA
GAGATGGCTTTAG

AAACTTCTTTGAACAGTAAGATTTTTGATGTAAGTCTATTAATGCTAATGCTGGAGTTTTGAAAGATTGTAT
TGAGAAAAGGCAG

CTGCTAAAAAATAA

SPy1821

Seq ID 76

ATGATTGAAGCAAGTAAGCTTAAAGCAGGTATGACATTTGAAGCAGAAGGAAAATTAATCCGTGTCCTTGAAG
CTAGCCACCAC

AAACCAGGTAAAGGAAACACTATCATGCGTATGAAACTACGTGATGTGCGTACAGGTTCTACTTTTGACACAA
CTTACCGCCCA

GATGAAAAATTTGAGCAAGCCATCATTGAAACTGTCCCAGCACAATACCTATACAAAATGGATGACACTGCTT
ACTTCATGAACA

CTGACACTTATGATCAGTACGAAATTCAGTTGCTAACGTTGAGCAAGAATTGCTTTACATTCTTGAAAACCTC
AGACGTGAAAAT

CCAATTTTATGGAAGTGAAGTGATTGGGGTAACGGTTCCAACAACGTGTTGAATTGACCGTTGCGGAAACACAA
CCATCTATTTAA

GGAGCGACAGTGACGGGTTGAGGAAACCTGCAACTCTTGAGACAGGACTTGTTGTTAACGTTCCAGACTTTA
TCGAAGCTGG

CCAAAACTAATCATTAACACTGCAGAAGGTACTTACGTTTCTCGTGCTTAA

SPy1916

Seq ID 77

ATGACTAAAACATTACCTAAAGATTTTATTTTTGGTGGTGCTACAGCTGCTTACCAGGCTGAAGGCGCTACCC
ACACAGATGGTA

AAGGACCAGTAGCTTGGGATAAAATACTTAGAAGACAACCTATTGGTACACAGCTGAGCCAGCAAGTGATTTTTTA
TAATCGTTACCC

TGTCGATTTGAAACTTAGTGAAGAATTTGGTGTCAACGGCATCCGTATCTCTATTGCCTGGTCTCGTATTTTT
CCAACAGGAAAA

GGAGAAGTTAACCTAAAGGAGTAGAATACTACCACAATCTTTTGCAGAGTGTATAAGCGTCATGTTGAGC
CTTTTGTTACAC

TTCACCATTTTGATACCCAGAAGCTCTCCACTCGGATGGTGACTTCCTCAATCGTGAGAACATTGAACATTT
TGTAATTATGC

AGAATTTTGTTTTAAAGAATTCTCAGAAGTTAACTATTGGACAACATTTAACGAAATTGGGCCTATTGGTGAT
GGCCAATACTTAG

TTGGTAAATTCCCTCCAGGTATCCAATATGATCTTGCTAAAAGTTTTCCAATCACACCATAACATGATGGTCTC
TCATGCTCGTGCA

GTCAAACTCTTTAAAGATAGTGGTTATTCAGGTGAAATTGGTGTGTCCATGCACTTCCAATAAGTATCCAT

TTGACGCTAACAA

TCCTGATGATGTTAGAGCAGCTGAACTTGAAGATATCATCCATAATAAATTTATCCTTGATGCTACTTATCTT
GGTAAGTATTTCAG

ATAAAACAATGGAAGGTGTTAACCATATCCTTGAGGTGAATGGCGGTGAACTTGATCTTCGCGAAGAAGATTT
TGCCGCACTAG

ACGCCGCAAAAGATTTGAATGATTTCCCTTGGTATTAATACTACTATATGAGTGATTGGATGCAAGCTTTTGATGG
TGAGACTGAAAT

CATTACAATGGCAAGGGTAAAAAGGCAGCTCTAAATACCAAATCAAGGGTGTGGTGAAGAAAAGCACCC
GTTGATGTTCC

AAAAACGGACTGGGACTGGATTATCTTCCCACAAGGCTTATATGATCAAATCATGCGTGTCAAAGCCGATTAT
CCTAATTACAAG

AAAATTTACATTACAGAGAATGGTCTTGGCTACAAAGATGAGTTTGTAGATAATACTGTCTATGATGGTGGAC
GTATCGATTATGT

GAAAAACACTTAGAAGTTATTTCTGATGCTATTTCTGATGGTGCAAATGTTAAAGGATACTTTATGTGGTCA
CTGATGGATGTCT

TTTCATGGTCAAATGGCTATGAAAAACGTTACGGTCTCTTCTATGTTGATTTTGAACTCAAGAACGTTATCC
TAAGAAGAGTGC

CTACTGGTATAAAAAAGTAGCAGAACTCAAGTGATTGAATGA

SPy1972

Seq ID 78

ATGAAAAAGAAAGTCAACCAAGGATCAAAGCGCTATCAATATCTGTTAAAAAAGTGGGGGATAGGTTTTGTAA
TCGCTGCAACTG

GGACTGTCGTGTTAGGGTGCACCCCTAGTATCTTAACACATCAAGTTGCTGCTAAAACCATTTGTTGGACTAGC
CCGCGATGAAG

CTCAACAAGGAGATGGCAATGCTAAATCTGGTGATGGTCTTCAATCGTCTAGCAAGGAGGCAAAACCAGTTTT
AGACAGCTCGT

CAGCTAATCCTGCTAGTATTGCTGAGCATCATTTGCGTATGCATTTTAAAACATTGCCAGCTGGTGAGTCGCT
AGGAAGCTTGG

GACTTTGGGTGTGGGGAGATGTGGATCAACCTTCAAAGGATTGGCCAAATGGTGCTATCACCATGACAAAAGC
GAAAAAAGAT

GACTATGGCTATTATCTAGATGTGCCACTAGCAGCTAAACACCGCCAGCAAGTGTCTTATCTCATTATAATA
AAGCTGGAGAGA

ATCTTTCAAAGGACCAGCACATCTCGCTTCTCACGCCAAAAATGAATGAAGTTTGGATAGACGAGAATTACCA
TGCGCACGCTT

ATCGACCTTTGAAAAAAGGTACCTTCGAATCAACTACCACAATCAATCGGGACACTACGATAACTTAGCTGT
CTGGACCTTTAA

AGATGTCAAAACCCCAACGACCGACTGGCCAAATGGACTTGACTTGTCACATAAAGGGCATTATGGAGCTTAT
GTTGATGTCCC

CTTAAAGAAGGAGCTAACGAAATCGGATTTTTAATCCTTGATAAAAGTAAGACAGGAGATGCTATTAAAGTG
CAACCAAAGAT
TATCTATTTAAAGAGTTAGACAATCATACTCAGGTTTTGTCAAAGACACTGACCCAAAAGTTTACAACAATC
CTTATTATATTGAT
CAGGTTAGTCTCAAAGGAGCTGAACAAACCACGCCAAATGAGATTAAAGCCATTTTTACGACCTTAGATGGGC
TTGATGAAGAT
GCGGTGAAACAAAACATCAAGATCACTGACAAAGCAGGGAAAACTGTTGCAATTGATGAGTTGACACTTGACA
GGGATAAGTCT
GTAATGACATTAAAGGGTGATTTTAAGGCGCAAGGTGCAGTCTACACGGTTACATTTGGAGAAGTTAGCCAAG
TCGCTCGCCAA
TCCTGGCAATTTAAAGATAAACTCTATGCTTACGATGGTGAACCTGGAGCTACCCTAGCTAAGGATGGTTCTG
TTGATTTAGCGC
TATGGTCTCCAAGTGCTGATACTGTTAAGGTTGTCGTTTATGATAAACAAGATCAGACAAGGGTGGTTGGTCA
AGCTGATTTGAC
CAAGTCGGACAAGGGTGTTTGGAGAGCTCATCTAACTTCTGACAGTGTCAAGGGCATTAGTGATTACACAGGC
TACTATTACCT
TTATGAAATCACGCGCGGTCAGGAAAAAGTCATGGTTTTGGATCCTTACGCCAAATCTCTCGCTGCCTGGAAT
GATGCGACTGC
TACTGATGACATCAAAACAGCAAAAGCTGCCTTTATTGATCCAAGCAAAGTAGGACCAACAGGCCTTGATTTT
GCCAAAATTAAC
AACTTTAAAAAGCGTGAAGACGCTATTATCTATGAAGCACATGTGCGAGATTTTACGTCAGATAAGGCTCTAG
AAGGCAAGTTAA
CACACCCTTTTGGGACTTTTTCAGCTTTCGTTGAACAGCTAGACTATCTCAAAGACTTGGGGGTACCCACGT
TCAATTGCTACC
GGTTTTGAGTTATTTTTATGCCAATGAGCTGGACAAGAGCCGCTCAACAGCCTACACGTCTTCAGACAATAAT
TACAACTGGGGT
TATGACCCACAACACTACTTTGCCCTTTCTGGCATGTATTCGGCAAATCCTAATGACCCTGCTTTACGTATCG
CAGAGCTTAAAA
ACCTTGTC AATGAGATTCACAAACGTGGTATGGGTGTTATTTTTGATGTGGTTTATAACCACACGGCTAGAAC
CTATCTCTTTGA
AGATTTGGAACCCAACTACTATCATTTTATGAATGCTGATGGTACAGCTAGAGAGAGTTTTGGCGGAGGTCGT
CTAGGAACGAC
ACATGCCATGAGTCGTCGTATCTTGGTGGATTCGATTACTTATCTGACTCGTGAATTCAAGGTAGATGGTTTT
CGTTTCGACATG
ATGGGTGACCATGATGCGGCAGCTATTGAGCAAGCCTTTAAGGCAGCCAAAGCCATTAATCCAAATACCATTA
TGATTGGCGAA
GGCT GGCGTACCTACCAAGGTGATGAGGGGAAAAAAGAAATTGCGGCAGATCAAGATTGGATGAAAGCAACC
AATACGGTCGG
TGTTTTCTCTGATGATATCAGAAATACCCTCAAGTCAGGTTTTCCAAATGAAGGCACAGCAGCCTTTATTACT

GGTGGCGCAAAA

AATCTAGAAGGTTTATTCAAAACGATCAAAGCACAGCCTGGTAACTTTGAAGCAGATGCCCCAGGAGATGTAG
TGCAGTATATT

GCAGCCCATGACAACCTGACCTTACATGATGTCATTGCCAAATCCATCAATAAGGATCCTAAAGTGGCTGAAG
AAGAGATTAC

AAGCGTATTCGTCTAGGAAATACCATGATTTTAACTGCTCAAGGGACTGCCTTTATCCATTCTGGTCAGGAAT
ATGGACGAACCA

AGCAGCTTCTAAATCCCGACTACAAGACAAAAGCGTCTGATGACAAGGTGCCAAATAAGGCGACTCTGATTGA
TGCTGTAGCG

CAATACCCTTACTTCATCCACGATTCTTATGATTTCGTCTGATGCGGTCAATCATTTTGACTGGGCAAAGGCAA
CAGATTCCATAG

CTCACCCGATTAGCAACCAAAACAAAAGCCTATACACAGGGACTAATTGCGTTGCGTCGCTCAACAGATGCCTT
TACAAAAGCAA

CCAAAGCTGAGGTAGATCGGGATGTGACCTTGATCACCCAAGCAGGACAAGATGGTATTCAACAAGAGGACCT
CATCATGGGT

TACCAAACAGTGGCATCAAATGGAGATCGCTATGCTGTCTTTGTCAATGCAGACAACAAGACCCGCAAGGTAG
TTTTACCTCAA

GCCTACCGCTATTTGCTAGGAGCCCAAGTGCTTGTTGATGCTGAGCAAGCTGGTGTTACTGCCATTGCTAAGC
CTAAGGGAGT

CCAGTTTACCAAAGAAGGCTTGACTATTGAAGGCCTAACTGCCCTGGTCCTCAAAGTATCCTCAAAAACGGCT
AATCCCTCTCA

GCAAAAGAGTCAGACAGACAATCATCAAACCAAAAACACCAGATGGCTCAAAAGACCTAGACAAATCATTAATG
ACTAGACCAA

AAGAGCTAAACAAACCAAAAGCTCCCAAAAACGGGTGAAGCCTCCTCAAAAGGCTTATTAGCAGCTGGAATA
GCTCTGCTTTT

ATTGGCTATTAGCCTGTTGATGAAGCGCCAAAAAGATTAG

SPy1979

Seq ID 79

ATGAAAAATTACTTATCTATTGGAGTGATTGCACTGCTGTTTGCAATTAACATTTGGAACAGTCAAGTCGGTCC
AAGCTATTGCTG

GGTATGGATGGCTACCAGACCGTCCACCTATCAATAACAGCCAGTTAGTTGTTAGTATGGCCGGTATCGTTGA
AGGTACCGATA

AAAAAGTTTTTATAAATTTTTTTGAAATCGATCTAACATCACAACCTGCTCACGGAGGAAAGACAGAGCAGGG
CTTAAGTCCAAA

ATCAAAACCATTTGCTACAGATAATGGCGCAATGCCACATAAACTTGAAAAAGCTGACTTATTAAGCTATT
CAAAAACAGCTG

ATCGCTAACGTTACAGTAACGACGGCTACTTTGAGGTCATTGATTTTGCAAGCGATGCAACCATTACTGATC
GAAACGGCAAG

GTCTACTTTGCTGACAAAGATGGTTCGGTAACCTTGCCGACCCAACCTGTCCAAGAATTTTTGTTAAAGGGAC
ATGTGCGCGTT

AGACCATATAAAGAAAAACCAGTACAAAATCAAGCAAAATCTGTTGATGTAGAATATACTGTACAGTTTACTC
CTTTAAACCCTGA

TGACGATTTTCAGACCAGGGCTCAAAGATACTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCACATCT
CAAGAATTACTA

GCTCAAGCACAAAGCATTTTAAACAAAAACCCACCCAGGCTATACGATTTATGAACGTGACTCCTCAATCGTCA
CTCATGACAATG

ACATTTTCCGTACGATTTTACCAATGGATCAAGAGTTTACTTACCATGTCAAAAATCGGGAACAAGCTTATGA
GATCAATCCTAAA

ACAGGTATTAAAGAAAAAACGAACAACACTGATCTGGTCTCTGAGAAATATTACGTCCTTAAACAAGGGGAAA
AGCCGTATGATC

CCTTTGATCGCAGTCACTTGAAACTGTTCAACCATCAAATACGTTGATGTCAACACCAACGAATTGCTAAAAAG
CGAGCAGCTCTT

AACAGCTAGCGAACGTAACCTTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTACTCTACAAC
AATCTCGATGCT

TTTGATATCATGGACTATACCTTAACTGGAAAAAGTAGAGGATAATCACGATAAGAATAATCGTGTGTTACAG
TTTATATGGGCA

AGCGCCCTAAAGGGGCAAAGGGTAGCTATCATTTAGCTTATGATAAAGATCTCTATACCGAAGAAGAACGAAA
AGCTTACAGCT

ACCTGCGTGATACAGGGACACCTATACCTGATAACCTAAAGACAAATAA

SPy1983

Seq ID 80

ATGTTGACATCAAAGCACCATAATCTCAACAAACTAGTCTGGCGCTACGGGCTAACCTCAGCCGCTGCCGTCC
TTCTAGCCTTT

GGAGGCGGGGCAAGCAGCGTTAAGGCTGAGGTTTCTTCTACGACTATGACGTCGAGTCAAAGAGAGTCAAAAA
TAAAAGAGAT

CGAAGAAAGTCTTAAAAAATATCCAGAAGTGCCAATGAGAAATTTTGGGAAAGAAAGTGGTATGGAACCTAT
TTTAAAGAAGAA

GATTTTCAAAGGAGCTAAAAAGATTTTACTGAGAAGAGGCTTAAGGAGATTCTAGATTTAATTGGTAAATCTG
GAATCAAGGGAG

ACCGTGGTGAGACTGGTCCTGCTGGCCCAGCCGACCACAAGGTAAAACTGGTGAGAGGGGCGCCCAAGGTCC
TAAAGGTGA

CCGCGGTGAGCAAGGAATCCAAGGTAAAGCTGGTGAAAAAGGTGAGCGCGGTGAAAAAGGCGACAAAGGTGAA
ACCGGTGAA

CGCGGTGAAAAAGGCGAAGCTGGAATCCAAGGCCCAAGGTGAAGCTGGTAAAGATGGCGCTCCAGGTAAAG
ATGGAGCTC

CAGGCGAAAAGGGTGAAAAAGGTGACCGCGGTGAAACCGGAGCTCAGGGTCCAGTAGGCCCAAGGTGAAAA

AGGTGAAAC

GGGCGCCCAAGGCCAGCAGGCCACAAGGTGAGGCAGGCAAACCAGGTGAGCAAGGCCAGCAGGCCACA
GGTGAAG

CAGGCCAACAGGCGAAAAAGCTCCAGAAAAAGAGCCAGAAAGGCGAAGCAGGCCAACAGGCGAAAAAGCTCC
AGAAAAGAG

CAAAGAGGTAAGTCCAGCTGCAGAAAAACCTGCTGACAAAGAAGCTAACCAAACGCCAGAACGCCGCAATGGC
AATATGGCTA

AGACACCTGTAGCCAACAACACAGACGTCTACCAGCAACTGGTGAGCAAGCCAACCCATTCTTTACAGCAGC
AGCAGTAGCA

GTGATGACAACAGCTGGTGTCTAGCCGTTACAAAACGCAAAGAAAACAACTAA

SPy1991

Seq ID 81

ATGATACTCTTAATTGATAATTACGATTCATTTACCTACAACCTCGCCCAATATTTAAGTGAATTTGACGAGA
CGATTGTCTTGAT

AACCAAGACCCAACTTATATGACATGGCCAAAAAGCTAACGCTCTAGTCCTCTCACCTGGTCCTGGTTGGC
CCAAGGAAGCC

AACCAATGCCAAACTCATTCAAGACTTTTACCAAACAAACCTATCTTAGGAGTGTGTCTGGGACACCAAG
CTATCGCTGAAA

CTTTAGGGGGAACCTTACGCTTGGCCAAACGCGTCATGCATGGGAGACAAAGCACCATTGAAACGCAAGGCC
TGCTAGTCTT

TTTCGCTCCCTGCCACAAGAGATCACCGTCA TGCGCTACCATTCCATCGTTGTGGATCAGTTACAAAAGGTT
TTAGCGTAACC

GCTAGAGACTGTGACGATCAAGAAATCATGGCATTGTAACACCACACCCTGCCACTTTTTGGGCTACAATTC
ACCCAGAAAGC

ATCGGAACTCCTGATGGCATGACCATGATTGCCAACTTCATCGCAGCCATTCCCCGTAA

SPy2000

Seq ID 82

GTGTCAAAATACCTAAAATACTTCTCTATTATCACGTTATTTTGGACTGGGCTTATTTAGTTGCATGTCAAC
AACAAAAGCCTCAA

ACAAAAGAACGTCAGCGCAAAACAGTCCAAAAGACGAACTTGTCGTTTCTATGGGGGCAAAGCTCCCTCATG
AATTCGATCCA

AAGGACCGTTATGGAGTCCACAATGAAGGGAATATCACTCATAGCACTCTATTGAAACGTTCTCCTGAACTAG
ATATAAAAGGAG

AGCTTGCTAAACATACCATCTCTCTGAAGATGGGCTGACTTGGTCGTTTGACTTGCATGATGATTTTAAATT
CTCAAATGGTGA

GCCTGTTACTGCTGATGATGTTAAGTTTACTTATGATATGTTGAAAGCAGATGGAAAGGCTTGGGATCTAACC
TTCATTAAGAAC

GTTGAAGTAGTTGGGAAAAATCAGGTCAATATCCATTTGACTGAGGCGCATTCGACATTTACAGCACAGTTGA
CTGAAATCCCAA
TCGTCCCTAAAAAACATTACAATGATAAGTATAAGAGCAATCCTATCGGTTTCAGGACCTTACATGGTAAAAGA
ATATAAGGCTGG
AGAACAAGCTATTTTTGTTGTAACCTTATTGGCATGGGAAAAAACCATCTTTAAAAAATGGACTTGGGTC
TTACTTGATGAAA
ACACAGCACTAGCAGCTTTAGAATCTGGTGATGTTGATATGATCTACGCAACGCCAGAACTTGCTGATAAAAA
AGTCAAAGGCA
CCCGCCTCCTTGATATTCCATCAAATGATGTGCGCGGCTTATCATTACCTTATGTGAAAAAGGGCGTCATCAC
TGATTCTCCTGA
TGGTTATCCTGTAGGAAATGATGTCACTAGTGATCCAGCAATCCGAAAAGCCTTGACTATTGGTTTAAATAGG
CAAAAAGTTCTC
GATACGGTTTTAAATGGTTATGGTAAACCAGCTTATTCAATTATTGATAAAACACCATTTTGAATCCAAAA
CAGCCATTAAAGA
TAATAAGTAGCTAAAGCTAAGCAATTATTGACAAAAGCGGGATGGAAAGAACAAGCAGACGGTAGCCGTAAA
AAAGGTGACCT
TGATGCAGCGTTTGATCTGTACTACCCTACTAATGATCAATTGCGAGCGAACTTAGCCGTTGAAGTAGCAGAG
CAAGCCAAGGC
CCTAGGGATTACTATTAAACTCAAAGCTAGTAACTGGGATGAAATGGCAACGAAGTCACATGACTCAGCCTTA
CTTTATGCCGG
AGGACGTCATCACGCGCAGCAATTTTATGAATCGCATCATCCAAGCCTAGCAGGGAAAGGTTGGACCAATATT
ACGTTTTATAA
CAATCCTACCGTGACTAAGTACCTTGACAAAGCAATGACATCTTCTGACCTTGATAAAGCTAACGAATATTGG
AAGTTAGCGCAG
TGGGATGGCAAAACAGGTGCTTCTACTCTTGGAGATTGCCAAATGTATGGTTGGTGAGCCTTAACCATACTT
ATATTGGTGATA
AACGTATCAATGTAGGTAAACAAGGCGTCCACAGTCATGGTCATGATTGGTCATTATTGACTAACATTGCCGA
GTGGACTTGGG
ATGAATCAACTAAGTAA

SPy2006

Seq ID 83

GTGAAGAAAACATATGGTTATATCGGCTCAGTTGCTGCTATTTTACTAGCTACTCATATTGGAAGTTACCAAC
TTGGTAAGCATC
ATATGGGTTTCAGCAACAAAGGACAATCAAATTGCCTATATTGATGATAGCAAAGGTAAGGCAAAAGCCCCTAA
AACAAACAAAAC
GATGGATCAAATCAGTGCTGAAGAAGGCATCTCTGCTGAACAGATCGTAGTCAAAATTACTGACCAAGGCTAT
GTGACCTCACA
TGGTGACCATTATCATTTTTACAATGGGAAAGTTCCTTATGATGCGATTATTAGTGAAGAGTTGTTGATGACG

GATCCTAATTACC

GTTTTAAACAATCAGACGTTATCAATGAAATCTTAGACGGTTACGTTATTTAAAGTCAATGGCAACTATTATGT
TTACCTCAAGCCA

GGTAGCAAGCGCAAAAACATTCGAACCAAAACAACAAATTGCTGAGCAAGTAGCCAAAGGAAGCTAAAGAAGCTA
AAGAAAAAGGT

TTAGCTCAAGTGGCCCATCTCAGTAAAGAAGAAGTTGCGGCAGTCAATGAAGCAAAAAGACAAGGACGCTATA
CTACAGACGAT

GGCTATATTTTTAGTCCGACAGATATCATTGATGATTTAGGAGATGCTTATTTAGTACCTCATGGTAATCACT
ATCATTATATTCCT

AAAAAGGATTTGTCTCCAAGTGAGCTAGCTGCTGCACAAGCCTACTGGAGTCAAAAACAAGGTCGAGGTGCTA
GACCGTCTGAT

TACCGCCCGACACCAGCCCCAGCCCCAGGTCGTAGGAAAGCCCCAATTCCTGATGTGACGCCTAACCTGGAC
AAGGTCATCA

GCCAGATAACGGTGGCTATCATCCAGCGCCTCCTAGGCCAAATGATGCGTCACAAAACAAACACCAAAGAGAT
GAGTTTAAAG

GAAAAACCTTTAAGGAACCTTTTAGATCAACTACACCGTCTTGATTTGAAATACCGTCATGTGGAAGAAGATGG
GTTGATTTTTGA

ACCGACTCAAGTGATCAAATCAAACGCTTTTGGGTATGTGGTGCCTCATGGAGATCATTATCATATTATCCCA
AGAAGTCAGTTA

TCACCTCTTGAAATGGAATTAGCAGATCGATACTTAGCCGGCCAACTGAGGACGATGACTCAGGTTTCAGATC
ACTCAAAACCA

TCAGATAAAGAAGTGACACATACCTTTCTTGGTCATCGCATCAAAGCTTACGGAAAAGGCTTAGATGGTAAAC
CATATGATACGA

GTGATGCTTATGTTTTTAGTAAAGAATCCATTCATTCAAGTGGATAAATCAGGAGTTACAGCTAAACACGGAGA
TCATTTCCACTAT

ATAGGATTTGGAGAACTTGAACAATATGAGTTGGATGAGGTCGCTAACTGGGTGAAAGCAAAAGGTCAAGCTG
ATGAGCTTGCT

GCTGCTTTGGATCAGGAACAAGGCAAAAGAAAAACCACTCTTTGACACTAAAAAAGTGAGTCGCAAAGTAACAA
AAGATGGTAAA

GTGGGCTATATGATGCCAAAAGATGGCAAGGACTATTTCTATGCTCGTGATCAACTTGATTTGACTCAGATTG
CCTTTGCCGAAC

AAGAATAATGCTTAAAGATAAGAAACATTACCGTTATGACATTGTTGACACAGGTATTGAGCCACGACTTGC
TGATAGATGTGTC

AAGTCTGCCGATGCATGCTGGTAATGCTACTTACGATACTGGAAGTTCGTTTGTATCCCTCATATTGATCAT
ATCCATGTCGTT

CCGTATTCATGGTTGACGCGCGATCAGATTGCAACAATCAAGTATGTGATGCAACACCCCGAAGTTCGTCCGG
ATATATGGTCT

AAGCCAGGGCATGAAGAGTCAGGTTCGGTCATTCCAAATGTTACGCCTCTTGATAAACGTGCTGGTATGCCAA
ACTGGCAAATT

ATCCATTCTGCTGAAGAAGTTCAAAAAGCCCTAGCAGAAGGTCGTTTTGCAACACCAGACGGCTATATTTTCG
ATCCACGAGAT
GTTTTGGCCAAAGAACTTTTGTATGGAAAGATGGCTCCTTTAGCATCCCAAGAGCAGATGGCAGTTCATTGA
GAACCATTAATA
AATCTGATCTATCCCAAGCTGAGTGGCAACAAGCTCAAGAGTTATTGGCAAAGAAAAACGCTGGTGATGCTAC
TGATACGGATA
AACCCAAAGAAAAGCAACAGGCAGATAAGAGCAATGAAAACCAACAGCCAAGTGAAGCCAGTAAAGAAGAAGA
AAAAGAATCA
GATGACTTTATAGACAGTTTACCAGACTATGGTCTAGATAGAGCAACCCTAGAAGATCATATCAATCAATTAG
CACAAAAAGCTA
ATATCGATCCTAAGTATCTCATTTTCCAACCAGAAGGTGTCCAATTTTATAATAAAAATGGTGAATTGGTAAC
TTATGATATCAAG
ACACTTCAACAAATAAACCCTTAA

SPy2009

Seq ID 84

ATGCGTAGAGCAGAAAAATAACAAACACAGCCGCTATTCCATTCGCAAACCTGAGCGTTGGGGTAACGAGTATAG
CAATTGCGAGT
CTCTTTTTTAGGAAAGGTTGCCTATGCCGTAGATGGCATCCCTCCAATCTCTCTTACTCAAAAGACTACAGCCA
CTACATCAGAAA
ATTGGCATCATATTGATAAGGATGGCCTTATTCCTTTAGGTATAAGCTTAGAAGCTGCCAAAGAGGAATTTAA
AAAAGAAGTAGA
AGAATCACGTTTATCTGAAGCACAAAAAGAAACGTATAAACAAAAAATTAAACTGCACCAGACAAAGATAAG
CTATTATTACGT
ATCATAGTGAGTATATGACAGCCGTTAAGGATCTTCCAGCGTCTACTGAGTCTACTACTCAGCCAGTTGAGGC
ACCCGTGCAGG
AGACACAGGCATCAGCTTCAGATTCGATGGTGACAGGTGATTCAACATCAGTTACGACTGATTCTCCTGAGGA
AACCCCATCTT
CGGAAAGTCCAGTGGCCCCAGCTTTATCTGAGGCTCCAGCTCAACCAGCTGAGAGTGAGGAACCTTCAGTAGC
AGCATCTTCT
GAGGAAACCCCATCTCCATCAACTCCAGCGGCCCCAGAACTCCTGAAGAACCAGCAGCTCCATCTCCATCAC
CTGAGAGTGA
GGAACCTTCAGTAGCAGCTCCTTCTGAGGAAACCCCATCTCCAGAACTCCTGAAGAACCAGCAGCTCCATCT
CAACCAGCTGA
GAGTGAAGAATCTTCAGTAGCAGCTACGACAAGCCCGTCTCCATCAACTCCAGCTGAATCAGAGACTCAGACG
CCACCAGCTG
TTAATAAGACTCTGATAAGCCATCTTCAGCAGCTGAAAAACCAGCAGCCTCTTCACTTGTTTCAGAACAAAC
CGTTCAACAACC
AACTTCAAAGAGATCTTCTGATAAAAAAGAAGAGCAAGAACAGTCTTACTCTCCAAATCGCTCATTTGTCAAGA

CAGGTTAGGGCC

CATGAGTCAGGTAAGTACTTGCCTTCAACAGGTGAAAAAGCACAGCCACTCTTTATAGCTACTATGACTTTGA
TGTCTCTATTTG

GCAGTCTTTTAGTCACAAAACGCCAAAAAGAACTAAAAAATAG

SPy2010

Seq ID 85

TTGCGTAAAAAACAAAAATTACCATTTGATAAACTTGCCATTGCGCTCATGTCTACGAGCATCTTGCTCAATG
CACAATCAGACAT

TAAAGCAAATACTGTGACAGAAGACACTCCTGCTACCGAACAAGCTGTAGAAACCCACAAACCAACAGCGGTT
TCTGAGGAAGC

ACCATCATCAAAGGAAACCAAAACCCACAACTCCTGATGACGCAGAAGAAACAATAGCAGATGACGCTAAT
GATCTAGCCCC

TCAAGCTCCTGCTAAAACTGCTGATACACCAGCAACCTCAAAAGCGACTATTAGGGATTTGAACGACCCTTCT
CAGGTCAAAAC

CCTGCAGGAAAAAGCAGGCAAAGGAGCTGGGACTGTTGTTGCAGTGATTGATGCTGGTTTTGATAAAAAATCAT
GAAGCGTGGC

GCTTAACAGACAAAACCAAAGCACGTTACCAATCAAAAGAAGATCTTGAAAAAGCTAAAAAGAGCACGGTAT
TACCTATGGCG

AGTGGGTCAATGATAAGGTTGCTTATTACCACGACTATAGTAAAGATGGTAAAACCGCTGTCGATCAAGAGCA
CGGCACACACG

TGTCAGGGATCTTGTCAGGAAATGCTCCATCTGAAACGAAAGAACCTTACCGCCTAGAAGGTGCGATGCCTGA
GGCTCAATTG

CTTTTGATGCGTGTCGAAATTGTAAATGGACTAGCAGACTATGCTCGTAACTACGCTCAAGCTATCATAGATG
CTGTCAACTTGG

GAGCTAAGGTGATTAATATGAGCTTTGGTAATGCTGCACTAGCCTATGCCAACCTTCCAGACGAAACCAAAAA
AGCCTTTGACTA

TGCCAAATCAAAAGGTGTTAGCATTGTGACCTCAGCTGGTAATGATAGTAGCTTTGGGGGCAAGACCCGTCTA
CCTCTAGCAGA

TCATCCTGATTATGGGGTGGTTGGGACACCTGCAGCGGCAGACTCAACATTGACAGTTGCTTCTTACAGCCCA
GATAAACAGCT

CACTGAACTGCTACGGTCAAAACAGCCGATCAGCAAGATAAAGAAATGCCTGTTCTTTCAACAAACCGTTTT
GAGCCAAACAA

GGCTTACGACTATGCTTATGCTAATCGTGGGATGAAAGAGGATGATTTTAAGGATGTCAAAGGTAAGATTGCC
CTTATTGAACGT

GGCGATATTGATTTCAAAGATAAGATTGCAAACGCTAAAAAAGCTGGTGCTGTAGGAGTCTTGATCTATGACA
ATCAGGACAAG

GGCTTCCCGATTGAATTGCCAAATGTTGATCAGATGCCTGCGGCCTTTATCAGTCGAAAAGATGGTCTCTTAT
TAAAAGAGAATC

CCCCAAAAACCATCACCTTCAATGCGACACCTAAGGTATTGCCAACAGCAAGTGGCACCAACTAAGCCGCTT
CTCAAGCTGGG

GTCTGACAGCTGACGGCAATATTAAGCCAGATATTGCAGCACCCGGCCAAGATATTTTGTCAATCAGTGGCTAA
CAACAAGTATG

CCAAACTTTCTGGAAGTAGTATGTCTGCGCCATTAGTAGCGGGTATCATGGGACTGTTGCAAAAGCAATATGA
GACACAGTATC

CTGATATGACACCATCAGAGCGTCTTGATTTAGCTAAAAAAGTATTGATGAGCTCAGCAACTGCCTTATATGA
TGAAGATGAAAA

AGCTTATTTTTCTCCTCGCCAACAAGGAGCAGGAGCAGTCGATGCTAAAAAAGCTTCAGCAGCAACGATGTAT
GTGACAGATAA

GGATAATACCTCAAGCAAGGTTACCTGAACAATGTTTCTGATAAATTTGAAGTAACAGTAACAGTTCACAAC
AAATCTGATAAAC

CTCAAGAGTTGTATTACCAAGCAACTGTTCAAACAGATAAAGTAGATGGAAAACCTTTGCCTTGGCTCCTAA
AGCATTGTATGA

GACATCATGGCAAAAAATCACAATTCCAGCCAATAGCAGCAAACAAGTCACCATTCCAATCGATGTTAGTCAA
TTTAGCAAGGAC

TTGCTTGCCCCAATGAAAAATGGCTATTTCTTAGAAGGTTTGTTCGTTTCAAACAAGATCCTACAAAAGAAG
AGCTTATGAGTAT

TCCCTATATTGGTTTCCGAGGTGATTTTGGCAATCTGTCAGCCTTAGAAAAACCAATCTATGATAGCAAAGAC
GGTAGCAGCTAC

TATCATGAAGCAAATAGTGATGCCAAAGACCAATTAGATGGTGATGGATTACAGTTTTACGCTCTGAAAAATA
ACTTTACAGCAC

TTACTACAGAGTCTAATCCATGGACGATTATTAAAGCTGTCAAAGAAGGGGTGAAAACATAGAGGATATCGA
ATCTTCAGAGAT

CACAGAAACCATTTTTGCAGGTACTTTTGCAAAACAAGACGATGATAGCCACTACTATATCCACCGTCACGCT
AATGGCAAGCCA

TATGCTGCGATCTCTCAAATGGGGACGGTAACAGAGATTATGTCCAATTCCAAGGTACTTTCTTGCGTAATG
CTAAAAACCTTG

TGGCTGAAGTCTTGGACAAAGAAGGAAATGTTGTTTGGACAAGTGAGGTAACCGAGCAAGTTGTTAAAACTA
CAACAATGACT

TGGCAAGCACACTTGGTTCAACCCGTTTTGAAAAAACGCGTTGGGACGGTAAAGATAAAGACGGCAAAGTTGT
TGCTAACGGAA

CATACACCTATCGTGTTGCTACACTCCGATTAGCTCAGGTGCAAAAGAACAACACACTGATTTTGATGTGAT
TG TAGACAATAC

GACACCTGAAGTCGCAACATCGGCAACATTCTCAACAGAAGATCGTCGTTTGACACTTGCATCTAAACCAAAA
ACCAGCCAACC

GGTTTACCGTGAGCGTATTGCTTACACTTATATGGATGAGGATCTGCCAACAAACAGAGTATATTTCTCAAAT
GAAGATGGTACC

TTTACTCTTCTGAAGAGGCTGAAACAATGGAAGGCGCTACTGTTCCATTGAAAATGTCAGACTTTACTTATG

TTGTTGAAGATA

TGGCTGGTAACATCACTTATACACCAGTGACTAAGCTATTGGAAGGCCACTCTAATAAACAGAACAAAGACGG
TTCAGATCAAG

CACCAGACAAAAAACAGAACTAAACCAGAACAAAGACGGTTCAGGTCAAGCACCAGATAAAAAACCAGAAAC
TAAACCAGAAC

AAGACGGTTCAGGTCAAACACCAGACAAAAAACAGAACTAAACCAGAACAAAGACGGTTCAGGTCAAACACC
AGATAAAAAAC

CAGAACTAAACCAGAAAAAGATAGTTCAGGTCAAACACCAGGTAAACTCCTCAAAAAGGTCAACCTTCTCG
TACTCTAGAGAA

ACGATCTTCTAAGCGTGCTTTAGCTACAAAAGCATCAACAAAAGATCAGTTACCAACGACTAATGACAAGGAT
ACAAATCGTTTA

CATCTCCTTAAGTTAGTTATGACCACTTCTTCTTGGGATTAGTAGCTCATATCTTTAAACAAAACGCACTG
AAGATTAG

SPy2016

Seq ID 86

ATGAATATTAGAAATAAGATTGAAAAATAGTAAACACTACTATTTACATCCCTTG TAGCCGTGGCTCTACTAG
GAGCTACACAACC

AGTTTCAGCCGAAACGTATACATCACGCAATTTTGACTGGTCTGGAGATGACTGGTCTGGAGATGACTGGCCT
GAAGATGACTG

GTCTGGAGATGGTTTGTCTAAATATGACCGGTCTGGAGTTGGTTTGTCTCAATATGGCTGGTCTAAATATGGC
TGGTCTAGCGA

TAAAGAAGAATGGCCTGAAGATTGGCCTGAAGATGACTGGTCTAGCGATAAAAAAGATGAGACAGAAGATAAA
ACGAGACCACC

ATATGGAGAAGCATTAGGTACAGGGTATGAAAAACGTGATGATTGGGGAGGACCTGGTACGGTGGCAACTGAC
CCTTACACTC

CACCATATGGAGGAGCATTAGGTACAGGGTATGAAAAACGTGATGATTGGGGAGGACCTGGTACGGTGGCAAC
TGACCCTTAC

ACTCCACCATATGGAGGAGCATTAGGTACAGGGTATGAAAAACGTGATGATTGGAGAGGACCTGGACATATTC
CTAAACCTGAG

AACGAACAATCACCAAAACCACTTCATATTCCTGAACCTCCTCAGATTGAGTGGCCTCAGTGGAATGGCTTTG
ATGGATTATCAT

TTGGCCCCTCTGATTGGGGCCAATCTGAGGACACCCCTCCAAGTGAACCTCGTGTGCCAGAAAAACCGCAACA
TACTCCTCAA

AAAAATCCACAAGAATCAGATTTTGATAGAGGGTTTCAGCTGGCTTGAAAGCAAAAACTCAGGTAGAGGTA
TTGATTTTGAAG

GTTTCCAGTATGGTGGCTGGTCAGACGAATATAAAAAAGGTTACATGCAAGCCTTCGGTACACCATATACACC
ATCAGCAACGT

AA

SPy2018
Seq ID 87
ATGGCTAAAAATAACACGAATAGACACTATTCGCTTAGAAAAATTAAAAACAGGAACGGCTTCAGTAGCGGTAG
CTTTGACTGTTT
TAGGGGCAGGTTTTGCGAATCAAACAGAGGTTAAGGCTAACGGTGATGGTAATCCTAGGGAAGTTATAGAAGA
TCTTGCAGCAA
ACAATCCCGCAATACAAAATATACGTTTACGTTACGAAAAACAAGGACTTAAAAGCGAGATTAGAGAATGCAAT
GGAAGTTGCAG
GAAGAGATTTTAAGAGAGCTGAAGAACTTGAAAAAGCAAAACAAGCCTTAGAAGACCAGCGTAAAGATTTAGA
AACTAAATTAAA
AGAACTACAACAAGACTATGACTTAGCAAAGGAATCAACAAGTTGGGATAGACAAAGACTTGAAAAAGAGTTA
GAAGAGAAAAA
GGAAGCTCTTGAATTAGCGATAGACCAGGCAAGTCGGGACTACCATAGAGCTACCGCTTTAGAAAAAGAGTTA
GAAGAGAAAAA
GAAAGCTCTTGAATTAGCGATAGACCAAGCGAGTCAGGACTATAATAGAGCTAACGTCTTAGAAAAAGAGTTA
GAAACGATTACT
AGAGAACAAGAGATTAATCGTAATCTTTTAGGCAATGCAAAACTTGAACTTGATCAACTTTCATCTGAAAAAG
AGCAGCTAACGA
TCGAAAAAGCAAACTTGAGGAAGAAAAACAAATCTCAGACGCAAGTCGTCAAAGCCTTCGTCGTGACTTGGA
CGCATCACGTG
AAGCTAAGAAACAGGTTGAAAAAGATTTAGCAAACTTGACTGCTGAACTTGATAAGGTTAAAGAAGACAAACA
AATCTCAGACGC
AAGCCGTCAAGGCCTTCGCCGTGACTTGACGCATCACGTGAAGCTAAGAAACAGGTTGAAAAAGATTTAGCA
AACTTGACTGC
TGAAGTTGATAAGGTTAAAGAAGAAAAACAAATCTCAGACGCAAGCCGTCAAGGCCTTCGCCGTGACTTGAC
GCATCACGTGA
AGCTAAGAAACAAGTTGAAAAAGCTTTAGAAGAAGCAAAACAGCAAATTAGCTGCTCTTGAAAACTTAACAAA
GAGCTTGAAGAA
AGCAAGAAATTAACAGAAAAAGAAAAAGCTGAACTACAAGCAAAACTTGAAAGCAGAAGCAAAAGCACTCAAAG
AACAATTAGCG
AAACAAGCTGAAGAACTTGCAAACTAAGAGCTGGAAAAGCATCAGACTCACAAACCCCTGATACAAAACCAG
GAAACAAAGCT
GTTCCAGGTAAAGGTCAAGCACCACAAGCAGGTACAAAACCTAACCAAAACAAAGCACCAATGAAGGAACTA
AGAGACAGTTA
CCATCAACAGGTGAAACAGCTAACCCATTCTTACAGCGGCAGCCCTTACTGTTATGGCAACAGCTGGAGTAG
CAGCAGTTGTA
AAACGCAAAGAAGAAAACTAA

SPy2025
Seq ID 88
ATGAAGAAAAGGAAATTGTTAGCAGTAACACTATTAAGTACCATACTCTTAAACAGTGCAGTGCCATTAGTTG
TTGCTGATACCT
CCTTGC GTAATAGCACATCATCCACTGATCAGCCTACTACAGCAGATACTGATACGGATGACGAGAGTGAAAC
ACCAAAAAAAG
ACAAAAAAGCAAGGAAACAGCGTCGCAGCACGACCCAAAAAGACCATAAGCCATCACACACTCACCCAAC
CCCCCCTTCA
AATGATACTAAGCAGACCGATCAGGCATCATCTGAAGCTACTGACAAACCAAATAAAGACAAAAACGACACCA
AGCAACCAGAC
AGCAGTGATCAATCCACCCCATCTCCCAAAGACCAGTCGTCTCAAAAAGAGTCACAAAACAAAGACGGCCGAC
CTACCCCATCA
CCTGATCAGCAAAAAGATCAGACACCTGATAAAACACCAGAAAAATCAGCTGATAAAACCCCTGAAAAAGGAC
CAGAAAAAGCA
ACTGATAAAACACCAGAGCCAAATCGTGACGCTCCAAAACCCATCCAACCTCCTTTAGCAGCTGCTCCTGTCT
TTATACCTTGGA
GAGAAAGTGACAAAGACCTGAGCAAGCTAAAACCAAGCAGTCGCTCATCAGCGGCTTACGTGAGACACTGGAC
AGGTGACTCT
GCCTACACTCACAACCTGTTGTACGCCGTTATGGGATTACTGCTGAACAGCTAGATGGTTTTTTGAACAGTC
TAGGTATTCACT
ATGATAAAGAACGCTTAAACGGAAAGCGTTTATTAGAATGGGAAAACTAACAGGACTAGACGTTTCGAGCTAT
CGTAGCTATTG
CAATGGCAGAAAGCTCACTAGGTACTCAGGGAGTTGCTAAAGAAAAAGGAGCCAATATGTTTGTTATGGCGC
CTTTGACTTCA
ACCCAAACAATGCCAAAAAATACAGCGATGAGGTTGCTATTCGTCCATGGTAGAAGACACCATCATTGCCAA
CAAAAACCAAA
CCTTTGAAAGACAAGACCTCAAAGCAAAAAAATGGTCACTAGGCCAGTTGGATACCTTGATTGATGGTGGGGT
TTACTTTACAG
ATACAAGTGGCAGTGGGCAAAGACGAGCAGATATCATGACCAAACTAGACCAATGGATAGATGATCATGGAAG
CACACCTGAG
ATTCCAGAACATCTCAAGATAACTTCCGGGACACAATTTAGCGAAGTGCCCGTAGGTTATAAAAGAAGTCAGC
CACAAAACGTTT
TGACCTACAAGTCAGAGACCTACAGCTTTGGCCAATGCACTTGGTACGCCTATAATCGTGTCAAAGAGCTAGG
TTATCAAGTCG
ACAGGTACATGGGTAACGGTGGCGACTGGCAGCGCAAGCCAGGTTTTGTGACCACCCATAAACCTAAAGTGGG
CTATGTCGTC
TCATTTGCACCAGGCCAAGCAGGAGCAGATGCAACCTATGGTCACGTTGCTGTTGTAGAGCAAATCAAAGAAG
ATGGTTCTATC
TTAATTTTCAGAGTCAAATGTTATGGGACTAGGCACCATTTCTATCGGACGTTTACAGCTGAGCAGGCTAGTT

TGTTGACCTATG

TCGTAGGGGACAACTCCCAAGACCATAA

SPy2039

Seq ID 89

ATGAATAAAAAGAAATTAGGTGTCAGATTATTAAGTCTTTTAGCATTAGGTGGATTGTCTTGCTAACCAG
TATTTGCCGATCA

AAACTTTGCTCGTAACGAAAAAGAAGCAAAAGATAGCGCTATCACATTTATCCAAAAATCAGCAGCTATCAAA
GCAGGTGCACGA

AGCGCAGAAGATATTAAGCTTGACAAAGTAACTTAGGTGGAGAACTTTCTGGCTCTAATATGTATGTTTACA
ATATTTCTACTGG

AGGATTTGTTATCGTTTCAGGAGATAAACGTTCTCCAGAAATCTAGGATACTCTACCAGCGGATCATTGAC
GCTAACGGTAAA

GAAAACATTGCTTCCTTCATGGAAAGTTATGTCGAACAAATCAAAGAAAACAAAAAATTAGACACTACTTATG
CTGGTACCGCTG

AGATTAAACAACCAGTTGTTAAATCTCTCCTTGATTCAAAAGGCATTCATTACAATCAAGGTAACCCTTACAA
CCTATTGACACCT

GTTATTGAAAAAGTAAAACCAGGTGAACAATCTTTGTAGGTCAACATGCAGCTACAGGATGTGTTGCTACTG
CAACTGCTCAAA

TTATGAAATATCATAATTACCCTAACAAAGGGTTGAAAGACTACACTTACACACTAAGCTCAAATAACCCATA
TTTCAACCATCCT

AAGAACTTGTTTGCAGCTATCTCTACTAGACAATACAACTGGAACAACATCTTACCTACTTATAGCGGAAGAG
AATCTAACGTTC

AAAAAATGGCGATTTTCAGAATTGATGGCTGATGTTGGTATTTTCAGTAGACATGGATTATGGTCCATCTAGTGG
TTCTGCAGGTAG

CTCTCGTGTTCAAAGAGCCTTGAAAGAAAACTTTGGCTACAACCAATCTGTTACCAAATCAACCGTGGCGAC
TTTAGCAAACAA

GATTGGGAAGCACAAATTGACAAAGAATTATCTCAAAACCAACCAGTATACTACCAAGGTGTCGGTAAAGTAG
GCGGACATGCC

TTTGTTATCGATGGTGCTGACGGACGTAACCTTCTACCATGTAACTGGGGTTGGGGTGGAGTCTCTGACGGCT
TCTTCCGTCTT

GACGCACTAAACCCTTCAGCTCTTGGTACTGGTGGCGGCGCAGGCGGCTTCAACGGTTACCAAAGTGCTGTTG
TAGGCATCAA

ACCTTAG

SPy2043

Seq ID 90

ATGAATCTACTTGGATCAAGACGGGTTTTTTCTAAAAAATGTCGGCTAGTAAATTTTCAATGGTAGCTCTTG
TATCAGCCACAAT

GGCTGTAACAACAGTCACACTTGAAAACTGCACTGGCACGACAAACACAGGTCTCAAATGATGTTGTTCTA
AATGATGGCGC
AAGCAAGTACCTAAACGAAGCATTAGCTTGGACATTCAATGACAGTCCCACTATTACAAAACCTTAGGTACT
AGTCAGATCACT
CCAGCACTCTTTCCTAAAGCAGGAGATATTCTCTATAGCAAATTAGATGAGTTAGGAAGGACGCGTACTGCTA
GAGGTACATTG
ACTTATGCCAATGTTGAAGGTAGCTACGGTGTAGACAATCTTTCGGTAAAAATCAAAACCCCGCAGGCTGGA
CTGGAAACCTT
AATCATGTCAAATATAAAATTGAATGGTTAAATGGTCTATCTTATGTGCGAGATTTCTGGAATAGAAGTCATC
TCATTGCAGATAG
TCTCGGTGGAGATGCACTCAGAGTCAATGCCGTTACAGGGACACGTACCCAAAATGTAGGAGGTCGTGACCAA
AAAGGCGGCA
TGCGCTATACCGAACAAGAGCTCAAGAATGGTTAGAAGCAAATCGTGATGGCTATCTTTATTATGAAGCTGC
TCCAATCTATAA
CGCAGACGAGTTGATTCCAAGAGCTGTCGTGGTATCAATGCAATCTTCTGATAATACCATCAACGAGAAAGTA
TTAGTTTACAAC
ACAGCTAATGGCTACACCATTAACCTACCATAAC GGTACACCTACTCAGAAATAA

SPy2059

Seq ID 91

ATGAGATTTCTAGAACTTTTACAAAAGAAA TTTTTCCTAAAGCATATCAGGAAAAACAATTCTTAATGCATC
AAAAAACGCGTTTA
ACGCCACAACACAATCAAAAGCAGTATTCGCCAAATGCCAATCATTTGGACTCATCAGCTACCAAAAACCTCAG
AACAAGACCCT
GCAACAGCTCTGCAACGCAGTAGAGCCTACGAAGGAAGCCCTAAAAGTCGGCCCGCTTGGTTGCAAAAGCTGG
AAGCTGTTTT
GCCGTCTCCTCAACGTCCAATTCGGCGTTTTTGGCGCCGCTATCACATCGGAAAACCTGCTAATGATTCTGATT
GGAACCTCTTGT
CTTACTCTTAGGATCATACTTGTTTTACTTATCAAAAACAGCTAAAGTATCTGATTTACAAGATGCCTTGAAG
GCTACAACGGTTA
TTTATGATCACAAAGGAGAGTATGCAGGCAGTTTATCTGGTCAAAAAGGGAGTTATGTTGAGCTCAACGCTAT
TTCAGATGATCT
TGAGAATGCTGTTATTGCCACTGAGGATAGGACTTTTTACAGTAATAGCGGTATTAATCTTAAACGCTTCTTA
TTGGCGGTAGTT
ACGGCGGGCCGCTTTGGAGGTGGCTCAACGATTACACAGCAACTGGCTAAAAATGCTTATCTCTCACAAGATC
AGACAATTA
CGAAAGGCCCGAGAGTTTTTTTTTGGCGTTAGAGTTGACCAAAAAATACAGTAAAAAAGATATTCTTACTATGT
ACCTTAACAACCTC
CTACTTTGGAATGGAGTTTGGGGAGTTGAAGATGCCAGTCAAAAATATTTTGGAAACCACAGCTGCTAACTTA

ACACTGGATGAA

GCTGCCACATTAGCAGGTATGCTCAAAGGACCTGAAATATATAACCCTTACCATTCTCTAAAAAATGCTACTC
ACCGTAGAGATA

CTGTTTTAGGAGCGATGGTTGATGCCAAAAAGATTACCCAAACAAAAGCTCAGCAAGCTAGAGCAGTAGGGCT
AAAAAATCGCT

TAGCTGATACTTATGTTGGTAAGACAGATGACTACAAATACCCATCCTACTTTGATGCTGTTATTAGTGAAGC
AATAGCAACTTAT

GGTCTTTCAGAAAAAGACATTGTTAATAATGGATACAAAGTTTACACTGAGCTAGATCAAAATTACCAAAGTG
GCATGCAGACGA

CTTTTAACAACGATGAACTATTTCTGTTTCAGCTTATGACGGTAGCTCTGCTCAAGCAGCTAGTGTTGCTTT
AGATCCTAAAACA

GGAGGTGTTAGAGGTCTGATTGGTCGTGTGAATAGTAGTGAAAAATCCGACTTTCAGAAGTTTTAACTATGCGA
CTCAAGCAAAA

CGTAGTCCCGCATCAACAATCAAACCACTCGTGGTTTACGCGCCAGCCGTTGCTTCAGGATGGTCAATTGAAA
AAGAACTACCA

AATACCGTTCAAGATTTTCGATGGCTATCAGCCACATAATTATGGAAATTATGAATCAGAAGATGTTCCCTATGT
ATCAAGCATTAGC

AAACTCTTATAATATTCCAGCAGTTTCTACATTGAACGATATCGGAATCGATAAAGCCTTTACCTATGGTAAA
ACATTTGGGTTAG

ATATGAGCTCTGCCAAAAAAGAGTTGGGGGTAGCTTTAGGTGGCAGCGTGACAACCAATCCATTGGAGATGGC
TCAGGCATAT

GCTGCCTTTGCCAATAATGGAGTAATCCATCCTGCGCACTTGATTAACCGGATTGAAAATGCCAGGGGTGAAG
TGCTTAAAACC

TTTACTGATAAGGCTAAACGTGTTGTCAGCCAGTCTGTTGCAGATAAGATGACAGCCATGATGCTAGGTACCT
TTTCAAATGGAA

CAGCAGTCAATGCTAACGTATATGGCTATACACTAGCTGGTAAAAACAGGGACGACAGAAACCAACTTCAATCC
CGACTTAGCAG

GCGATCAGTGGGTTATTGGTTATACGCCAGATGTTGTTATTAGTCAATGGGTAGGATTTAATCAGACCGATGA
AAATCATTATCT

AACGGATTCAAGTGCAGGCACGGCCTCAGCTATTTTTAGCACTCAGGCATCTTACATTTTGCCTTATACCAAG
GGCAGCCAATTT

CATGTAGATAATGCCTACGCTCAAAATGGTATTTTTCAGCTGTTTATGGAGTCAATGAAACAGGTAATCAATCAG
GAGTTGATACTC

AATCTATTATTGATGGTTTAAAGAAAAATCAGCACAAAGAAGCTTCGCAATCACTATCAAAAGCAGTCGATCAGTC
AGGGTTACGTGA

TAAAGCCCAATCTATTTGGAAAGAGATTGTTGACTATTTTAGATAG

SPy2110

Seq ID 92

ATGGTAAGTTT TAGAAGAAGACAAGGTGACTGTTCAACCTGATATTAAAGTGATTAAACGAGATGGTCGCCTTG
TTAATTTTGATA
GTACAAAAATCTATAGTGCTTTATTA AAAAGCAAGCATGAAAGTAACTCGGATGTCGCCACTTGTTGAGGCTAA
ATTAGAGGCTAT
TTCTGATCGCATTATAGCAGAAATTATTGAGCGTTTTCCAACTAATATCAAAATTTATGAAATCCAAAATATT
GTAGAGCATAAGCT
TCTTG CAGCTAATGAATATGCTATTGCAAAAAG AATACATTAATTATCGTACTCAGCGTGACTTTGCACGTTCA
CAAGCAACAGATA
TCAATTTTTCTATTGATAAAATTAATTAATAAAGATCAAACAGTTGTTAATGAAAATGCTAACAAAGATAGCGA
TGTTTTTAATACTC
AACGAGATTTAACTGCTGGAATCGTAGGGAAATCGATTGGTTTTAAAAATGTTACCTTCGCATGTTGCTAATGC
TCATCAAAAAGG
AGATATCCATTACCATGATTTGGATTACAGTCCTTATACACCGATGACGAACTGCTGTTTAATTGACTTTAAG
GGCATGTTAGCC
AATGGCTTTAAAAATTGGTAATGCTGAAGTGGAAGTCCCAAGTCTATTCAAAC TGCAACAGCTCAGATCTCAC
AGATTATTGCGA
ATGTAGCATCAAGTCAGTACGGCGGATGCACAGCTGATCGCATTGACGAGTTTTTAGCCCCATATGCGGAGCT
TAACTTTAAAA
AACATATGGCTGATGCTAAGAAATGGATCGTTGAGACTAAGAGAGAAAGCTATGCTTTTGAAAAGACTCAAAA
AGATATTTATGA
TGCGATGCAGTCTTTGGAGTATGAAATTAATACGCTCTTTACGTCTAATGGTCAAACACCATTTACTTCTTTA
GGATTTGGTTTGG
GGACGTCTTGGTTTGAACGTGAGATTCAAAAAGCTATTTTGACCATTCCGATTAATGGTCTTGGTAGTGAACA
TCGCACGGCTAT
TTTCCCTAAATTAATTTTACGGTTAAACGTGGCTTGAATTTAGAACCAGATTCACCAA CTATGATATTAAG
ACTTTGGCTTTAG
AATGTGCGACTAAGCGGATGTACCCGGATATGTTATCTTATGATAAAATTATTGATTTGACAGGATCTTTCAA
ATCTCCAATGGG
ATGCCGCTCTTTCCTTCAAGGCTGGAAAGATGAAAATGGGCAAGATGTGACCTCAGGCCGTATGAATCTTGGG
GTTGTCACCT
CAATTTACCTCGCATTGCCATGGAATCAAATGGCGATATGGATAAGTTTTGGGAGCTGTTTAATGAGAGGATG
CTAATTAGTAAG
GATGCTTTAATTTATCGTGTCGAACGTGTCACAGAAGCAAAACCAGCAAATGCTCCTATTCTTTATCAATATG
GTGCTTTTGAAA
GCGTTTGGAGAAGACAGGGAATGTAAATGATCTCTTTAAGAATCGTCGTGCAACAGTCTCTCTTGGCTATATT
GGTCTTTATGAA
GTGGCGTCTGTTTTTTATGGTGGTCAATGGGAAGGTAATCCAGATGCTAAAGCTTTTACCTTGTC AATTGTCA
AGGCAATGAAAC
AGGCCTGTGAGGATTGGTCAGATGAATATGGTTATCATTTCTCTGTTTATTCGACTCCATCAGAAAGTTTGAC

AGATCGCTTTTG

TCGTTTAGATACGGAAAAATTTGGCATTGTGACAGATATTACGGATAAAGAATACTATACAAATTCTTTTCAC
TATGATGTGCGTA

AGAGTCCGACACCTTTTGAAAAATTAGATTTTGAAAAAGATTATCCAGAAGCAGGTGCTTCAGGTGGTTTTAT
CCACTACTGTGA

GTATCCTGTTTTGCAACAAAATCCAAAGGCCTTGGAAGCGGTTTGGGACTATGCTTATGATCGTGTGGGGTAT
TTGGGAACCAA

TACGCCTATTGATAAATGCTATAATTGCCAATTTGAAGGCGATTTACCCCAACAGAACGTGGTTTTACTTGC
CCAACTGTGGC

AATAATGACCCTAAACAGTTGATGTGGTCAAACGTACATGTGGTTACTTGGGGAATCCTCAGGCCCGCCCAA
TGGTTAACGGT

CGCCATAAGGAAATCTCTGCGCGTGTA AACATATGAATGGTTCTACTATAAAATACCCAGGCCTGTAA

SPy2127

Seq ID 93

ATGAGGAGGAATTACAGTAGAGTCATTGACGAACTGCGTACTGACTACGGGCTGAATTTAGTTGCTATTGGTC
AACGTTTGGGT

ACCGACCCCGAACAGTTGGTAAATGGTGGCAGGGTAAACATAACCCGAACCAAGAAAGCAGAAAGAACTGA
ATAGGCTATA

TAGAGAGGTGAAAGAAACTATGATGACACAAGTAAATATTTTTGAAGAAGCTAATGACAACACAAAGCAGGTT
ATGCAAGTTATT

ACAACGACAAATTTTCATGGACAACCTTTAGACATTTACGGTGATATTCAGGAGCCTTTATTTTTGGCTAGGG
CAGTCGCTGAAA

TGATTGATTACACAAAACTAGCCAAGGGTACTATGACGTACAAGCTATGCTAAGAAAAGTAGATGAGGATGA
AAAGCTTAAAG

GAATGGCTTTAGAAGGTACTACGAAAAATTTTCGTAGTGGTCAAAAAGTTTGGTTTTTAAGTGAAGTGGACT
TTATGAAGTGCT

TATGCGTTCAAACAAACCAAAAGCCAAAGAGTTTAGAAAAGCAGTCAAAAACATTCTAAAAGAAATCCGCTTG
AATGGCTATTAC

ATGCAAGGCGAATTGGTGCAAGAACTAGCTCAACCAAGCACCCAAAACTACCAGGTATAAGTGACCTAACTT
ATATACTAAATA

AGCTAGCTGATTTAGTTGATATGGATAATCTAGCTGATATTTCAAATGGGATTGACCGAGTTCAGCAACTAGT
GAAGCTGATCAG

CTTGTAG

SPy2191

Seq ID 94

ATGTTTAAGAAAGAAAAATTTAAACAACGTTATTTAATTTTGGATTAGTAGCGTTAGCTCTAACAATATTAG
CCATCATTTTTGCC

TTCTCAAGTAAAAATGCTGATACTAAGTCTTATGCTAAGAAGTCAGAAAGTAAAAATGGTAACAATCGACAAGG
CTCCAAAAAATA

ATCATGCTATTACTAAAGAAGAAAGCAAAGAAAAAGCAAAGAGCATTGCTTCGGAGCCTATTTCCACAGTAGA
AAACTCTGTAGC

TCCGACAGTAACAGAGGAAGTACCGGTTGTTTCAGCAAGAAGTGACTCAAACCTGTTTCAGCAGGTATCTTCAGTA
GCCTATAATCC

AAACAATGTGGTACTTTCCAATGGAAATACTGCTGGTATTGTAGGAAGTCAAGCGGCGGCACAGATGGCAGCA
GCAACAGGTG

TTCCACAATCAACTTGGAACATATAATTGCGCGTGAATCTAATGGAAATCCTAACGCAGCTAATGCTTCTGG
GGCATCAGGGT

TGTTCCAGACAATGCCAGGTTGGGGTTCTACAGCAACGGTTGAAGATCAAGTCAATGCAGCCTTGAAAGCCTA
TAGTGCACAAG

GTTTATCAGCTTGGGGTTACTAA

SPy2211

Seq ID 95

ATGAAAAATAATAATAAATGGATAATTGCTGGACTTGCTAGTTTTTTGTTCCTCTTAGTATTATATTTATCA
TCCTTCTATCGATG

GGCATTATTATAATAGTGATAAAACAATTCTAGCTAGTGCTTTTCATCAGTATGTTATTTTTGCGCAGA
ACTTTCGTAACATC

ATGCACGGTTCTGATAGTTTTTTTTATACCTTTACAAGCGGACTAGGGATAAATTTTTATGCTTTAATGTGTT
ATTATCTTGGCAGT

TTCTTTTCTCCATTACTTTTCTTTTTTAATTTAACCTCTATGCCAGATGCTATCTATTTGTTTACCTTGATAA
AATTTGGGTAAATAG

GATTAGCTGCATGCTATTCTTTTCATAGATTATATCCAAAAATCAGTGCTTTCTTGATGATTTCCATCTCAGT
TTTTTATAGCTTAA

TGAGCTTCTTGACAAGTCAAATGGAACATAAATTC TTGGTTAGATGTTTTCACTCTTCCACTTGTTATAC
TTGGATTAAATAAAC

TTATCACAGAAAAATAAAACCAGAACTTATTATCTTTCGATATCATTATTATTCATTCAAAATTACTACTTTGG
CTACATGATTGCTCT

TTTTTGATTCTTTACGCCTTAGTTTGCTTTTACGTCTCAATGATTTTAACAAAATGTTTATCGCTTTTGTT
AGGTTTACAGCTGT

GTCAATATGTGCTGCTTTAACAAGTGCTCTAGTAATACTTCCTACCTATCTAGATTTGTCAACTTATGGAGAG
AATCTATCCCCGA

TAAACAGTTAGTTACGAACAATGCTTGGTTTTTGGATATACCTGCTAAGCTCTCAATAGGAGTGACGATAC
TACCAAGTTTAAT

GCTCTGCCTATGATTTACGTAGGATTATTTCCCCTAATGCTTAGTGTTATTTATTTTACTTTAGAAAGTATCC
CTTTAAAAATAAAA

TTAGCCAATGCCTGCTTGTTAACTTTTATTATAATAAGTTTTTACCTACAGCCACTTGATCTTTTTTGGCAGG

GGATGCACTCACC

AAATATGTTTTTGCATCGCTACGCTTGGTCTTTTTCCATAGTTATCCTATTACTCGCATGTGAGACTCTCTCT
CGACTAAAAGAAG

TGACTCAAATAAAAGCAGGTTTTGCTTTTATTTTCCTCATTATACTGACATCTCTTCCTTATAGCTTTTCTCA
ACAATATAATTTTCT

ACCTTTAACTCTTTTTTTACTTAGTGTTTTTTTATTATTAGGTTATACTATTTCACTATTTTCGTTTAGAAAT
TCTCAAATCCCATCTA

CTTTTATTTCTGCTTTCATACTTATCTTTAGCCTTCTTGAATCAGGGTTAAACACCTACTACCAGCTTCAAGG
AATTAATAAGGAG

TGGGGATTCCCATCACGACAGATATATAATAGTCAATTAAAGGATATTAACAACCTTGTCAACTCTGTGTCAA
AAAATAGTCAACC

TTTTTTTAGAATGGAAAGGCTACTTCCCCAAACAGGGAACGATAGCATGAAATTTAATTATTACGGCATTTC
CAATTTTCCTCTG

TAAGAAATAGACTATCTAGTTCCTTATTGGATCGATTGGGATTTCACTCTAAAGGCACAAATTTAAACCTTAG
ATACCAAAACAAT

ACTATTATTATGGACAGTCTACTTGGTATAAAAATATAATCTTAGCGAAGGACCTCCAAATAAATTTGGATTTA
CAAAACTAAAACT

AGCGGGAATACTACTCTTTATCAAAATCACTATAGTAGCCCTTTAGCTATATTAACACGTAATGTTTACAAAG
ATGTCAACCTAAA

TGTCAATACCCTTGATAACCAAAACCAAAATTACTTAACCAACTAAGTGGGAAATCTTTAACCTATTTTAACTTA
CAGCCAGCTCAAC

TTATTTCTGGTGCTAATCAATTTAACGGACAAAATATCTGCACAAGCTTCTGATTATCAAACTCCGTTACCCT
TAATTATCAATTA

ACATCCCTAAACATAGTCAACTCTATGTTAGCATACCCAATATTATATTTTCAAATCCTGATGCTAAAGAGAT
GCGTATTCAGACA

GATAATCATAATTTCATATATACTACAGATAACGCTTACTCTTTTTTTGATTTAGGATATTTGCGCGATGCCA
AAGTTGCTACATTT

TCGTTTGTTTTTCCAAAAAATAAACAAATTAGTTTTAAGGAACCTCATTTTTATAGTTTGTCTATTGAATCTT
ACCTTGAAGCAATG

AATAGCATTAAACAAAAAAATGTTCACTTACGCTAAAAGTAATACGGTAATCACTGATTATAATTCAAAAA
CGAAAGGTTCTCTT

ATTTTTACACTTCCTTACGATAAAGGTTGGTCAGCACAAAAAGATGGGAAAAATCTTCCAGTCAAAAAAGCAC
AAGGAGGATTTT

TATCAGTTACTATTTCCTAAAGGAAAGGGACGTGTTATCCTTACCTTTATTCCTAATGGTTTTAAATTAGGGTT
ATCTCTATCTTGTG

TAGGAATTATCGCTTATATGCTTTTGTATAAGTACATAGATATAAAGTCTAAATTACTTTAG

ARF0450

Seq ID 96

ttttccagattttctgcccacacggagggaactattcctcgctatggtcagcctcttgtcgaaatgaacactaca
actctcaacaccatcacggagttgggacggtttcgtctaagcagaatccac
gaccactt

ARF0569

Seq ID 97

tcgtttattttgggaaaaaagaaaccccgaaggtagc

ARF0694

Seq ID 98

aaggtgaggaaaagacagaagttacaaaagagaagcttttgggaattggctagatggattaaagacatctcaga
cgataccgacgaaaagacagaagatgaggcgctactgtgacgga
gatggaacagaagaacaacccgttg

ARF0700

Seq ID 99

ctttaccagaagaagcgattaaaaaaatcacagaggctatcaaagttgatgacatcccgtctattaaacaaag
ctttgatgacctcaaaaatagag

ARF1007

Seq ID 100

tttgttctacaaaaatattccttatggcaa

ARF1145

Seq ID 101

ccaataagtatgcagaaggctatccaggtaatcgctattatgggtggaacagaatgtgtggatattgtcgaaac
gcttgctattgaacgtgctaagaaacttttcggtgccgcatTTGCCAATgTT
caagcccattccggaagtcaagccaatgcagcagcttatatggcct

ARF1208

Seq ID 102

tttcggaacatcttttTGTGactTTtagTTgctgttcctcttttGTatcctgttatcaaaaacttaaaagaaaag
gttacaatcggacaagcaaaaagaggtttttg

ARF1262

Seq ID 103

tgggttactgggtgtaaatggTggTTTTTTTTgc

ARF1294

Seq ID 104

cta atgctgaaagccaagaaaaccaattcaaaactcgtaacattgtcccagccaacaaagaagttcaatcttc
agaagctgttcaatcaaacgaacttgctaaaaccgttatcactatgggt
tcttcttcagactacactgtcg

ARF1316

Seq ID 105

ccccgaaatgggttggggacgttttgaacgttatgaaagacttggcagaacaaggcatgaccatgttaattgtt
actcacgaaatgggatttgcctcgccaagtagc

ARF1352

Seq ID 106

ttgatgaactgcccatcattgcacttcttgcaaccaagcacaaggaacaacctgtattaaagatgctcaaga
attacgagtc aaagaaacagatcgtattcaag

ARF1481

Seq ID 107

aaaacaacgaagtgtactacctcaaacgtcccaagctggtcgaatccagactacaaagaaccagatttagaa
gaatttgctctaggaagcatggacggtatagaagatggatcagga
gatttttaattttttaacaaacaaatcaaaaaagatttttggtaaaacggcgagtaaagagacttttgctaag
tttgctagttagtgcgccgaaggaatcgaaaaaatggagttaagccaatttt
taattgga

ARF1557

Seq ID 108

ggcagaagactaccaccaagattacctcaagaaaaatccaaatggctactgccatat

ARF1629

Seq ID 109

ttttggctgccaggaagcagatattttgttagggacgcaaatgattgccaaaggactggattttccaaatgtg
acttttagttgggtactaaatgcagatacttccttaaatgtgccgattttcgag
cgtcagaaaaaacgtttcaattgttaacacaggttgctggctcgggcaggccgtgccca

ARF1654

Seq ID 110

tgcgtgccatcggtgaaatgctctattatgatccagatcaacacacctttatcaattttatttccaaatacat
tggtccaggctggggttatttttcgggtctatcctattggatttcccttatttttatttga
atggcagaaatcacagcag

ARF2027

Seq ID 111

tctgattatTTTtagacatcatgctcccttccttaagtggcttagaagtgctaaaaacaattcgaaagacatca
gatgtccctattattatgctaacggcattaga

ARF2093

Seq ID 112

ttactgctgctgaagcagacatcaaagctgaacttgctgctgaaggcaaaccagaaaagatctgggacaaaat
catcccaggtaaaatggaccgcttcatgcttgacaacactaaagttg
accaagcttacactcttcttgcacaagtttacatcatggacgacagcaaaactgttgaagcttaccttgactc
ag

ARF2207

Seq ID 113

cacctattcgtgaaagacgTTTggagtacgctaaagatatgggagaggtgttccgtatgttacaagaaggtag
tcaaaaagcaagaactgtggcagccaagactttatcagaag

CRF0038

Seq ID 114

ctgtacgctccacaatctctgtcaaatecTTTTcttgatagtattccatgtgaccag

CRF0122

Seq ID 115

aacagacgtgacttccttgacaatgtggtaaatecttatcagggtcacggagccacttggcaccagcctcgTTT
ttgacaactTTTattaacagggcagcccatattaatatcaacaatatcggcct
tggtattggTTTgaatgaaatcagctgcacgTTTcaacccttctgcgtcaccacccaaa

CRF0406

Seq ID 116

caagcaccacttgacgataccataacaagccaacatattgggtcagggtatctt

CRF0416

Seq ID 117

tatTTTctcaagaaaacacTTTgaaggctgcaaaatecttggttggtgtcaccatttggagaaatggctaaga
cagggtTTTccttgggcattTTTTtctaagataaatttggcatctgcattTTTaaaggt
gccaaagtgtgtgtagaccattgtcagaaatcgTTTTggttgacacattggtctctgaccaagagtgacaagtc
ctgtaaagtgtttaccgaca

CRF0507

Seq ID 118

agcaaaaacaaaaggaacaccgataccggttgtaataacggtaaaggcagcaaaagcgtttcccataatcata
gtaaaaataaccatgccaaagcacataagecgagaacgccaaag
aagcggctatcagcaggtacaatctcccgaatgagagaagcaatcacatcaccaacaccagcagctgcaaaaa
tagctcc

CRF0549

Seq ID 119

ctttttatacatcgatcgcggttgatattggattttctcgtaatcaatttttcttggtttgtccaaatttacg
atgattttttgaatggt

CRF0569

Seq ID 120

tcgtttatattgggaaaaaagaaaccccgaaggtagc

CRF0628

Seq ID 121

atcaaacatttgaccaagcaaaacaacggatgccagtatctaaggttctagtagctaacccttaggggtcta
aaggaatagctgattcaatccaattgaggatgaaacctttagctgttaag
agataccggttcattctctttcaacaagg

CRF0727

Seq ID 122

ccgcctccaccaaatttaccacctgcgtgtaagactgtaaaaactgtttcaacggcgggacgtcctgtcttgg
cttgatatcaactggaattccacggccatcatctactactgtaatggaatt
atctgcttcaataaagactttaata

CRF0742

Seq ID 123

gaaaacgagttcaatcagtactatcaagacgcaaaatacaagtcccataaagaacgtcttactatcaatactt
tcaagagacaagggttatattgagg

CRF0784

Seq ID 124

acaagggtagccccatattttccacaagccttggaatatcttgagcctgttacattaatagcctcattca
aggcttttcttctgtcttctgctgcatcaactgccgta

CRF0854

Seq ID 125

tcaccttttacaaaatcagttttaacgccttctgcactcaaactttcttcgataaaacgtcctgtaaaaccac
ccaagaaccctgtcgccgtgcctagcaatgtctaaacgttgcaagatgcgact
cacattaatgcctttacctccagcaaacttatcatcactcgccatgcgattcactgaccecaagatttatttgg
tcaattctgacaataaagtcataagaagggtttaaggtcacggta

CRF0875

Seq ID 126

gatcacaaatatttattggcattatcaactaaaattgtcacaaagtccaacaatgacttttcctccgctt

CRF0907

Seq ID 127

ccaaatctgctggaccattttctgccaataacccgcatcaaaaccataaagcaaagctggat

CRF0979

Seq ID 128

caatgtctcataattatcaataatatcaattttcgccataataaaaacactccttcgtgtatcttgaaaagag
cgtctctacatgatataatatttcatgcagaaacacttctgtgg

CRF1068

Seq ID 129

gaaaaactcttcagaacggcgcggaacgggtacaacttcaagtgggtttccaaaagcaaatacatgggtctgg
tgcttatcgtttttttgaagtccecgcaacaggaatagtgagtcgaagt
tttttctatcacatctttggccacattgacttgtccaacgcctccatcaacaataatcaagttaggagcttgc
agtcctctctttttcacccgactgtaccttcgaaaaagcacttctctcactactggc
ataatcatctgggtccaccactgttttaactctaaatttacgataatctttcttactaggtttgccatcaaca
aaaacaaccatagccgcaactggacttgtcccttgaatgttagaattatcaaagg
cttcaatgcgcacaggcttgtcaatcctaag

CRF1152

Seq ID 130

gataactgtcgcgatgcctctcgacgaatctcgataacaaaacgcactccttcaggctagattcgtcacga
actgctgtaataccttc

CRF1203

Seq ID 131

ataggttcataaaccaactagtgagcttcgcttttagtaacaatggatgaggtaaaaacattattttaaacc
taataactcctacttttgaggatgc

CRF1225

Seq ID 132

taccaagtgtgtcagctgcctgaggggtgttccccaactgatcgtaaacgcagaccaaacctgggtttataaag
caaaaaccaagcaaagaacgaaaaagcaatggcgaagtaaccaa

CRF1236

Seq ID 133

atacttgctaaaatagaaattagccaaaaacactcataactataataataatgcgctca

CRF1362

Seq ID 134

tttccagatgtcaatgataaggtcattgcgcgtgatagtcgggtttttcggtaaagagagaacctgccttttga
gcaatcgcttctgggaagtcgggtgacaaaaagttgagaccaaccccaata
atgacatcgggtgat

CRF1524

Seq ID 135

aaattcaatgttacgcttttgccactctgtcaaacgagcttgtcatcactttgtttctctttatcttttgcc
atttacttctccttaatcgcgagctaataatcagctctcaaaaggtcataaaagaagtc
cgg

CRF1525

Seq ID 136

gagacctctgctttgtagcagctcgcttcatacgctctgcggattctccaaaataatcatctgcttaagtc
ctaaaaggctctggcaccaaattcgtaaaattcattgccacgatctagaccgcc
agcaatcaaaatcaagcgactgttatcaaaacctgataaagctttt

CRF1527

Seq ID 137

aagagactggatgattgcaccctgatattcttcttggtgtatcaatttgaacacgctcaaacgggttcacatttg
acaccatcaatttctttgatgataacttctggacg

CRF1588

Seq ID 138

gaatggctaagtcagcaacactctcatcaatttttacatgataagcgaagctatctggataccgcatttgata
aaggtaaattgtactttttcacaaaaagccagttcttcttttaagggtggatcatctta
tcttggc

CRF1649

Seq ID 139

catgaccatttgtcacatcaacaatcggtgaagcaacttggaatcttggtctggatagtaaataatcacc
agaatgataaatattataaagagtctgctgcgcacgcggga

CRF1749

Seq ID 140

gtcatgtttaaagcaccgggtcaaattgatgtttcaagacgcgctcaaaatcctcttccgtcattttaagcatca
atttatcattagtaatgccagcattgttaactaaaacatcaatcgaacctaaac
tttcgatagcttcattgaccatacgttttagcttctgatgcttctgaaacatccccagagatggtaacaactgt
aacgccg

CRF1903

Seq ID 141

agggaattatccagatttagtagctctgcaatcctgctaaaaccactagttaaggtgcctcctaataaggacaa
tggcacctcctaatacctccagctgctccacttcttggtattgtctgtaaatca
atggttgtttggcaaaaaacttttttagcaaagttacttgcaatt

CRF1964

Seq ID 142

gagcactttcaaccaaatcaccaaattgggcaaaaatggaaaaagagcgaccaaaccaacatggtcaaaaa
gtgatgtggcgcataaacagacctatcaatcacca

CRF2055

Seq ID 143

aaatcaaggtctttaaggacttcttcgatcacctcgtccacggttttttctctccaactatttttagtattaa
gcccatcacttggttaacatctgctcaggcgctttttgttgggattcgggattggcaca
ccatggacatctcaaggggcaaccttttaaaaaaacggttgccgaataccaggaccgtca

CRF2091

Seq ID 144

cgaaatgaaatcaacacttctctctacgttttcttctaccaaata

CRF2096

Seq ID 145

tgtctcttagcagcttttaccaaattttcaaaatcagccattgttccaaaatcagggttgaccgccgta

CRF2104

Seq ID 146

gaggccatgtcttgccagcgttgctgttcttttttagctgactcttctttgatacgatatgatttggccaagt
cctccgtgggtcatagctgccaatgatttttaaaatggcttggctgtcttgtggaaggag

gtgtggg

CRF2116

Seq ID 147

aatttcaaagctcgaccacctggtgttggttcaggatttccaaacatcacaccaactttttcacgcaattggt
taataaagatagcgatagtttttgttttattaatagaagctgataatttacgcatg
gcc

CRF2153

Seq ID 148

aaagagacctgccactacagccaccaacatcatgtcattttccaagccaagttcgataggaactttttgccac
atccagataaaaagcggacatcagaactgttccatagatggt

NRF0001

Seq ID 149

attacctgttgaccaagcaaacgcagcaactgttcaggaagcccagtcctttcaacaatctgttgaagcatc
tcttggtaaagagaatgtcattgtcaatgttcttgaaacag
aaacatcaactcacgaagcccaaggcttctatgctgagaccccagaacaacaagactacgatatcatttcac
atggtggggaccagactatcaagatccacggacctac
cttgacatca

NRF0003

Seq ID 150

tcgggaagacaggattcgaacctgcgacaccttggtcccaaaccaagtactctaccaagc

SPy0012

Seq ID 151

MRKLLAAMLMTFFLTPLPVISTEKKLIFSNAVYQLKQDVVQSTQFYNQIPSNPNLYQETCAYKDSDLTLPAG
RLGVNQPLLIKSLVLNK
ESLPVFELADGTYVEANRQLIYDDIVLNQVDIDSYFWTQKKLRLYSAPYVLGTQTIPSSFLFAQKVHATQMAQ
TNHGTYYLIDDKGWA
SQEDLVQFDNRMLKVQEMLLQKYNPNYSIFVKQLNTQTSAGINADKKMYAASISKLAPLYIVQKQLQKKKLA
ENKTLTYTKDVNHFY
GDYDPLGSGKISKIADNKDYRVEDLLKAVAQQSDNVATNILGYLCHQYDKAFRSEIKALSGIDWDMEQRLLT
SRSAANMMEAIIYHQK
GQIIISYLSNTEFDQQRITKNITVPVAHKIGDAYDYKHDVAIVYGNTPFILSIFTNKSTYEDITAIADDVYGIL
K

SPy0019

Seq ID 152

MKKRILSAVLVSGVTLGAATTGVAEDLSTKIAKQDSIIISNLTTEQKAAQNQVSALQAQVSSLQSEQDKLTARN
TELEALSKRFEQEIKAL

TSQIVARNEKLKNQARSAYKNNETSGYINALLNSKSISDVVNRLVA INRAVSANAKLLEQQKADKVSLEEKQA
ANQTAINTIAANMAMA

EENQNTLRTQQANLVAATANLALQLASATEDKANLVAQKEAAEKAEEALAEQAAKVKAQEQAQAASVEA
AKSAITPAPQATPA

AQSSNAIEPAALTAPAAPSAGPQTSYDSSNTYPVGQCTWGAKSLAPWAGNNWGNNGGQWAYSQAAGYRTGSTP
MVGAIIVWND

GGYGHVAVVVEVQSASSIRVMESNYSGRQYIADHRGWFNPTGVTFIYPH

SPy0025

Seq ID 153

MSSYFPVAPLSDLVSYMNKRI FVEKKADFGIKSASLVKELTHNLQLTSLKALRIVQVYDVFNLAEDLLARAEK
HIFSEQVTDCLLTETEIT

AELDKVAFFAIEALPGQFDQRAASSQEALLLFGSDSQVKVNTAQLYLVNKDITEAELEAVKNYLLNPVDSRFK
DITLPLEEQAFSVSDK

TIPNLDFFETYQADDFATYKAEQGLAMEVDDLLFIQNYFKSIGCVPTETELKVLDITYWSDHCRHTTFETELKN
IDFSASKFQKQLQTTY

DKYIAMRDELGRSEKPQTLMDMATIFGRYERANGRLDDMEVSDEINACSVIEVDVDGVKEPWLLMFKNETHN
HPTEIEPFGGAATCI

GGAIRDPLSGRSYVYQAMRISGAGDITTPIAETRAGKLPQQVISKTAAHGYSSYGNQIGLATTYVREYFHPGF
VAKRMELGAVVGAAP

KENVVREKPEAGDVVILLGGKTGRDGVGGATGSSKVQTVESVETAGAEVQKGNAIEERKIQR LFRDGNVTRLI
KKSNDFGAGGVCVA

IGELADGLEIDLKVPLKYQGLNGTEIAISESQERMSVVVRPNVDFAIAACNKENIDAVVVATVTEKPNLVM
TWNGEIIVDLERRFLDT

NGVRVVVDAKVVDKDLTVPEARTTSAETLEADTLKVLSDLNHASQKGLQTFDSSVGRSTVNHPIGGRYQITP
TESSVQKLVPVQHGVT

TTASVMAQGYNPYIAEWSPYHGAAYAVIEATARLVATGADWSRARFSYQEYFERMDKQAERFGQPVSALLGSI
EAQIQGLPSIGGK

DSMSGTFEDLTVPPTLVAFGVTTADSRKVLSPFEKAAAGENIYYIPGQAISEDIDFDLIKDNFSQFEAIQAQHK
ITAASAAKYGGVLESLAL

MTFGNRIGASVEIAELDSSLTAQLGGFVFTSAEEIADAVKIGQTQADFVTVNGNDLAGASLLAAFEGKLEEV
YPTEFEQTDVLEEVPA

VVSDTVIKAKETIEKPVVYIPVFPGTNSEYDSAKAFEQVGASVNLVPFVTLNEVAIAESVDTMVANIAKANI I
FFAGGFSAADEPDGSAKF

IVNILLNEKVRAAIDSFIEKGGLIIGICNGFQALVKSGLLPYGNFEEAGETSPTLIFYNDANQHVAKMVETRIA
NTNSPWLAGVEVGDIHAI

PVSHGEGKLVVSASEFAELRDNGQIWSQYVDFDGQPSMDSKYNPNNGSVNAIEGITSKNGQIIIGKMGHSERWED
GLFQNIPGNKDQIL
FASAVKYFTGK

SPy0031

Seq ID 154

MKKFHRFLVSGVILLGFNGLVPTMPSTLISQQENLVHAAVLGDNYPSKWKKGNGIDSWNMYIRQCTSFAAFRL
SSANGFQLPKGYGN
ACTWGHIAKNQGYPVNKTPSIGAIAWFDKNAYQSNAAYGHVAWVADIRGDTVTEEYNNAGQGPERYHKRQI
PKSQVSGYIHFKDL
SSQTSHSYPRQLKHISQASFDPSGTYHFTTRLPVKGQTSIDSPDLAYYEAGQSVYYDKVVTAGGYTWLSYLSF
SGNRRYIPIKEPAQS
VVQNDNTKPSIKVGDVTFPGVFRVDQLVNNLIVNKELAGGDPTPLNWIDPTPLDETDNQKVLGDQILRVGE
YFIVTGSYKVLKIDQP
SNGIYVQIGSRGTWVNADKANKL

SPy0103

Seq ID 155

MINQWNNLRHKKLKGFTLLEMLLVILVISVLMLLFVPNLSKQKDRVTTETGNAAVVKLVENQAELYELSQGSKP
SLSQLKADGSITEKQE
KAYQDYDKHKNEKARLSN

SPy0112

Seq ID 156

MKIGIIGVGKMASAIIKGLKQTPHELIISGSSLERSKEIAEQALALPYAMSHQDLIDQVDLVILGIKPQLFETV
LKPLHFKQPIISMAAGISLQR
LATFVGQDLPLLRIMPNNMAQILQSSTALTGNALVSQELQARVRDLTDSFGSTFDISEKDFDTFTALAGSSPA
YIYLFIEALAKAGVKNG
IPKAKALEIVTQTVLASASNLKTSSQSPHDFIDAICSPGGTTIAGLMELERLGLTATVSSAIDKTIDKAKSL

SPy0115

Seq ID 157

MTDLFSKIKEVTELDGIAGYEHSVRDYLRTKITPLVDRVETDGLGGIFGIRDSKAEEKAPRILVAAHMDEVGFM
VSDIKVDGTLRVVGIGG
WNPLVVSSQRFTLYTRTGQVIPLISGSVPPHFLRGANGSASLPHIEDIVFDGGFTDKAEAERFGITPGDIIIP
QSETILTANQKNIIISKAWD
NRYGVLMIITEMLEALKGQDLNNTLIAGANVQEEVGLRGAHVSTTKFDPELFFAVDCSPAGDIYGNPGTIGDGT
LLRFYDPGHVMLKD
MRDFLLTTAAEEAGVNFQYYCGKGGTDAGAAHLQNGGVPSTTIGVCARYIHSHTLYAMDDFVEAQAFLQAIK

KLDRSTVDLIKCY

SPy0166

Seq ID 158

MEDISDPEVILEYGVYPAFIKGYTQLKANIEEALLEMSNSGQALDIYQAVQTLNAENMLLNYYESLPFYLNRRQ
SILANMTKALKDAHIRE
AMAHYKLGEFAHYQDTMLDMVERTIKTF

SPy0167

Seq ID 159

MSNKKTFFKYSRVAGLLTAALIIGNLVTANAESNKQNTASTETTTTNEQPKPESELTTTEKAGQKTDDMLNSN
DMIKLAPKEMPLESA
EKEEKKSEDKKKSEEDHTEEINDKIYSLNYNELEVLAKNGETIENFVPKEGVKKADKFIVIERKKKNINTTPV
DISIIDSVTDRTPAALQL
ANKGFTENKPDVVTKRNPQKIHIIDLPGMGDKATVEVNDPTYANVSTAIDNLVNQWHDNYSGGNTLPARTQYT
ESMVYSKSQIEAAL
NVNSKILDGTLGIDFKSISKGEKKVMIAAYKQIFYTVSANLPNNPADVFDKSVTFKELQRKGSNEAPPLFVS
NVAYGRTVFVKLETSSK
SNDVEAAFSAAKGTDVKTNGKYSDILENSSFTAVVLGGDAAEHNKVVTKDFDVIRNVIKDNATFSRKNPAYP
ISYTSVFLKNNKIAGV
NNRTEYVETTSTEYTSKGKINLSHQGAYVAQYEILWDEINYDDKGKEVITKRRWDNNWYSKTSFPSTVIPLGAN
SRNIRIMARECTGLA
WEWWRKVIDERDVKLSKEINVNISGSTLSPYGSITYK

SPy0168

Seq ID 160

MKQQSYQPLRFVYLLVALFAALLLIARPVMADEGTNSADAAYYKGQSAGKKAGKKAGKEATWTDLTPTVPTNP
ETPSDIGETTNKQL
YKEGYKDGYPEGYNEGWKSQYPVLTVPKVIWDLISYWLQRLFPNNQSSTAAQSMS

SPy0171

Seq ID 161

VKNKLFLVALATVTVLGPLATPHHQTVHASDVTLTETCDKNGTVCFGYENVIDGEVCKLTADGKGITICVGYEN
RDIKESETSSTKNDC
SNWFWCFLNYLWTTIKSWVS

SPy0183

Seq ID 162

METILEVKHLSKIFGKKQKAALEMVKTGKNKSEIFKKTGATVGVDASFEVKKGEIFVIMGLSGSGKSTLVRM

LNRLIEPSAGSILLEGK

DISTMSADQLREVRHDINMVFSFALFPHKTILENTEFGLLELRGVPKEERQRLAEKALDNSGLLDFKDQYPN
QLSGGMQQRVGLAR

ALANSPKILLMDEAFSALDPLIRREMQDELLDLQDSMKQTIIFISHDLNEALRIGDRIALMKDGQIMQIGTGE
EILTNPANDFVREFVEDV

DRSKVLTAQNIMIKPLTTTVELDGPQVALNRMHNEEVSMLEATNRRRQLVGSLTADAAIEARKKGLPLSEVID
RDVRTVSKDTIITDILP

LIYDSSAPIAVTDDNNRLLGVIIRGRVIEALANISDEDLN

SPy0230

Seq ID 163

MKTARFFWFYFKRYRFSFTVIAVAVILATYLQVKAPVFLGESLTELKGIGQAYYVAKMSGQTHFSPDLSAFNA
VMFKLLMTYFFTVLAN

LIYSFLLTRVSSHSTNMRKGLFGKLERLTVAFFDRHKDGEILSRFTSDLDNIQNSLNQSLIQVVTNIALYIG
LVWMMFRQDSRLALLTIA

STPVALIFLVINIRLARKYTNIQQQEVSALEAFMDETISGQKAIIVQGVQEDTMTAFLKHNERVRQATFKRRL
FSGQLFPVMNGMSLINT

AIVIFVGSTIVLSDKSMPLAAAAALGLVVTFFVQYSQQYYQPMQIASSWGELQLAFTGAHRIQEMFDETEEVRPQ
NAPFTSLKEAVAIN

HVDFGYLPGQKVLSDVSIVAPKGKMIAVVGPTGSGKTTIMNLINRFYDVDAGSITFDGRDIRDYDLDSLQKQV
GIVLQESVLFSGTITDN

IRFGDQTISQDMVETAARATHIHDFIMSLPKGYNTYVSDDDNVSTGQKQLISIARTLLTDPEVLILDEATSN
VDTVTESKIQRAMEAIVA

GRTSFVIAHRLKTIILNADHIIVLKDGKVIIEQGNHHELLHQKGFYAELYHNQFVFE

SPy0269

Seq ID 164

MDLEQTKPNQVKQKIALTSTIALLSASVGVSHQVKADDRASGETKASNTHDDSLPKPETIQEAKATIDAVEKT
LSQQKAELTELALT

KTAEINHLKEQQDNEQKALTSAQEIYTNLTASSEETLLAQGAHQRELATETELHNAQADQHSKETALSEQ
KASISAETTRAQDLVE

QVKTSEQNIAKLNAMISNPDAITKAAQTANDNTKALSSELEKAKADLENQKAKVKKQLTEELAAQKAALAEKE
AELSRLKSSAPSTQDS

IVGNNTMKAPQGYPLEELKKLEASGYIGSASYNYYKEHADQIIAKASPGNQLNQYQDIPADRNRFVDPDNL
PEVQNELAQFAAHMI

NSVRRQLGLPPVTVTAGSQEFARLLSTSYKKTHGNTGPSFVYGQPGVSGHYGVGPHDKTIIEDSAGASGLIRN
DDNMYENIGAFNDV

HTVNGIKRGIYDSIKYMLFTDHLHGNTYGHAINFLRVDKHNPNAPVYLGFTSNVGSLEHFMFPESENIANH
QRFNKTPIKAVGSTKD

YAQRVGTVSDTIAATKGVSSLENRLSAIHQEADIMAAQAKVSQ LQGKLASTLKQSDSLNLQVRQLNDTKGSL
RTELLAAKAKQAQLE

ATRDQSLAKLASLKAALHQTEALAEQAAARVTALVAKKAHLQYLRDFKLNPNRLQVIRERIDNTKQDLAKTTS
SLLNAQEALALQAKQ

SSLEATIATTEHQLTLLKTLANEKEYRHLDEDIATVPDLQVAPPLTGVKPLSYSKIDTTPLVQEMVKETKQLL
EASARLAAE NTSLVAEA

LVGQTSEMVASNAIVSKITSSITQPSSKTSYSGSGSSTSNLISDVDESTQRALKAGVVMLAAVGLTGFRFRKE
SK

SPy0287

Seq ID 165

MTKEKLVAFSQAHAEPAWLQERRLAALAEIPNLELPTIERVKFHRWNLGDGTLTENESLASVPDFIAIGDNPK
LVQVGTQTVLEQLPM

ALIDKGVVFSDFYTALEEIPEVIEAHFGQALAFDEDKLAAYHTAYFNSSAAVLYVPDHLEITTPIEAIFLQDSD
SDVPFNKHVLVIAGKESKF

TYLERFESIGNATQKISANISVEVIAQAGSQIKFSAIDRLGPSVTTYISRRGRLEKDANIDWALAVMNEGNVI
ADFDSDLIGQGSQADLKV

VAASSGRVQGIDTRVTNYGQRTVGHILQHGVILERGTLTFNGIGHILKDAKGADAQQESRVLMMLSDQARADA
NPILLIDENEVTAGH

AASIGQVDPEDMYLMSRGLDQETAERLVIRGFLGAVIAEIPISVRQEIIKVLDEKLLNR

SPy0292

Seq ID 166

MIKRLISLVVIALFFAASTVSGEEYSVTAKHAIADVLESGKVLYEKDAKEVVPVASVSKLLTTYLVYKEVSKG
KLNWDSPTISNYPYELT

TNYTISNVPLDKRKYTVKELLSALVVNNANSPAIALAEKIGGTEPKFVDKMKKQLRQWGISDAKVVNSTGLTN
HFLGANTYPNTEPDD

ENCFCATDLAIIARHLLLEFPEVLKLSSKSSTIFAGQTIYSYNYMLKGMPCYREGVDGLFVGYSKKAGASFVA
TSVENQMRVITVVLNA

DQSHEDDLAIFKTTNQLLQYLLINFQKVQLIENNKPVKTLYVLDSPKTVKLVAQNSLFFIKPIHTKTKNTVH
ITKKSSTMIAPLSKGQVLG

RATLQDKHLIGQGYLDTPPSINLILQKNISKSFFLKVWWNRVRYVNTSL

SPy0295

Seq ID 167

MESIDKSKFRFVERDSEASEVIDTPAYSYWKS VFRQFFSKKSTVFMLVILVTVLMMSFYIPMFANYDFNDVSN
INDFSKRYIWPNAEY

WFGTDKNGQSLFDGVWYGARNSILISVIATLINITIGVVLGAIWGVSKAFDKVMIEIYNIISNIPSMIIIVL
TYSLGAGFWNLILAFCTGWIG

VAYSIRVQILRYRDLEYNLASQTLGTPMYKIAVKNLLPQLVSVIMTMLSQMLPVYVSSEAFLSFFGIGLPTTT
PSLGRFIANYSSNLTTNA
YLFWIPLVTLILVSLPLYIVGQNLADASDPRSHR

SPy0348

Seq ID 168

LALTDKDKDQQDQQRSFKEQILAELEKANQIRKEKEEELFQKELEAKEAARRTAQLYAEYKRQDAFQKESIA
HNNKTAKHFQAIKGA
VMTSEALKPTLLSEKENSSLKTTNKRVVQANELQETASKESQVPLTIEKGHSVRRKLSKRQQTERAAKKISTV
LISSIIITLLAVTLAGAG
YVYSALNPVDKNSDAFVQVEIPSGSGNKLIGQILQKKGLIKNSTVFSFYTKFKNFTNFQSGYYNLQKSMSLEE
IASALQEGGTAEPTKP
SLGKILPEGYTIKQIAKAVEHNSKGKTKKAKTPFNEKDFDLVTDEAFIQDMVKRYPKLLATIPTKEKAIYR
LEGYLFPATYNYKETTM
RELVEDMLAAMDATLVPYYDKIAASGKTVNEVLTLASLVEKEGSTDDRRQIASVFYNRLNSGMALQSNIAIL
YAMGKLGEKTTLAEDA
TIDTTINSPYNIYTNTGLMPGPVASSGVSAIEATLNPASTDYLYFVANVHTGEVYYAKTFEEHSANVEKYVNS
QIQ

SPy0416

Seq ID 169

VEKKQRFSLRKYKSGTFSVLIGSVFLVMTTVAADELSTMSEPTITNHAQQQAQHLTNTELSSAESKSQDTSQ
ITLKTNREKEQSQDL
VSEPTTTELADTDAASMANTGSDATQKSASLPPVNTDVHDWVKTKGAWDKGYKGQKVVAVIDTGIDPAHQSM
RISDVSTAKVKSK
EDMLARQKAAGINYGSWINDKVVFHANYVENSJNIKENQFEDFDEDWENFEFDAEAEPKAIKKHKIYRPQSTQ
APKETVIKTEETDGS
HDIDWTQTDDDTKYESHGMHVTGIVAGNSKEAAATGERFLGIAPEAQVMFMRVFANDIMGSAESLFIKAIEDA
VALGADVINLSLGT
NGAQLSGSKPLMEAIEKAKKAGVSVVVAAGNERVYGSDDHDDPLATNPDYGLVGSPSTGRTPTSVAAINSKWVI
QRLMTVKELENRAD
LNHGKAIYSESVDKDIKDSLGYDKSHQFAYVKESTDAGYNAQDVKGKIALIERDPNKTYDEMIALAKKHGAL
GVLIFNNKPGQSNRS
MRLTANGMGIPSAFISHEFGKAMSQNLNGNGTGSLEFDSVVS KAPSQKGNEMNHFSNWGLTSDGYLKPDI TAPG
GDIYSTYNDNHYG
SQTGTSMAS PQIAGASLLVKQYLEKTQPNLPKEKIADIVKNLLMSNAQIHVNPETKTTTSPRQQGAGLLNIDG
AVTSGLYVTGKDNYG
SISLGNITDTMTFDVTVHNL SNKDKTLRYDTELLTDHVPDQKGRFTLTSHSLKTYQGGEVTPANGKVTVRVT
MDVSQFTKELTKQMP

NGYYLEGFVRFRDSQDDQLNRVNIPFVGFKGQFENLAVAEESIYRLKSQGKTGFYFDESGPKDDIYVGKHFTG
LVTLGSETNVSTKTI

SDNGLHTLGTfKNADGKFILEKNAQGNPVLAI SPNGDNNQDFAAFKGVFLRKYQGLKASVYHASDKEHKNPLW
VSPEsfKGDKNFN

SDIRFAKSTTLGTAFSGKSLTGAELPDGHYHYVVSYPDVVGAKRQEMTFDMILDRQKPVLSQATFDPETNR
FKPEPLKDRGLAGV

RKDSVFYLERKDNKPYTVTINDSYKYVSVEDNKTfVERQADGSFILPLDKAKLGDFYfMVEDFAGNVAIAKL
GDLHPQTLGKTPIKLL

TDGNYQTKETLKDNLMTQSDTGLVTNQAQLAVVHRNQPSQLTKMNQDFFISPNEGDKDFVAFKGLKNNVY
NDLTVNVYAKDDH

QKQTPIWSSQAGASVSAIESTAwyGITARGSKVMPGDYQYVVTYRDEHGKEHQKYTISVNDKKPMITQGRFD
TINGVDHFTPDKTK

ALDSSGIVREEVFYLAKKNGRKFDVTEGKDGITVSDNKVYIPKNPDGSYTISKRDGVTLSDYfYLVEDRAGNV
SFATLRDLKAVGKDK

AVVNFGLDLPVPEDKQIVNFTYLVRDADGKPIENLEYYNNSGNSLILPYGKYTVELLTYDTNAAKLESdKIVS
FTLSADNNFQQVTFKIT

MLATSQITAHFDHLLPEGSrvSLKTAQDQLIPLEQSLYVPKAYGKTvQEGTYEVVSLPKGYRIEGNTKVNTL
P NEVHELsRLVKVGD

ASDSTGDHKVMSKNNsqALTASATPTKSTTSATAKALPSTGEKMGLKLRIvGLVLLGLTCVfSRKKSTKD

SPy0430

Seq ID 170

MKWSGFMKTKSKRFLNLATLCLALLGTTLLMAHPVQAEVISKRDYMTRFGLGDLEDDsANYPSNLEARYKGYL
EGYEKGLKGDDIPE

RPKIQVPEDVQPSDHGDYRDGYEEGFEGQHkRDPLETEAEDDSQGGRQEGRQGHQEGADSSDLNVEESDGLS
VIDEVVGVIYQA

FSTIWTYLSGLF

SPy0433

Seq ID 171

MKKTLTLLLALFAIGVTSSVRAEDEQSSTQKPVKFDLDGPQQKIKDYSGNTTTLEDLYVGSKVVKIYIPQGWW
VYLYRQCDHNSKERGI

LASPILEKNITKTDpyRQYYTGVPYILNLGEDPLKKGEKLTFsfKGEDGFYVGSYIYRSDTIKKEKEAEEAL
QKKEEEKQKQLEESM

LKQIREEDHKPWHQRLSESIQDQWWNFKGLFQ

SPy0437

Seq ID 172

MKKTLTLLLALFAIGVTSSVRAEDEQNKfILDGLQEKVKEVSVDfSVGESKIKVWLPQAWSVKISREHSPKS

SISNSGEQKPLSNSSE

NKEGQFSKRLPYGTQHTIKLSSQLTKGERVTLTFRDEDFWGAGYCFYRDSLSIKEDKQYEEEIKKIEDDLERQ
DLENDALFMFKKQTE
REANKPWHQRLSENIQDQWWNFKGLFQ

SPy0469

Seq ID 173

MIITKKSLFVTSVALSLVPLATAQAQEWTPRSVTEIKSELVLVDNVFTYTVKYGDTLSTIAEAMGIDVHVLGD
INHIANIDLIFPDTILTANY
NQHGQATNLTVQAPASSPASVSHVPSSEPLPQASATSQPTVPMAPPATPSDVPTTPFASAKPDSSVTASSELT
SSTNDVSTELSSSES
QKQPEVPQEAVPTPKAAETTEVEPKTDISEAPTSANRPVPNESASEEVSSAAPAQAPAEKEETSAPAAQKAVA
DTTSVATSNGLSYA
PNHAYNPMNAGLQPQTAAFKEEVASAFGITSFSGYRPGDPGDHGKGLAIDFMVPENSALGDQVAQYAI DHMAE
RGISYVIWKQRFY
APFASIYGPAYTWNMPDRGSITENHYDHHVHSFNA

SPy0488

Seq ID 174

LRQIQSIRLIDVLELAFGVGYKEETTSQFSSDQPSQVVLRYGEANTVRFAYTNQMSLMKDIRIALDGSDKSLT
AQIVPGMGHVEYEGFQT
SARGIFTMSGVPESTVPVANPNVQTKYIRYFKVIDDMHNTMYKGTVFLVQPQAWKYTMKSVDQLPVDDLNIHIG
VAGIERMTTLIKNAG
ALLTTGGSGAFPDNIKVSINPKGRQATITYGDGSTDIPPAVLWKKGSVKEPTQADQSVGTPTPGIPGKFKRD
QSLNEHEAMVNVEPLS
HVVKDNIKVIDEKSTGRFEPFRPNEDEKEKPASDVKVRPAEVGSWLEPATALPSVEMSAEDRLKS

SPy0515

Seq ID 175

MKVLLYLEAENYLKSGIGRAIKHQAKALSLVGQHFTTNPRETYDLVHLNTYGLKSWLLMIKAQKAGKKVIMH
GHSTEEDFRNSFIFSN
LLSPWFKKYLCHFYNKADAIITPTLYSKSLIESYGVKSPIFAVSNGIDLEQYGADPKKEAAFRRYFDIKEGEK
VVMGAGLFFLRKGIDDF
VKVAQAMPDVRFIWFGETNKWVIPAQVRQMVNGNHPKNLIFPGYIKGDVYEGAMTGADAFFFPSREETEGIVV
LEALASRQHLVLRDI
PVYYGWVDQSSAELATDIPGFIEALKKVFSGASNKVEAGYKVAQSRRLETVGHALVDVYKKVMEL

SPy0580

Seq ID 176

MENNNHNIAEALSVSLHQIEQVLALTAQGNTIPFIARYRKEVTGNLDEVVIXSIIDMDKSLTTLNERKATIL
AKIEEQGKLTDQLRTSIEA

TEKLADLEELYLPYKEKRRTKATIAREAGLFPLARLILQNAQNLETAAPFVTEGFASPQEALAGAVDILVEA
MSEDAKLRSWTYNEIW

QYSRLVSTLKDEQLDEKKVFQIYYDFSDQVSNMQGYRTLALNRGEKLGILKVSFEHNLEKMQRFFSVRFKETN
PYIEEVINQTIKKKIV

PAMERRVRSELSDAEDGAIHLFSENLRHLLVSPKKGKMLGFDPAFRTGAKLAIVDQTGKLLTTQVIYPVA
PASQTKIQAAKETLTQ

LIETYQIDI I AIGNGTASRESEAFVADVLKDFPNTSYVIVNESGASVYSASELARHEFPDLTVEKRSAISIAR
LRQDPLAELVKIDPKSIGV

GQYQHDVSQKKLSENLFVVDTVVNQVGVNVTASPSLLAHVSGLNKTI SENIVKYREENGALTSRADIKKVP
RLGAKAFEQAAGFLR

IPGAKNILDNTGVHPESYPAVKELFKVLGIQDLDDAAKATLAAVQVPQMAETLAIGQETLKDI IADLLKPGRD
LRDDFEAPILRQDILDLK

DLEIGQKLEGTVRNVVDFGAFVDIGVHEDGLIHISEMSKTFVNHP SQVVSVDLTVVWSKIDLRHKVNLSL
LPPRDTH

SPy0621

Seq ID 177

MNEKVFRDPVHNYIHIDNPLIYDLINTKEFQRLRRIKQVPTTAFTFHGAEHSRFSHCLGVYEIARRVTAIFEE
KYADIWNKDESLVTMTA

ALLHDIGHGAYSHTFEVL FHTDHEAFTQE I ITNPETEINAILVRHAPDFPDKVASVINHTYPNKQVVQLISSQ
IDCDRMDYLLRDSYFSAA

NYGQFDLMRILRVIRPVEDGIVFEHSGMHAVEDYIVSRFQMYMQVYFHPASRAVELILQNLLKRAQHLYPEQQ
AYFQKTAPGLIPFFE

KKANLADYIALDDGVMNTYFQVWMASEDHILSDLASRFINRKILKSVTFDQDSQGELERLRQLVESVGFDPDY
YTGIIHINFDLPYDIYR

PELENPRTQIEMMQKDGLAELSQLSPIVKALTGTTYGDRRFYFPKEMLELDDL FAPSKETFMSYISNGHFHF
SQ

SPy0630

Seq ID 178

MDINLLQALLIGLWTAFCFSGMLLGIYTNRCIILSFGVGII LGDLPTALSMGAISELAYMGFGVGAGGTVPPN
PIGPGIFGTLMAITSAGKV

TPEAALALSTPIAVAIQFLQTFAYTAFAGAPETAKKQLQKGNIRGFKFAANGTIWAFAFIGLGLLLGALSMD
TLLHLVDYIPPVLLNGLT

VAGKMLPAIGFAMILSVMAKKELIPFVLIGYVCAAYLQIPTIGIAIIGIIFALNEFYNKPKQVDATTVQGGQQ
DDWI

SPy0681

Seq ID 179

LTPRSKTTAGHFRYARYLIESEDENHLVTAYNQEQAYRLFIDGDGTGLMHIFDGNCEIKHDERGDHLLITTP
KGNKRVIYKGGGKVN

SVGAITGMSLGSVVFCEINLLHMDFIQECFRRTWAAKLRVHLADLNPPAPQHPVIKDVFDVQNTRWTHWTMDD
NPILTAERKQNIINS

LKKNPYLYKRDVLGQRVMPQGVIYGLFDTEKNVLDALIGEPVEMYFCADGGQSDATSMSCNIVTRVRDNGRIS
FRLNRVAHYHSGA

DTGQVKAMSTYALELKVFIDWCVKKYQMRYTEVFVDPACKSLREELHKLGVFTLGAPNNSKDVSSKAKGIEVG
IERGQNIISDGAFFL

VNHSEEEYDHYHFLKEIGLYSRDDNGKPIDKDNHAMDEFKRYSVNVFVHRYYN

SPy0683

Seq ID 180

MKKKPIKLNDEQLLLEASQLSDMYHQLTDLDFDQVIERIKARGSSASLADNPYLWQANKLHDVGLLNADNIKLI
AKYSGIAEAQLRYIIKNE

GFKIYKNTSEQLLEALGRESGVNSTIQDDLSNYARQAIDDVHNLNTTLPFSVIGAYQGI IQDAVAGVVTGLK
TPDQAINQTVIKWFKKG

FYGFTDKAGRKWRADSYARTVINTTTWRVFNEAKEAPAREFGIDTFYYSKKATAREMCAPLQHQIVTTGEARE
EGGIKILALSDYGHG

EPDGCLGINCKHTKTPFVVGVSNSKPELPEHLKNITPAQAKANANAQAKQRAIERSIRKSKELLHVAKQLGDKE
LIRQQSDVRSKQDA

LNLYLINNAFLHRNQAREKRYNNPYTKTQSEVEVRKEKAKLDKRRDVESAIIGVETSEGIPLKITKHLAERAV
LRNIAPIDIVDSIKEPLKI

APIKYDNLDRPSQKYIGKCVSTVINPIDGNIVTVHATSTRIRKKYGGN

SPy0702

Seq ID 181

MSRDPTLILDESNLVIGKGRVHYTFTTEDDNPVKVRLASKCLGTAHFNQLMIERGDQATSYVAPVVVEGTGNP
TGLFKDLKEISLELTD

TANSQLWSKIKLTNRGMLQEYYDGKIKTEIVNSARGVATRISEDTDKKLALINDTIDGIRREYRDADRKLSAS
YQAGIEGLKATMANDKI

GLQAEIKASAAQLSQKYDDELKLSAKITTTSSGTTEAYESKLAGLRAEFTRSNQGTRELESQISGLRAVQQ
STASQISQEIRDREGA

VS RVQQSLESYQRRMQDAEENYSSLTHTVRGLQSDVGSPGKIQSRLTQLAGQIEQRVTRDGVMSIISGAGDS
IKLAIQKAGGINAKM

SGNEIISAINLSYGVTIAGKHIALDGNTTVNGTFTTKIAEAIKIRADQIIAGTIDAARIRVINLNASSIVGL
DANFIKAKIGYAITDLLEGKVIK

ARNGAMLIDLNTAKMDFNSDATINFNSKNALVRKDGTHTAFFVHFSNATPKGYTGSALYASIGITSSGDGVNS

ASSGRFAGLRSFRYA

 TGYNHTAAVDQTEIYGDNLVVDDFNITRGFKFRPDKMQKMLDMNDLYAAVVALGRCWGHLANVGWNTAHSNF
TSAVNRELNNYIT
 KI

SPy0710

Seq ID 182

 MTFLDKIKQGCLDGWAKYKILPSLTAAQAILESGWGKHAPHNALFGIKADSSWTGKSFDTKTQEEYQAGVVTD
IVDRFRAYDSWDESI
 ADHGQFLVDNPRYEAVIGETDYKKACYAIIKAAGYATASSYVELLIQLIEENDLQSWDREALKNNKEETMTTAN
EIVQYCVNLANS GMG
 VDKDGAHGTQCCDLPCFVAKNWFGVDLWGNAIDLLDSASAQGWVHRMPTEANPKAGATFVQSVYPYHQFGHTG
IVIEDSDGYTMR
 TVEQNIDGNPDALYVGAPARFNTRDFTGVIGWFYPPYQGDVTQPVSTEPQTSDTIVETAKTGTFTLDVAEIN
IRRWPSLASEVVGIYK
 QGDTVSFDSSEGYANGYYWISYVGGSGMRNYLGIGQTDKDGNRISLWGKLN

SPy0711

Seq ID 183

 MKKINIIKIVFIITVILISTISPIIKSDSKKDISNVKSDLLYAYTITPYDYKNCRVNFSTHTLNIDTQKYRG
KDYYISSEMSYEASQKFKRDDH
 VDVFLFYILNSHTGEYIYGGITPAQNNKVNHLKLLGNLFISGESQQNLNKKIILEKDIVTFQEIDFKIRKYL
M
DNYKIYDATSPYVSGRIEIG
 TKDGKHEQIDLFDSPNEGTRSDIFAKYKDNRIINMKNFSHFDIYLEK

SPy0720

Seq ID 184

 MITTFETILDKIKAHQTI IHRHQNPDPDALGSQAGLKEIIAQNFPDKKVLMTGFDEPSLAWISQMDQVTDKD
YKEALVIITDTANRPRIDD
 ERYTLGKCLIKIDHHPNDDVYGDFYYVDTSSASSASEIIADFAFSQNLTLSDKAAKLLYTGIVGDTGRFLYAST
TSKTLSIASQLRHFEDF
 AAISRQMSFPLKIAKLQSYVFEHLTIDESGAAYVLVSQETLKHFDVTLAESSAIVCAPGKIDNVQAWAIFVE
LTDGNYRVRMRSKEII
 NGIAKRHGGGGHPLASGANSANLEENQAIFRELIAVCQEI

SPy0727

Seq ID 185

 MIEENKHFEKKMQEYDASQIQVLEGLEAVRMRPGMYIGSTAKEGLHHLVWEIVDNSIDEALAGFASHIKVFIE
ADNSITVVDDGRGIPV

DIQAKTGRPAVETVFTVLHAGGKFGGGGYKVSGLHGVGSSVVNALSTQLDVRVYKNGQIHYYQEFKRGAVVAD
LEVIGTTDVTGTTV

HFTPDPEIFTETTQFDYSVLAKRIQELAFNLNRGLKISITDKRSGMEQEEHFLYEGGIGSYVEFLNDKKDVIFE
TPIYTDGELEGI AVEVAM

QYTTSYQETVMSFANNIHTHEGGTHEQGFRAALTRVINDYAKKNKILKENEDNLTGEDVREGLTAVISVKHPN
PQFEGQTKTKLGNSE

VVKITNRLFSEAFQRFLLENPQVARKIVEKGI LASKARIAAKRAREVTRKKSGLEISNLPGLADCSSNDANQ
NELFIVEGDSAGGSAKS

GRNREFQAILPIRGKILNVEKATMDKILANEEIRSLFTAMGTGFGADFDVSKARYQKLVIMTDADV DGAHIRT
LLLTLIYRFMRPVLEAG

YVYIAQPPIYGVKVGSEIKEYIQPGIDQEDQLKTALEKYSIGRSKPTVQRYKGLGEMDDHQLWETTMDPENRL
MARVTVDDAAEADKV

FDMLMGDRVEPRRDFIEENAVYSTLDI

SPy0737

Seq ID 186

MRKVKKVFVSSCMLLTVGLGVAVPTGFSQSNGVMVVKAAEVPATDL SRQASDSERVDESSLQKENLSVDSFK
LENLNGWEAENDT

AGNLGKFKDPDSSGYQNILTSSGKNISVAVAPKGS GKMNIKVTKRSNFQGGYYVGGLRTQTPVLKLN DVYRYS
FTTKKLSGNSSEFK

TRVKPVESNNKLGKELVIRVDNKNVSTKHDWLPDISDGTHTVDFTGLDKKLSVAFRFS PRQTSNVVYEF SNIN
IKNISPASVPAIPSKVL

EGTSVLSGTAISSGDTLEKRKSFDGDILRVYKDSKI IARTVIKGNKWDVKLSKPLIAGEKLD FEILHPRSQNV
SKKISKQVEAKPFD PASY

KEKVIAKLKPVYEATS EKITNDAWL DENAKDLQKQKLEEQYISGKVAISEAGTKQE AIDAAYNKYSSQ TDPD
SLPSQYKQGNKENEQE

KGRQDLIQTRDLTLKAIQEDKWLTEQE KTIQKEEALKAFETGIESVNQTVSLEQLKQRLIVYKASEKDSEKKE
YPESIPNQHIPGKEKEV

KAQKQEELKKLHDTTLEKINQDKWLTPDQQA EQLKQAEVTFKKGQEA IKS AQTLTQLETDLADYVSENEGKN
SIPDKYKSGNKDDL

VNKA EVKLKEAHEATKQAI EKDPWLSPEQKKAQKEKAKARLDEGLKALKAADSLEILKVTEEAFVDKEKNPDS
IPNQHKAGTADQARK

QALDSL DKEVQKELESIDNDNTLT TDEKAAAKKKVNDAYDVAKQTAMEANSYEDLT TIKDEF LSNLPHKQGTP
LKDQQSDAIAELEKK

QQEIEKAI EGDKTLPRDEKEKQIADSKERLKS DTQKVKDAKNADA IKKAFE EGKVNIPQA HIPGDLNKDKEKL
LAELKQKADDTEKAIDV

DKTLTEDEKKEQKVKTAELEKAKTDVKNTQTREELDKKVP ELKKAIEDTHVKG NLEGVKNKAIEDLKKAHTE
TVAKINGDDTL DKATK

EAQVKEADKALAAGKDAITKADDADKVSTAVTEHTPKIKAAHKTGDLKKAQVDANTALDKAAEKERGEINKDA
TLTTEDKAKQLKEVE

TALTKAKDNVKA AKTADAINDARDKGVATIDAVHKAGQDLGARKSGQVAKLEEAAKATKDKISADPTLTSKEK
EEQSKAVDAELKKAIE

AVNAADTADKVDDALGEGVTDIKNQHKSGDSIDARREAHGKELDRVAQETKGAIEKDPTLTTEEKAKQVKDVD
AAKERGMAKLNEAK

DADALDKAYGEGVTDIKNQHKSGDPVDARRGLHNKSIDEVAQATKDAITADTTLTEAEKETQRGNVDKEATKA
KEELAKAKDADALD

KAYGDGVTSIKNQHKSGKGLDVRKDEHKKALEAVAKRVTAIEADPTLTPEVREQQKAEVQKELELATDKIAE
AKDADEADKAYGDG

VTAIENAHVIGKIEARKDLAKKDLAEAAAKTKALIIEDKTLTDDQRKEQLLGVDTEYAKGIENIDAAKDAAG
VDKAYS DGVRDILAQYKE

GQNLNDRRNAAKEFLLKEADKVTKLINDDPTLTHDQKVDQINKVEQAKLDAIKSVDDAQTADAINDALGKGIE
NINNQQYQHGDGVDVR

KATAKGDLEKEAAKVKALIAKDPTLTQADKDKQTAAVDAAKNTAIAAVDKATTTEGINQELGKGITAINKAYR
PGEGVKARKEAAKADL

EKEAAKVKALITNDPTLTKADKAKQTEAVAKALKAAIAAVDKATTAEGINQELGKGITAINKAYRPGE
GVKARKEAAKADLEREAAKVR

EAIANDPTLTKADKAKQTEAVAKALKAAIAAVDKATTAEGINQELGKGITAINKAYRPGEVGAHKEAAKANL
EKVAKETKALISGDRYL

SETEKAVQKQAVEQALAKALGQVEAAKTVEAVKLAENLGTVAIRSAYVAGLAKDTDQATAALNEAKQAAIEAL
KQAAAETLAKITTDK

LTEAQKAEQSENVSLALKTAIATVRSAQSIASVKEAKDKGITAIRAAYVPNKAVAKSSSANHLPKSGDANSIV
LVGLGVMSLLLGMVLYS

KKKESKD

SPy0747

Seq ID 187

MINKKCIIPVSLTLAITLTSVEEVTSRQNLTANEIVTQRPKRESVISDKSNFPVISPYLASVDFGERKTPL
PTPDKGVKVTTEQSIAQVR

KGPEERPYTVTGKITSVINGWGGYGFYIQDSEGIGLYVYPQKDLGYSGDIVQLTGTLTRFKGDLQLQQVTAH
KKLELSFPTSVKEAVI

SELETTTSTLVKLSHVTVGELSTDQYNNTSFLVRDDSGKSIVVHIDHRTGVKGADVVTKISQGDLINLTAIL
SIVDGGQLQLRPFSLQLE

VVKKVTSSNSDASSRNIVKIGETQGASHTSPLLKKA VTVEQVVVTYLDDSTHFYVQDLNGDGLATSDGIRVF
AKNAKVQVGDVLTISG

EVEEFFGRGYEERKQTDLTITQIVAKAVTKTGTAQVPSPLVLGKDRIAPANIIDNDGLRVFDPEEDAIDYWES
MEGMLVAVDDAKILGP

MKNKEIYVLPGSSTRPLNNSGGVLLPANSYNTDVIPVLFKKGKQIIKAGDSYKGRLAGPVSYSYGNYKVFVDD
SKNMPSLMDGHLKPE

KTNLQKDL SKLSIASYNIE NFSANPSSTKDEKVKRIAESFIHDLNAPDIIGLIEVQDNNGPTDDGTTDATQSA

QRLIDAIKKLGGPTRYV

DIAPENNVDGGQPGGNI RTGFLYQPERVSLSDKPKGGARDALTWVNGELNLSVGRIDPTNAAWKDVRKSLAAE
FIFQGRKVVVVAN

HLNSKRGDNALYGCVPVTFKSEQRRHVLANMLAQFAKEGAKHQANIVMLGDFNDFEFTKTIQLIEEGDMVNL
VSRHDISDRYSYFH

QGNNQTLDNILVSRHLLDHYEFDMVHVNSPFMEAHGRASDHDP LLLQLSFSKENDKAESSKQSVKAKKTSK GK
LLPKTGDSL VYVITL

LGTASLLVPILL LTKGKKES

SPy0777

Seq ID 188

VISFAPFLSPEAIKHLQENERCRDQSQKRTAQQIEAIY TSGQNILVSASAGSGKTFVMVERILDKILRGV SID
RLFISTFTVKAATELRERI

ENKLYSQIAQT TDFQMKVYLTEQLQSLCQADIGTMDAFAQKVVSRYGYSIGISSQFRIMQDKAEQDVLKQEVF
SKLFNEFMNQKEAPV

FRALVKNFSGNCKDTSAFREL VYTCYSFSQSTENPKIWLQENFLSAAKTYQRLEDIPDHDIELLLLAMQDTAN
QLRDVTDMEDYGQLT

KAGSRSAKYTKHLTIIEKLSDWVRDFKCLYGKAGLDRLIR DVTGLIPSGNDVTVSKVKYPVFKTLHQKLKQFR
HLETILMYQKDCFSLLE

QLQDFVLAFSEAYLAVKIQESAFEFS DIAHFAIKILEENTDIRQSYQQHYHEVMVDEYQDNNHMQERLLTLLS
NGHNRFMVGDIKQSIY

RFRQADPQIFNQKFRDYQKKPEQKGK VILLKENFRSQSEVLNVSNAVFSHLMDES VGDVLYDEQHQLIAGSHAQ
TVPYLD RRAQLLLY

NSDKDDGNAPSDSEGISFSEVTIVAKEI IKLHNDKGVPFEDITLLVSSRTRNDIISHTFNQYGIPIATDGGQQ
NYLKSVEVMVMLDTLRTI

NNPRNDYALVALLRSPMFAFDEDDLARIALQKDNELDKDCLYDKIQRAVIGRGAHPELIHDTLLGKLN VFLKT
LKS WRRYAKLGSLYDL

IWKIFNDRFYFDFVASQAKAEQAQANLYALALRANQFEKSGYKGLYRFIKMIDKVLETQNDLADVEVATPKQA
VNLMTIHKSKGLQFPY

VFILNCDKRFSMTDIHKSFILNRQH GIGIKYLADIKGLLGETTLNSVKVSMETLPYQLNKQELRLATLSEMR
LLYVAMTRA EKKVYFIGK

ASKSKSQEITDPKKLGKLLPLALREQLLTFQDWLLAIADIFSTEDLYFDVRFIEDSDLTQESVGR LQTPQLLN
PDDLKDN RQSETIARAL

DMLEAVSQLNANYEAAIHLPTVRTPSQLKATYEP LLEPIGVDIIEKSSRSLSDFTLPHFSKKAKVEASHIGSA
LHQLMQVLPLSKPINQQ

TLLDALRGIDSNEEVKTALDLKKIESFFCDTSLGQFFQTYQKHLYREAPFAILKLDPI SQEEYVLRGIIDAYF
LFDDHIVLV DYKTDKYKQP

IELKKRYQQQLELYAEALTQTYKLPVTKRYLVLMGGGKPEIVEV

SPy0789

Seq ID 189

MVKTDFKLRYQGS AIGYLSILKPLMMFTIMYLVFIRFLRLGGNVPHFPVALLLANVIWSFFSEATSMGMVSI
VSRGDLLRKLNF SKHII

VFSAVLGALINFLINLVVVLIFALINGVTISGYAYLSLFLFIELVVLVLGIALLLSNVFVYYRDLAQVWEVLL
QAGMYATPIIYPITFVLD SHPL

AAKLLMLNPVAQMIQDFRYLLIDRANVTIWQMSTNWFYIVIPYLVFPVILFIGIFVFKNADRFAEII

SPy0839

Seq ID 190

MTFLSDLISLMTKIRLSWVIKAGIFQLLFVTIANIVLSEFFYFILDVTGQYHLDKDNVVTFLKNPIALALLGA
YLFLLAAFIHLEFFALYRIIAD

QEISFYLFRRKQFSYYLRGLWKTFSGYQLLLFLLYILLTIPVLHIGLSSVITQKLYLPEFIVGELSKITSTKYL
LYGSLILVFYLNRLVYFLPLI

AINHRTVAQAWRESWQTKKKHVLLWMKLFAINGLTIVVLSLAISMILIFVDMFNPKGNNIIVQLGALTFTWE
LIFTTTIFFKLCSAMILKE

AIEPQKQYDEPRRSNKAYVVIFIVVTVGFAYSLERLTFFDTSKTVIAHRGLVSAGVENSLEALEGAKKAG
SDYVELDLILT KDNHFV

VSHDNRLKRLAGVNKTIRNLTKEVEHLSHQGHFSGRFVSFDTFYQKAKKLNMPLLIELKPIGTEPGNYVDL
FLETYHRLGISKDNKV

MSLDLEVIEAIKKKNPSITTYIIPIQFGFFGDEFVDFYVIEDFSYRSYLSSQAFWNNKEIYVWTINDPKRIE
HYLLKPIQGIITDQPALTNQ

LIKDLKQDNSYFSRLVRIISSLY

SPy0843

Seq ID 191

MKKHLKTVALTLTTSVVTNQEVSFSLVKEPILKQTQASSSISGADYAESSGKSKLKINETSGPVDDTVTDLF
SDKRTTPEKIKDNLAKG

PREQELKAVTENTESEKQITSGSQLEQSKESSLNKTVPSTSNWEICDFITKGNTLVGLSKSGVEKLSQTDHL
VLPSQAADGTQLIQVA

SFAFTPDKKTAIAEYTSRAGENGEISQLDVDGKEINEGEVFN SYLLKKVTIPTGYKHIGQDAFVDNKNIAEV
NLPESLETISDYAFAHLA

LKQIDLPDNLKAIGELAFFDNQITGKLSLPRQLMRLAERAFKSNHIKTIEFRGNSLKVIGEASFQDNDLSQLM
LPDGLEKIESEAF TGNP

GDDHYNNRVVLWTKSGKNPSGLATENTYVNPDKSLWQESPEIDYTKWLEEDFTYQNSVTGFSNKG LQKVKN
KNLEIPKQHNGVT

ITEIGDNAFRNVDFQNKTLRKYDLEEVKLPSTIRKIGAFQSNLKSFEASDDLEEIKEGAFMNNRIETLEL
KDKLVTIGDAAFHINHIYA

IVLPESVQEIGRSAFRQNGANNLIFMGSKVKTLGEMAFLSNRLEHLDLSEQKQLTEIPVQAFSDNALKEVLLP

ASLKTIREEAFKKNHLK

QLEVASALSHIAFNALDDNDGDEQFDNKVVVKTHHNSYALADGEHFIVDPDKLSSTIVDLEKILKLEGLDYS
TLRQTTQTQFRDMTTA

GKALLSKSNLRQGEKQKFLQEAQFFLGRVDLDKAIKAIEKALVTKKATKNGQLLERSINKAVLAYNNSAIKKA
NVKRLEKELDLLTGLV

EGKGPLAQATMVQGVYLLKTPLPLPEYYIGLVNYFDKSGKLIYALDMSDTIGEGQKDAYGNPILNVEDNEGY
HALAVATLADYEGLDI

KTILNSKLSQLTSIRQVPTAAYHRAGIFQAIQNAAAAEAEQLLPKPGTHSEKSSSESANSKDRGLQSNPKTNR
GRHSAILPRTGSKGSF

VYGILGYTSVALLSLITAIKKKKY

SPy0872

Seq ID 192

MKKYFILKSSVLSILTSFTLLVTDVQADQVDVQFLGVNDFHGALDNTGTAYTPSGKIPNAGTAAQLGAYMDDA
EIDFKQANQDGTIRV

QAGDMVGASPANSALLQDEPTVKVFNMKFEYGTIGNHEFDEGLDEFNRIMTGQAPDPESTINDITKQYEHEA
SHQTIVIANVIDKKT

KDIPYGWKPYAIKDIAINDKIVKIGFIGVVTTEIPNLVLKQNYEHYQFLDVAETIAKYAKELQEQHVHAIIVL
AHVPATSKDGVVDHEMAT

VMEKVNQIYPEHSIDIIFAGHNHQYTNGTIGKTRIVQALSQKGAYADVRLDTRDNDTDFIKTPSANVVAVAPG
IKTENSIDIAIINHNDIV

KTVTERKIGTATNSSTISKTENIDKESPVGNLATTAAQLTIAKKTFTVDFAMTNGGIRSDLVVKNDRTITWG
AAQAVQPFQGNILQVIQM

TGQHIYDVLNQYDENQTYFLQMSGLTYTYTDNDPKNSDTPFKIVKVYKDNNGEINLTTTYTVVVNDFLYGGG
DGFSAFKKAKLIGAIN

TDTEAFITYITNLEASGKTVNATIKGVKNYVTSNLESSTKVNSAGKHSIISKVFRNRDGNVTSSEVISDLLTS
TENTNNSLGKKETTTNKN

TISSSTLPITGDNYKMSPIMTILALISLGGLNAFIKKRKS

SPy0895

Seq ID 193

MTNNQTLDILLDVYAYNHAFRIAKALPNIPKTALYLLE MLKERRELNLAFLEHAAENRTIEDQYHCSLWLN
QSLEDEQIANIYILDLEVKV

KNGAIIDFVRSVSPILYRLFLRLITSEIPNFKAYIFDTKNDQYDTWHFQAMLESDEHVFAYLSQKQSRNVT
KSLADMLTLTSLPQEIKD

LVFLLRHFKAVERNPLAHLIKPFDEEELHRTTHFSSQAFLENIITLATFSGVIYRREPFIYFDDMNAIKKELS
LWRQSIV

SPy0972

Seq ID 194

MKTTSLIKVDLPSTIGIGYGAFWRSRNFYRVVKGSRGSKKSKTTALNFIVRLLKYPWANLLVIRRYSENTNKQS
TYTDFKWACNQLKVT

HLFKFNESLPEITVKATGQKILFRGLDDELKITSITVDVGALCWAWFEEAYQIETEDKFSTVVESIRGSLDAP
DFFKQITVTFPWSEH

VVLKRVFFDEETKRADTFSGTTTFRVNEWLDDVDKRRYEDLYKTNPRRARI VCDGEWGVAEGLVFDNFEVDF
DVEKTIQRVKETS

GMDFGFTQDPTTLCVAVDLANKELWLYNEHYQKAMLTDHIVKMIRDKNLHRSYIAGDSAERLIEIKSKGV
SGIVPSIKGKGSIMQGI

QFMQGFKIYIHPSCHEHTIEEFNTYTFKQDKEGNWLNPIDKNNHVIDAIRYALEKYHIRSNESNQFEVLRAGF
GY

SPy0981

Seq ID 195

MAEETQTVETVEEQVVPEAKQPQDEKKYTDADVDAIIDKKFAKWKSEQEAKESEAKKMAKMNEKEKADYEKQK
LLDELQELKNDKT

RNELTAVARQMFASEINVNDVLGLVVTLDAEQTKANVTTLANAFKVIADDRKALVRQTTTPSTGGGLSKQT
NYGANLASKAAQQS

TKLF

SPy1008

Seq ID 196

MRYNCRYSHIDKKIYSMIICLSFLLYSNVVQANSYNTNRHNLESYKHDSNLIEADSIKNSPDIVTSHMLKY
SVKDKNLSVFFEKDWIS

QEFKDKEVDIYALSAQEVCECPGKRYEAFGGITLTNSEKKEIKVPVNVWDKSKQQPPMFITVNPKVTAQEV
IKVRKLLIKKYDIYNN

REQKYSKGTVTLDLNSGKDIVFDLYYFGNGDFNSMLKIYSNNERIDSTQFHVDSIS

SPy1032

Seq ID 197

VNTYFCTHHKQLLLYSNLFSLFAMMGQGTAIYADTLTNSSEPNTYFQTQTLTTTDSKKVVPQQKDYYTEL
LDQWNSIIAGNDAYD

KTNPDMVTFHNKAEKDAQNIKSYQGPDHENRTYLWEHAKDYSASANITKTYRNIKIAKQITNPESCYYQDS
KAIAIVKDGMAFMYEH

AYNLDRENHQTGKENKENWWVYEIGTPRAINNTLSLMYPYFTQEEILKYTAIEKFVPDPTRFRVRAANFSP
FEANSGNLIDMGRVK

LISGILRKDDLEISDTIKAIEKVFTLVDEGNFYQDGLIDHVVTNAQSPLYKKGIAYTGAYGNVLIDGLSQL
IPIIQKTKSPIKADKMATIYH

WINHSFFPIIVRGEMMDMTRGRSISRNFNAQSHVAGIEALRAILRIADMSEEPHRLALKTRIKTLVTQGNFYN

VYDNLKTYHDIKLMKEL

LSDTSVPVQKLDSYVASFNSMDKLALYNNKHDFAFGLSMFSNRTQNYEAMNNENLHGWFTSDGMFYLYNNDLG
HYSENYWATVNP

YRLPGTTETEQQPLEGTPENIKTNYQQVGMTGLSDDAFVASKKLNNTSALAAMFTNWNKSLTLNKGWFILGN
KIIFVGSNIKNQSSH

KAYTTIEQRKENQKYPYCSYVNNQPVDLNNQLVDFNTKSFLESDDPAQNIGYYFFKPTTLSISKALQTGKW
QNIKADDSPEAIKEV

SNTFITIMQNHTQDGDYAYMMLPNMTRQEFETYISKLDIDLLENNDKLAAYVDHDSQQMHVHYGKKATMFS
NHNLSHQGFYSFPH

PVRQNNQ

SPy1054

Seq ID 198

LLTFGGASAVKAEENEKVREQEKLIIQQLSEKLVEINDLQTLNGDKESIQLVDYLTRRGKLEEEWMEYLNSGI
QRKLFVGPKGPGAGEK

GEQGPTGKQGERGETGPAGPRGDKGETGDKGAQGPVGPAGKDGQNGKDGLPGKDGDGQNGKDGLPGKDGDG
QDGDGLP

GKDGDGQNGKDGLPGKDGPQGPAPKTPEVPQNPDTAPHTPKTPRIPGQSKDVTPAPQNPSNRGLNKPQTQG
GNQLAKTPAAH

DTHRQLPATGETTNPFFTAATAVAIMTTAGVVAVAKRQENN

SPy1063

Seq ID 199

MYIFSSSKDSAKELVILTPNSQTILTGTIPAFEEKYGVKVRLLIQQGTGQLIDQLGRKDKPLNADIFFGGNYT
QFESHKDLFESYVSPQV

STVISDYQLPSHRATPYTINGSVLIVNNELARGLHITSYEDLLQPALKGKIAFADPNSSSSAFSQLTNILLAK
GGYTNADAWAYMKRLLV

NMNSIRATSSSEVYQSVAEKMIIVGLTYEDPCINLQKSGANVSIVYPKEGTVFVPSSVAIIKHAPNMTEAKLF
INFMLSRDVQNAFGQS

TSNRPIRQDAQTSHDMKALETIATLKEDYAYVTKHKKKIVATYNQLRQRLEKAK

SPy1162

Seq ID 200

MPTSIKAIKESLEAVTSLLDPLFQELATDTRSGVQKALKSRQKVIQAEIAEEERLEAMLSYEKALYKKGYKAI
AGIDEVGRGPLAGPVVA

ACVILPKYCKIKGLNDSKKIPKAKHETIYQAVKEKALAIGIGIIDNQLIDEVNIYEATKLAMLEAIKQLEGQL
TQPDYLLIDAMTLDAISQQSI

LKGDANSLSIAAASIVAKVTRDQMANYDRIFPGYDFAKNAGYGTKEHLQGLKAYGITPIHRKSFEVPVKSMMC
DSTNP

SPy1206

Seq ID 201

MTVKEETMSILEVKQLSHGFGDRAIFENVSFRLKGEHIGLVGANGEGKSTFMSIVTGHLQPDEGKVEWSKYV
TAGYLDQHTVLESG

QTVRDVLRТАFDELFKTENRINEIYASMADDDKADIAVLMEEVGELQDRLESRDFYTLDAKIDEVARALGVMDF
GMESDVTSLSGGQRT

KVLLAKLLEKPDILLDEPTNHLDAEHIEWLKRYLQHYENAFVLISHDISFLNDVINIVYHVENQSLVRYTG
DYYQFQAVYEMKQSQLE

AAYERQQKEIANLQDFVNRNKARVATRNAMMSRQKKLDKMDI IELQAEKPKPNFEFKQARTPSRFIFQTKNLV
IGYDYPLTKEPLNITF

ERNQKIAIVGANGIGKSTLLKSLGVIIEPLEGHIIVTGDFLEVGYFEQEVTVGNRQTPLEVWDAFPALNQAEV
RAALARCGLTSKHIES

QIQVLSGGEQAKVRFCLLMNRENNVLIILDEPTNHLIDAKNELKRALKAYKGSILMVCHPEPDFYNGWVTDTDW
FSKLT

SPy1228

Seq ID 202

MNKKFIGLGLASVAVLSLAACGNRGASKGGASGKTDLKVAMVTDGTGGVDDKSFNQSAWEGLQSWGKEMGLQKG
TGFDYFQSTSE

SEYATNLDTAVSGGYQLIYGIGFALKDAIAKAAGDNEGVKFVI IDDI IEGKDNVASVTFADHEAAYLAGIAAA
KTTKTKTVGFVGGMEGT

VITRFEKGFEAGVKSVDDTIQVKVDYAGSFGDAAKGKTIAAAQYAAGADV IYQAAGGTGAGVFNEAKAINEKR
SEADKVWVIGVDRD

QKDEGKYTSKDGKEANFVLASSIKEVGKAVQLINKQVADKKFPGGKTTVYGLKDGGEIATTNVSKEAVKAIK
EAKAKIKSGDIKVPEK

SPy1245

Seq ID 203

MKMKKKFFLLSLLALSTFFLSACSSWIDKGESITAVGSTALQPLVEAVADEFSSNLGKTVNVQGGGSGTGLS
QVQSGAVQIGNSDV

FAEEKDGDASKLVLDHQA VAVAGLAVIANPKVKVSNLSSQQLKIFSGEYTNWKQVGGEDLAISVINRAASSGS
RATFDSVIMKGVNAK

QSQEQDSNGMVKSIVSQTPGAISYLSFAYVDSSVKSLLNGFKANAKNVATNDWPIWSYEHMYTKDKPTGLTK
EFLDYMFSDDEVQQ

NIVTHMGYISINDMEVVKSHDGKVTKR

SPy1315

Seq ID 204

MTHKIKVLLLAIMSIFLTCNIAAETIAIVSDTAYAPFEFKDSDQIYKGIDVDI INEVAKRQSWDFSMSFPGF
DAAVNAVQSGQASALMAG

TTITNARKKVFHFSEPPYYDTKIVIATRKANAIKKYSDLGKGTGVKNGTAAQAFLLNNYKKKYDYTVKTFDTGD
LMYNSLSAGSIAAVMD

DEAVIQYAISQNQDIAINMKGEPIGSFGFAVKKGSGYDYLVDNFNTALKAMKADGTYQAIMTKWLGTDKATT
SQATGNPSAKATPTK

DSYKIVSDSSFAPFEFQNGKGKYVGIDIELIKAIKQQGFKIEIANPGFDAALNAVQSSQADGVIAGATITDA
RKAI FDFSDPYYSNIILA

VKAGKNIKNYEDLDRKTVGAKNGTSSYSWLKENAPKYGYNVKAFDDGSSMYDSLNSGSVDAIMDDEAVLKYAI
SQGRRFETPLEGIS

TGEVGFVKKGTNPELIEMFNGLAALKKSGQYDDIIDKYLDSKKAATPSEKGADESTISGLLSNNYKQLLAG
LGTTLSTLISFAIAIIIGI

IFGMMAVSPTKSLRLISTVFVDVVRGIPLMIVAIFWGVNLIESTGHQSPINDFLAATIALSLNGGAYIA
EIVRGGIEAVPAGQMEAS

RSLGLSYGTTMRKVILPQAVKLMLPNFINQFVISLKDITIVSAIGLVELFQTGKII IARNYQSFRMYAILAI I
YLIMIILLTRLAKRLEKRLN

SPy1357

Seq ID 205

MGKEIKVKCFLRRSAFGLVAVSASVLVGSTVSAVDSPIEQPRIIPNGGTLTNLLGNAPEKLALRNEERAIDEL
KKQAIEDKEATTAIEAAS

SDALEALADQTDALQSEEA VVKADNAASDALEALADQTDALQSEEA EVVQSDNAASDAWEKAATPIALDVKK
TKDTKPVVKKEERQ

NVNTLPTTGEESNPFFTAALAIMVSTGVLVVSSKCKEN

SPy1361

Seq ID 206

MKTKKVIIILVGLLLSSQLTLIACQSRGNGTYPITKQSRKGMTSNKIKPIKKSCKTNKTHKGVAGVDFPTDDG
FILTKDSKILSKTDQGIV

VDHDGHSHFIFYADLKGSPFEYLIPKGASLAKPAVAQRAASQGTSKVADPHHHYEFNPADIVAEDALGYTVRH
DDHFHYILKSSLSGQ

TQAQAKQVATRLPQTSSLVSTATANGIPGLHFPTSDGFQFNGQGIVGVTKDSILVDHDGHLHPISFADLRQGG
WAHVADQYDPAKKA

EKPAETHQTPELSEREKEYQEKLAYLAEKLGIDPSTIKRVETQDGKLGLEYPHHDHAHVLMMLSDIEIGKDIPD
PHAIEHARELEKHKVG

MDTLRALGFDEEVILDIVRTHDAPTFPSNEKDPNMMKEWLATVIKLDLGSRKDPLQRKGLSLLPNLETLGIG
FTPIKDISPVLQFKKLK

QLLMTKTGVTDYRFLDNMPQLEGIDISQNNLKDIFLSKYKNLTLVAAADNGIEDIRPLGQLPNLKFLVLSNN
KISDLSPLASLHQLQELH

IDNNQITDLSPVSHKESLTVVDLSRNADVDLATLQAPKLETLMVNDTKVSHLDLFLKNNPNLSSLSINRAQLQS
LEGIEASSVIVRVEAEG

NQIKSLVLKDKQGSLTFLDVTGNQLTSLEGVNFTALDILSVSKNQLTNVNLSPKNKTVTNIDISHNNISLAD
LKLNEQHIPEAIAKNFPA

VYEGSMVGNGTAEKAAMATKAKESAQEASESHDYNHNHTYEDEEGHAHEHRDKDDHDHEHEDENEAKDEQNH
AD

SPy1371

Seq ID 207

LAKQYKNLVNGEWKLSENEITTIYAPATGEELGSPAMTQAEVDAVYASAKKALSDWRALSYVERAAYLHKAAD
ILVRDAEKIGAILSKE

VAKGHKAAVSEVIRTAEIINYAAEEGLRMEGEVLEGGSFEEASKKKIAIVRREPVGVLVAISPFNYPVNLGS
KIAPALIAGNVVALKPPT

QGSISGLLLAEAFABEAGIPAGVFNTITGRGSVIGDYIVEHEAVSFINFTGSTPIGEGIGKLAGMRPIMLELGG
KDSAIVLEDADLALAAKNI

VAGAFGYSGQRCTAVKRVLVMDKVADQLAAEIKTLVEKLSVGMPEDDADITPLIDTSAADFVEGLIKDATDKG
ATALTAFNREGNLISP

VLFDHVTDMRLAWEEPFGPVLP IIRVTTVEEAIKISNESEYGLQASIFTTNFPKAFGIAEQLEVGTVHLNNK
TQRGTDNFPFLGAKKSG

AGVQGVKYSIEAMTTVKSVVFDIQ

SPy1375

Seq ID 208

MSLKDLGDISYFRLNNEINRPVNGKIPLHKDKKEALKAFSAENVLPNTMSFTSITEKIEYLISNDYIESAFIQK
YRPEFITELDSI IKSENFRF

KSFMAAYKFYQQYALKTNDGEHYLENLEDRVLFNALYFADGQEDLAKDLAVEMINQRYQPATPSFLNAGRSRR
GELVSCFLIQVTDD

MNSIGRSINSALQLSRIGGGVGITLSNLREAGAPIKGYAGAASGVVPVMKLFEDSFYSYNQLGQRQGAGVVYL
NVFHPDIIAFLSTKKE

NADEKVRVKTLSLGITVPDKFYELARKNEDMYLFSYPNVEKEYGIPFNYLDITNMYDELVANPKITKTIKAR
DLETEISKLQQESGYPII

INIDTANKANPIDGKIIMSNLCSEILQVQTPSLINDAQEFVEMGTDISCNLGSTNILNMMTSPDFGRSIKTMT
RALTFVTDSSSIEAVPTIK

HGNSQAHTFGLGAMGLHSYLAQHHEYGSPESIEFTDIYFMLLNWTLVESNNIARERQTTFVGFENSKYANG
SYFDKYVTGHFVPKS

DLVKDLFKDHFIPQASDWEALRDAVQKDGLYHQNRLAVAPNGSISYINDCSASIHPIQRIEERQEKKIGKIY
YPANGLSTDTIPYYTSA

YDMDMRKVIDVYAAATEHVDQGLSLTLFLRSELPMELEYEWKTQSKQTTRDLSILRNYAFNKGIKSIYYIRTFT
DDGEEVGANQCESCVI

SPy1389

Seq ID 209

MKELSSAQIRQMWLDFWKSKGHCVEPSANLVPVNDPTLLWINSGVATLKKYFDGSSIPENPRITNAQKSIRTN
DIENVGKTARHHTMF
EMLGNFSIGDYFRDEAIEWGFELLTSPDWDFPKDKLYMTYYPDDKDSYNRWIACGVEPSHLVPIEDNFWEIG
AGPSGPDTEIFFDR
GEDFDPENIGLRLLAEDIENDRYIEIWNIVLSQFNADPAVPRSEYKELPNKNIDTGAGLERLAAVMQGAKTNF
ETDLFMPIIREVEKLSG
KTYDPDGDNMSFKVIADHIRALSFAIGDGALPGNEGRGYVLRRLRRRAVMHGRRLGINETFLYKLVPTVGQIM
ESYYPEVLEKRDFIEK
IVKREETFARTIDAGSGHLDSLLAQLKAEGKDTLEGKDIFKLYDTYGFVPELTEELAEDAGYKIDHEGFKSA
MKEQQDRARA AVVKG
GSMGMQNETLAGIVEESRFEYDTYSLESSLSVI IADNERTEAVSEGQALLVFAQTPFYAEMGGQVADTGRIKN
DKGDTVAEVVDVQK
APNGQPLHTVNVLASLSVGNTYTLINKERRLAVEKNHTATHLLHAALHNVIGEHAQAGSLNEEEFLRFDF
HFEAVSNEELRHIEQE
VNEQIWNALITTTETDVETAKEMGAMALFGEKYGKVVRVQIGNYSVELCGGTHLNNSEIGLFKIVKEEGI
GSGTRRIIAVTGRQAF
EAYRNQEDALKEIAATVKAPQLKDAAAKVQALSDSLRDLQKENAELKEKAAAAAGDVFKDVQEAKGVRFIAS
QVDVADAGALRTFA
DNWKQKDYSVDVLVLVAAIGEKVNVLVASKTKDVHAGNMIKELAPIVAGRGGGKPDMAAGGSDASKIAELLAA
VAEIV

SPy1390

Seq ID 210

MKNSNKLIASVVTLASVMALAAACQSTNDNTKVISMKGDTISVSDFYNETKNTEVSQKAMNLVISRVFEAQYG
DKVSKKEVEKAYHKT
AEQYGASFSAALAQSSLTPETFKRQIRSSKLVEYAVKEAAKKELTTQEYKKAYESYPTMAVEMITLDNEETA
KSVLEELKAEGADFTA
IAKEKTTTPEKKVITYKFDGATNVPTDVVKAASSLNEGGISDVISVLDPTSYQKKFYIVKVTKKAEEKSDWQE
YKKRLKAI IIAEKSKDM
NFQNKVIANALDKANVKIKDKAFANILAQYANLGQKTKAASESSTSESSKAAEENPSESEQTQTSSAEPT
TEAQTQEPAAQ

SPy1422

Seq ID 211

VLYPTPIAKLIDSYSKLPGIGIKTATRLAFYTIGMSNEDVNDFAKNLLAAKRELYCSICGNLTDDDPCHICT
DTSRDQTTILVVEDAKDV

SAMEKIQEYHGYHVLHGLISPMNGVGPDDINLKSLITRLMDGKVSEVIVATNATADGEATSMYISRVLKPAG
IKVTRLARGLAVGSDIE
YADEVTLLRAIENRTEL

SPy1436

Seq ID 212

MDMSKSNRRTWQGLVVILIAILTTFTTSTVTAARKIRNFPDTTEILLGKATETPGILPFTGSYQLVLGDLN
LQRPTFAHIQLKDQDEPN
IKRKGLKFNPPGWHNYKLTANGKTTWMDRGHLVGYQFSGLNDEPKNLVTMTKYLNTGFSDKNPLGMLYYEN
RLDSWLALHPNF
WLDYKVTPVYHKNELVPRQVVLQYVGIDENGDLQIKLGSEKESVDNFGVTSVTLDNVSPLAELDYQTGMMLD
STQNEEDSNLETEE
FEEAA

SPy1494

Seq ID 213

MTSKKACLSSIIVLASLTCGNDTVSANHLSATGDKFDDCSTLVEKDVAPKDELEMLAWSSSQTTDDADRDYED
FLDDDSFISQNETDK
MFENLTDDRLLNELDELDEENEDEEDTIEPEQNVIMPSDDELFDLTDVETRLTVSSAPHLEAELPKPHLRS
LSDTALRSGEIRGHLD
NKLDALSVTATKLALTMAQKFDLTTHVYSIGESFSEVLAHYEDRKAESAFSKKKRFHLPATPDVVIEELRR
LVSSIGSSKEDVSVYPYS
RKLGMAMAKRKIALPQTGERFSYYPVLLGLMILGLTPIMIPKKINN

SPy1523

Seq ID 214

MAKDKEKQSDDKLVLTEWQKRNIEFLKKKKQQAEEKKLKEKLLSDKKAQQAQNAASEAVELKTDEKTDSEI
ESETTSKPKKTKKV
RQPKEKSATQIAFQKSLPVLLGALLMAVSIFMITPYSKKKEFSVRGNHQTNLDELKASKVKASDYWLTLLT
SPGQYERPILRTPWVK
SVHLSYQFPNHFLFNVIEFEIIAYAQVENGFQPILENGKRVDKVRASELPKSFLILNLKDEKAIQQLVKQLTT
LPKKLVKNIKSVSLANSKT
TADLLLIEMHDGNVVRVPQSQLTLKLPYYQKLKKNLENDISVDMEVGITYTTQEIQPEVPLTPEQNAADKE
GDKPGEHQEQTDNDS
ETPANQSSPQQTTPPETVLEQAHG

SPy1536

Seq ID 215

MKRLKKIKWWLVGLLALISLLLALFFPLPYYIEMPGGAYDIRTVLQVNGKEDKRKGAYQFVAVGISRASLAQL

LYAWLTPFTEISTAEDT

TGGYSADDFLRINQFYMETSQNAAIYQALSLAGKPVTLTDYKGVYVLDVNNESTFKGTLHLADTVTGVNGKQFT
SSAELIDYVSHLKLKLG

DEVTVQFTSDNPKPKGVGRIIKLKNGKNGIGIALTDHTSVNSEDTVIFSTKGVGGPSAGLMFTLDIYDQITKE
DLRKGRTIAGTGTIGKD

GEVGDIGGAGLKVVAAAAEAGADIF FVPNNPVDKEIKKVNPNAISNYEEAKRAAKRLKTKMKIVPVTTVQEAL
VYLRK

SPy1564

Seq ID 216

MLEHKIDFMVTLEVKEANANGDPLNGNMPRTDAKGYGVMSDVSIRKIRNRLQDMGKSIFVQANERIEDDFRS
LEKRFSQHFTAKTP

DKEIEEKANALWFDVRAFGQVFTYLKKSIGVRGPVSISMAKSLEPIVIVSSLQITRSTNGMEAKNNSGRSSDTM
GTKHFVDYGVYVLKG

SINAYFAEKTGFSQEDAEAIKEVLVSLFENDASSARPEGSMRVCEVFWFTHSSKLGNVSSARVFDLLEYHQSI
EEKSTYDAYQIHLNQ

EKLAKYEAKGLTLEILEGL

SPY1604

Seq ID 217

MATKKVHIISHSHWDREWMAYEQHHMRLINLIDDLLEVFQTDPDFHSFHLDGQTIILDDYLKVRPEREPEIR
QAIASGKLRIQPFYILQ

DDFLTSSSESNVRNMLIGKEDCDRWGASVPLGYFPDFTFGNMGQTPQLMLKAGLQAAAFGRGIRPTGFNNQVDT
EKYSSQFSEISW

QGPDNSRILGLLFANWYSNGNEIPTTEAEARLFWDKKLADAERFASTKHLLMMNGCDHQPVQLDVTKAIALAN
QLYPDYEFVHSCFE

DYLADLADDLPENLSTVQGEITSQETDGWYTLANTASARIYLKQANTRVSRQLENITEPLAAMAYEVTSTYPH
DQLRYAWKTLMQNH

PHDSICGCSVDSVHREMMTRFEKAYEVGHYLAKEAAKQIADAIDTRDFPMDSQPFVLFNTSGHSKTSVAELSL
TWKKYHFGQRFPKE

VYQEAQEYLARLSQSFIIDTSGQVRPEAEILGTSIAFDYDLPKRSFREPYFAIKVRLRLPITLPAMSWKTLA
LKLGNETTPSETVSLYD

DSNQCLENGFLKVMIIQTDGRLTITDKQSGLIYQDLLRFEDCGDIGNEYISRQPNHDQPFYADQGTIKLNIISN
TAQVAELEIQQTFAIPIS

ADKLLQAEMEAVIDITERQARRSQEKAELTLTTLIRMEKNPRLQFTTRFDNQMTNHLRLVLFPTHKTDHHL
ADSIFETVKRPNHPDA

TFWKNPSNPQHQECFVSLFDGENGVTIGNYGLNEYEILPDTNTIAITLLRSVGEMGDWGYFPTPEAQCLGKHS
LSYSFESITKQTQFA

SYWRAQEGQVPVITTTQTNQHEGTLAAEYSYLTGTNDQVALTAFKRRLADNALITRSYNLSNDKTCDFSLSLPN

YNAKVTNLLEKDSKQ

STPSQLGKAEILTLAWKKQ

SPy1607

Seq ID 218

MKITKIEKKRLYLIELDNDESLYVTEDTIVRFMLSKDKVLDNDQLEDMKHFAQLSYGKNLALYFLSFQQRSN
KQVADYLRKHEIEEHIIA

DIITQLQEEQWIDDTKLADTYIRQNQLNGDKGPQVLKQKLLQKGIASHDIDPILSQTDFSQLAQKVSQKLFDK
YQEKLPKALKDKITQA

LLTKGFSYDLAKHSLNHLNFDQDNQEIEDLLDKELDKQYRKLSRKYDGYTLKQKLYQALYRKGYNSDDINCKL
RNYL

SPy1615

Seq ID 219

MICLLCQQISQTPISITEIIFLRRISSPICQQCQKSFQKIGKSVCATCCANSDIACRDCLKWENKGYNVNHR
SLYCYNAAAMKAYFSQYKF

QGDYLLRKVF AVELADVITKYKGYIPVPVPVSPGCFRERQFNQVSAILEAANVSYSLFEKLDNTHQSSRTK
KERLLVEKSYRLLKVS

NIPDKILIVDDIYTTGSTIIALRKQLAKVANSDIKSLSIAR

SPy1666

Seq ID 220

MKSFSLTFSFLNLLKYGTIKVMTKEFHHTVLLHETVDMLDIKPDGIYVDATLGGSGHSAYLLSKLGEEGHL
CFDQDQKAIDNAQVTL

KSUIDKGQVTFIKDNFRHLKARLTALGVDEIDGILYDLGVSSPQLDERERGF SYKQDAPLDMRMDRQSLLTAY
EVVNTYPFNDLVKIF

KYGEDKFSKQIARKIEQARA IKPIETTTTELAE LIKAAKPAKELKKKGHPAKQIFQAIRIEVNDELGAADESIQ
DAMELLALDGRISVITFHSL

EDRLTKQLFKEASTVDVPKGLPLIPEDMKPKFELVSRKPILPSHSELTANKRAHSAKLRVAKKIRK

SPy1727

Seq ID 221

VTTTEQELTLTPLRGKSGKAYKGTYPNGECVFIKLNTTPILPALAKEQIAPQLLWAKRMGNGDMMSAQEWLNG
RTLTKEDMNSKQIIH

ILLRLHKS KKL VNQLQLNYKIENPYDLLVDFEQNAPLQIQQNSYLQAI VKELKRSLPEFKSEVATIVHGD
HSNWWITTSGMIFLVDWD

SVRLTDRMYDVAYLLSHYIPRSRWSEWLSYYGYKNNDKVMQKIIWYGQFSLTQILKCFDKRDMEHVNQEIIYA
LRKFREIFRKK

SPy1785

Seq ID 222

MILTAPMSNLKGFPGKSAEFQKLDIYTVEDLLLYPPFRYEDFKSKSVFDLVDGEKAVITGLVVTPANVQYYG
FKRNRLSFKLRQGEAV

LNVSFFNQPYLADKIELGQEVAVFGKWDATKSAITGMKVLAQVEDDMQPVYRVAQGISQSTLIKAIKSAFEID
AHLELKENLPATLLEKY

RLMGRSQACLAMHFPKDI TEYKQALRRIKFEELFYFQMNQLVKAENKSETNGLPILYSKRAMETKISSLPFI
LTNAQKRSLDDILSDMS

SGAHMNRLQLQDVGSGKTVIAGLSMYAAYTAGFQSALMVPTEILAEQHYISLQELFPDLSIAILTSGMKA AVK
RTVLAAIANGSVDMIV

GTHALI QDSVQYHKLGLVITDEQHRFGVKQRRIFREKGENPDVLMMTATPIPRTLAITAFGEMDVSIIDELPA
GRKPIMTRVVVKHEQLG

TVLEWVKGELQKDAQVYVISPLIEESEALDLKNAVALHAELSTYFEGIAKVALVHGRMKNDEKDAIMQDFKDK
KSHILVSTTVIEVGVNV

PNATIMIIMDADRFGLSQLHQLRGRVGRGYKQSYAVLVANPKTDSGKKRMTIMTETTDGFVLAESDLKMRGSG
EIFGTRQSGIPEFQV

ADIVEDYPILEEARKVSAAIVSDPNWIYEKQWQLVAQNIRKKEVYD

SPy1798

Seq ID 223

MKKISKCAFVAISALVLIQATQTVKSQEPLVQSQLVTTVALTQDNRLLEEIGPYASQSAGKEYYKHIEKIIIV
DNDVYEKSLEGERTFDIN

YQGIKINADLIKDGKHELTIVNKKDGDILITFIKKGDKVTFISAQKLGTTDHQDSLKKDVLSDKTVPQNQGTQ
KVVKSGKNTANLSLITKL

SQEDGAILFPEIDRYSDNKQIKALTQQITKVTVNGTVYKD LISDSVKDTNGWVS NMTGLHLGTKAFKD GENTI
VISSKGFEDVTITVTKK

DGQIHVSAKQKQHVTAE DRQSTKLDVTTLEKAIKEADAI IAKESNKDAVKDLAEKLQVIKDSYKEIKDSKLL
ADTHRLLKDTIESYQAGE

VSINNLTGTYTLNFKANKENSEESSMLQGAFDKRAKLVVKADGTMEISMLNTALGQFLIDFSIESKGTYPAA
VRKQVGQKDINGSYIR

SEFTMPIDDLDKLHKGAVLVSAMGGQESDLNHYDKYTKLDMTFSKTVTKGWSGYQVETDDKEKGVGTERLEKV
LVKLKGKDLGDGK

LSKTELEQIRGELRLDHYELTDISLLKHAKNITELHLDGNQITEIPKELFSQMKQLRFLNLR SNHLYLDKDT
FKSNAQLRELYLSSNFIH

SLEGGFLQSLHHLEQLDLSKNRIGRLCDNPFEGLSRLTSLGFAENSLEEIPEKALEPLTSLNFIDLSQNNLAL
LPKTIEKLRLALSTIVASR

NHITRIDNISFKNLPKLSVLDLSTNEISNLPNGIFKQNNQLTKLDDFFNNLLTQVEESVFPDVETLNLDVKFNQ
IKSVSPKVRALIGQHKLTP

QKHIAKLEASLDGEKIKYHQA FSLLDLYYWEQKTNSAIDKELVSVEEYQQLLQEKGS DTVSLLNDMQVDWSIV

IQLQKKASNGQYVTV

DEKLLSNDPKDDL TGEFSLKDPGTYRIRKALITKKFATQKEHIYLT SNDILVAKGPHSHQKDLVENGLRALNQ
KQLRDGIYYLNASMLKT

DLASESMSNKAINHRVTLVVKKGVSYLEVEFRGIKVGKMLGYLGELSYFVDGYQRDLAGKPVGR TKKAEVVSY
FTDVTGLPLADRYG

KNYPKVLRMKLI EQAKKDGLVPLQVFVPIMDAISKGSGLQTVFMRLDWASLTTEKAKVVKETNNPQENSHLTS
TDQLKGPQNRQKEK

TPTSPPSAATGIANLTDLLAKKATGQSTQETSKTDDTDKAEKLKQLVRDHQTSIEGKTAKDTKTKKSDKKHRS
NQQSNGEESSRYH

LIAGLSSFMIVALGFIIGRKTLFK

SPy1801

Seq ID 224

MNKNKLLRVAMLLSLLAPTAESMTVLAQDVMLETHKATTNETSDSSSKEENNKNAAPTTS DKTDQGPLDASAE
TNSNSLVNADDKKR

SDSSQSAIGSSDNKAEAEENQVDDKSTDHKSSTDHKSPTDQPKPSKVD TAPASSLSKQLPEARTPIQSLSPY
VSDLDLSEIDIPSVN

TYAAYVEHWSGKNAYTHHLLSRRYGIKADQIDSYLKSTGIA YDSTRINGEKLLQWEKKSGLDVRAI VAIAMSE
SSLGTQGIATLLGANM

FGYAAF DLDPTQASKFNDDSAIVKMTQDTIIKNKNSNFALQDLKAAKFSRGQLNFASDGGVYFTDTTGSGKRR
AQIMEDLDKWIDDH

GGTPAIPAELKVQSSASFASVPAGYKLSKSYDVLGYQASSYAWGQCTWYVYNRAKELGYQFDPFMGNGGDWKY
KVGYALSKTPKV

GYAISFAPGQAGADGTYGHVSIVEDVRKDG SILISESNCIGLGKISYRTFTAQQAEQLTYVIGKSKN

SPy1813

Seq ID 225

MDKHLLVKRTLGCVCAATLMGAALATHHDSLNTVKAEEKTVQVQKGLPSIDSLHYLSENSKKEFKEELSKAGQ
ESQKVKEILAKAQQA

DKQAQELAKMKIPEKIPMKPLHGSLYGGYFRTWHDKTS DPTEKDKVNSMGELPK EVDLAFIFHDWTKDYSLFW
KELATKHVPKLNKQ

GTRVIRTIPWRFLAGGDNSGIAEDTSKYPNTPEGNKALAKAI VDEYVYKYNLDGLDVDVEHDSIPKVDKKEDT
AGVERSIQVFEEIGKLI

GPKGVDKSRLFIMDSTYMADKNPLIERGAPYINLLL VQVYGSQGEKGGWEPVSNRPEKTMEERWQGYSKYIRP
EQYMIGFSFYEEEN

AQEGNLWYDINSRKDEDKANGINTDITGTRAERYARWQPKTGGVKGGIFS YAI DRDGV AHQPKKYAKQKEFKD
ATDNIFHSDYSVSK

ALKTVMLKDKSYDLIDEKDFDPKALREAVMAQVGTRKGD LERFNGTLRLDNPAIQSLEGLNKFKKLAQLDLIG
LSRITKLD RSVLPANM

KPGKDTLETVLETYKKDNKEEPATIPPVSLKVSGLTGLKELDLSGFDRETLAGLDAATLTSLEKVDISGNKLD
LAPGTENRQIFDTMLST

ISNHVGSNEQTVKFDKQKPTGHYPDTYGKTSRLRPVANEKVDLQSQLLFGTVTNQGTLINSEADYKAYQNHKI
AGRSFVDSNYHYNN

FKVSYENYTVKVTDSTLGTTTDKTLATDKEETYKVDFSPADKTKAVHTAKVIVGDEKTMNVNLAEGATVIGG
SADPVNARKVFDGQL

GSETDNISLGWDSKQSIIFKLKEDGLIKHWRFFNDSARNPETTNKPIQEASLQIFNIKDYNLDNLENPNKFD
DEKYWITVDTYSAQGE

RATAFSNTLNNITSKYWRVVFDTKGDYSSPVPELQILGYPLPNADTIMKTVTTAKELSQQKDKFSQKMLDE
LKIHEMALETSLSKI

FDVTAINANAGVLKDCIEKRQLLKK

SPy1821

Seq ID 226

MIEASKLKAGMTFEAEGKLIRVLEASHHKPGKGNTIMRMKLRDVRTGSTFDTTYRPDEKFEQAI IETVPAQYL
YKMDDTAYFMNTDTY

DQYEIPVANVEQELLYILENSDVKIQFYGSEVIGTVPTTVELTVAETQPSIKGATVTGSGKPATLETGLVVN
VPDFIEAGQKLIINTAEG

TYVSRA

SPy1916

Seq ID 227

MTKTLPKDFIFGGATAAYQAEGATHTDGKGPAWDKYLEDNWYTAEPASDFYNRYPVDLKLSEEFVNGIRI
SIAWSRIFPTGKGEV

NPKGVEYYHNLFAECHKRHVEPFVTLHHFDTPREALHSDGDFLNRENIEHFVNIAEFCKEFSEVNYWTTFNEI
GPIGDGQYLVGKFPP

GIQYDLAKVFQSHHNMVSHARAVKLFKDSGYSGEIGVVHALPTKYPFDANNPDDVRAAELEDI IHNKFILDA
TYLGKYSDKTMEGVN

HILEVNGGELDLREEDFAALDAAKDLNDFLGINYMSDWMQAFDGETE I IHNGKGEKGSSKYQIKGVGRRKAP
VDVPKTDWDWIIFP

QGLYDQIMRVKADYPNYKKIYITENGLGYKDEFVDNTVYDGGRIDYVKKHLEVISDAISDGANVKGYFMWSLM
DVFSWSNGYEKRYG

LFYVDFETQERYPKKSAYWYKKVAETQVIE

SPy1972

Seq ID 228

MKKKVNQGSKRYQYLLKKWGIGFVIAATGTVVLGCTPSILTHQVAAKTIVGLARDEAQQGDGNAKSGDGLQSS
SKEAKPVLDSSSAN

PASIAEHHLRMHFKTLPAGEISLGSGLWVGVDQPSKDWPNGAITMTKAKKDDYGYLDVPLAAKHRQQVSY

LINNKAGENLSKD

QHISLLTPKMNEVWIDENYHAHAYRPLKKGYLRYNYHNQSGHYDNLAVWTFKDVKTPTTDWPNGLDLSHKGHY
GAYVDVPLKEGAN

EIGFLILDKSKTGDAIKVQPKDYLFKELDNHTQVFVKDTPKVVNNPYYIDQVSLKGAEQTPNEIKAIFTTL
DGLDEDAVKQNIKITD KA

GKTVAIDELTLDRDKSVMTLKGDFAQGAIVTVTFGEVSQVARQSWQLKDKLYAYDGELGATLAKDGSVDLAL
WSPSADTVKVVVY

DKQDQTRVVGQADLTSDKGVWRAHLTSDSVKGISDYTGYYLYEITRGQEKVMVLDPYAKSLAAWNDATATD
DIKTAKAAFIDPSK

LGPTGLDFAKINNFKKREDAI IYEAHVRDFTSDKALEGKLTHPFGTFSAFVEQLDYLKDLGVTHVQLLPVLSY
FYANELDKSRSTAYTS

SDNNYNWGYDPQHYFALSGMYSANPNDPALRIAELKNLVNEIHKRGMGVIFDVVYNHTARTYLFEDLEPNYYH
FMNADGTARESG

GGRLGTTHAMSRRLVDSITYLTREFKVDGFRFDMMGDHDAAAIEQAFKAAKAINPNTIMIGEGWRTYQGDEG
KKEIAADQDWMKAT

NTVGVFSDDIRNTLKSGFPNEGTAAFITGGAKNLEGLFKTIKAQPGNFADAPGDVVQYIAAHDNLTLHDVIA
KSINKDPKVAEEEEIHKR

IRLGNTMILTAQGTAFIHSGQEYGRTKQLLNPDYKTKASDDKVPNKATLIDAVAQYPYFIHDSYDSSDAVNHF
DWAKATDSIAHPISNQ

TKAYTQGLIALRRSTDAFTKATKAEVDRDVTLITQAGQDGIQQEDLIMGYQTVASNGDRYAVFVNADNKTRKV
VLPQAYRYLLGAQVL

VDAEQAGVTAIAKPKGVQFTKEGLTIEGLTALVLKVSSKTANPSQQKSQTDNHQTKTPDGSKDLKSLMTRPK
RAKTNQKLPKTGEA

SSKGLLAAGIALLLLAISLLMKRQKD

SPy1979

Seq ID 229

MKNYLSIGVIALLFALTFGTVKSVQAIAGYGWLPRPPINNSQLVVSMAGIVEGTDKKVFINFFEIDLTSQPA
HGGKTEQGLSPKSKPFA

TDNGAMPHKLEKADLLKAIQKQLIANVHSNDGYFEVIDFASDATITDRNGKVYFADKGSVTLPTQPVQEFL
KGHVRVRPYKEKPVQ

NQAKSVDVEYTVQFTPLNPDDDFRPLKDTKLLKTLAIGDTITSQELLAQAQSILNKTHPGYTIYERDSSIVT
HDNDIFRTILPMDQEFTY

HVKNREQAYEINPKTGIKEKTNNTDLVSEKYYVLKQGEKPYDPFDRSHLKLFTIKYVDVNTNELLKSEQLLTA
SERNLDFRDLYDPRDK

AKLLYNNLDAFDIMDYTLTGKVEDNHDKNRVTVMGKRPKGAKGSYHLAYDKDLYTEEERKAYSYLRTGT
PIPDNPKDK

SPy1983

Seq ID 230

MLTSKHHNLNKLWVRYGLTSAAAVLLAFGGGASSVKAEVSSTTMTSSQRESKIKEIEESLKKYPEVSNEKFEW
RKWYGTYFKEEDFQ

KELKDFTEKRLKEILDIGKSGIKGDRGETGPAGPAGPQGKTGERGAQGPKGDRGEQGIQ GKAGEKGERGEKG
DKGETGERGEKG

EAGIQGPQGEAGKDGAPGKDGAPGEKGEKGD RGETGAQGPVGPQGEKGETGAQGPAGPQGEAGKPGEQGPAGP
QGEAGQPGE

KAPEKSPEGEAGQPGEKAPEKSKEVTPAAEKPADKEANQTPERRNGNMAKTPVANNHRRLPATGEQANPFFTA
AAVAVMTTAGVL

AVTKRKENN

SPy1991

Seq ID 231

MILLIDNYDSFTYNLAQYLSEFDETI VLYNQDPNLYDMAKKANALVLSPGPGWPKEANQMPKLIQDFYQTKPI
LGVCLGHQAIAETLGG

TLRLAKRVMHGRQSTIETQGPASLFRSLPQEITVMRYHSIVVDQLPKGFSVTARDCDDQEIMAFEHHTLPLFG
LQFHPESIGTPDGMT

MIANFIAAIPR

SPy2000

Seq ID 232

VSKYLKYFSIITLFLTGLILVACQQQKPQTKERQRKQRPKDEL VVSMGAKLPHEFDPKDRYGVHNEGNITHST
LLKRSPELDIKGELAK

TYHLS EDGLTWSFDLHDDFKFSNGEPVTADDVKFTYDMLKADGKA WDLTFIKNVEVVGKNQVNIHLTEAHSTF
TAQLTEIPIVPKKHY

NDKYKSNPIGSGPYMVKEYKAGEQAIFVRNPYWHGKKPYFKKWTWVLLDENTALAALESGD VDMIYATPELAD
KKVKGTRLLDIPSN

DVRGLSLPYVKKGVITDSPDGYPVGNDVTS DPAIRKALTIGLNRQKVLDTVLNGYGKPAYSII DKTPFWNPKT
AIKDNKVAKAKQLLTK

AGWKEQADGSRKKGDLDAAFDLYYPTNDQLRANLAVEVAEQAKALGITIKLKASNWDEMATKSHDSALLYAGG
RHHAQQFYESHHP

SLAGKGWTNITFYNNPTVTKYLDKAMTSSDL DKANEYWKLAQWDGKTGASTLGDLPNVWL VSLNHTYIGDKRI
NVGKQGVHSHGHD

WSLLTNIAEWTWDESTK

SPy2006

Seq ID 233

VKKTYGYIGSVAAILLATHIGSYQLGKHHMGSATKDNQIAYIDDSKGKAKAPKTNKTM DQISAEEGISAEQIV
VKITDQGYVTSHGDHYH

FYNGKVPYDAI I SEELLMTDPNYRFKQSDVINE ILDGYVIKVNNGYVYLKPGSKRKNIRTKQQIAEQVAKGT
KEAKEKGLAQVAHLSK

EEVAAVNEAKRQGRYTTDDGYIFSPTD I IDDLGDAYLVPHGNHYHYIPKKDLSPSELAQAQYWSQKQGRGAR
PSDYRPTPAPAPGR

RKAPIPDVTPNPGQGHQPDNGGYHPAPPRNDASQNKHQDEFKGGKTFKELLDQLHRLDLKYRHVEEDGLIFE
PTQVIKSNAFGYVV

PHGDHYHIIPRSQLSPLEMELADRYLAGQTEDDDSGSDHSPSDKEVTHTFLGHRIKAYGKGLDGKPYDTSDA
YVFSKESIHSVDKS

GVTAKHGDHFHYIGFGELEQYELDEVANWVKAKGQADELAAALDQEQGKEKPLFDTKKVSRKVTKDGKVGMM
PKDGKDYFYARD

QLDLTQIAFAEQELMLKDKKHRYD I VDTGIEPRLAVDVSSSLPMHAGNATYDTGSSFVIPHIDHIHVVPYSWL
TRDQIATIKYVMQHPEV

RPDIWSKPGHEESGSVIPNVTPDKRAGMPNWQI IHSAAEVQKALAEGRFATPDGYIFDPRDVLAKETFWWD
GSFSIPRADGSSLRT

INKSDLSQAEWQQAQELLAKKNAGDATDTDKPKEKQQADKSNENQQPSEASKEEEKESDDFIDSLPDYGLDRA
TLEDHINQLAQKA

NIDPKYLIFQPEGVQFYKNKNGELVTYDIKTLQQINP

SPy2009

Seq ID 234

MRRANENKHSRYSIRKLSVGVTSIATIASLFLGKVAYAVDGIPPISLTQKTTATTSENWHHIDKDGLIPLGISL
EAAKEEFKKEVEESRLSE

AQKETYKQKIKTAPDKDLLFTYHSEYMTAVKDLPASTESTTQPVEAPVQETQASASDSMVTGDSTSVTTDSP
EETPSSESPVAPALS

EAPAQPAESEEPSVAASSEETPSPSTPAAPETPEEPAAPSPSPSEEPSVAAPSEETPSPETPEEPAAPSQA
ESEESSVAATTSPS

PSTPAESETQTPPAVTKDSKPSAAEKPAASSLVSEQTVQQPTSKRSSDKKEEQEQSYSPNRSLSRQVRAHE
SGKYLPSTGEKAQ

PLFIATMTLSLFGSLLVTKRQKETKK

SPy2010

Seq ID 235

LRKKQKLFPDKLAIALMSTSILLNAQSDIKANTVTEDTPATEQAVETPQPTAVSE EAPSSKETKTPQTPDDA
EETIADDANDLAPQAPA

KTADTPATSKATIRDLNDPSQVKTLQEAKAGGAGTVVAVIDAGFDKNHEAWRLTDKTKARYQSKEDLEKAKKE
HGITYGEWVNDKVA

YYHDYSKDGKTAVDQEHGTHVSGILSGNAPSETKEPYRLEGAMPEAQLLLMRVEIVNGLADYARNYAQAIIIDA
VNLGAKVINMSFGNA

ALAYANLPDETKKAFDYAKSKGVSI VTSAGNDSSFGGKTRLPLADHPDYG VVGTPAAADSTLTVASYS PDKQL

TETATVKTADQQDK

EMPVLSTNRFEPNKAYDYAYANRGMKEDDFKDVKGKIALIERGDIDFKDKIANAKKAGAVGVLIYDNQDKGFP
IELPNVDQMPAAFISR

KDGLLLKENPQKTITFNATPKVLPTASGTKLSRFSSWGLTADGNIKPDIAAPGQDILSSVANNKYAKLSGTSM
SAPLVAGIMGLLQKQY

ETQYP DMTPSERLDLAKKVLMSATALYDEDEKAYFSPRQQGAGAVDAKKASAATMYVTDKDNSSKVHLNN
VSDKFEVTVTVHNK

SDKPQELYYYQATVQTDKVDGKLFALAPKALYETSWQKITIPANSSKQVTIPIDVSQFSKDLLAPMKNGYFLEG
FVRFKQDPTKEELMSI

PYIGFRGDFGNLSALEKPIYDSKDGSSYYHEANSDAKDQLDGDGLQFYALKNNFTALTTESNPWTIIKAVKEG
VENIEDIESSEITETIFA

GTFKQDDDDSHYYIHRHANGKPYAAISPNGDGNRDYVQFQGTFLRNAKNLVAEVLCKEGNVVWTSEVTEQVVK
NYNNDLASTLGST

RFEKTRWDGKDKDGKVVANGTYTYRVRYTPISSGAKEQHTDFDVIDNTTPEVATSATFSTEDRRLTLASKPK
TSQPVYRERIAITY

MDEDLPTTEYISPNEGTFTLPEEAETMEGATVPLKMSDFTYVVEDMAGNITYTPVTKLLEGHSNKPEQDGS
QAPDKKPETKPEQD

GSGQAPDKKPETKPEQDGSQTPDKKPETKPEQDGSQTPDKKPETKPEKDSSGQTPGKTPQKGQPSRTLEKR
SSKRALATKAST

KDQLPTTNDKDTNRLHLLKLVMTTFFLGLVAHIFKTKRTED

SPy2016

Seq ID 236

MNIRNKIENSKTLLFTSLVAVALLGATQPVSAETYTSRNFDSGDDWSGDDWPEDDWSGDGLSKYDRSGVGLS
QYGWSKYGWSS

DKEEWPEDWPEDDWSDDKDETEKTRPPYGEALGTGYEKRDDWGGPGTVATDPYTPPYGGALGTGYEKRDDW
GGPGTVATDP

YTPPYGGALGTGYEKRDDWRGPGHIPKPENEQSPNPLHIPEPPQIEWPQWNGFDGLSFGPSDWGQSEDTPPSE
PRVPEKPQHTP

QKNPQESDFDRGFSAGLKAKNSGRGIDFEGFQYGGWSDEYKKGYMQAFGTPYTPSAT

SPy2018

Seq ID 237

MAKNNTNRHYSRLKLTGTASVAVALTVLGAGFANQTEVKANGDGNPREVIEDLAANNPAIQNIRLRYENKDL
KARLENAMEVAGR

FKRAEELEKAKQALEQQRKDLETKLQDYDLAKESTSWDRQRLEKELEEKKEALELAIDQASRDYHRATA
LEKELEEKKKALELA

IDQASQDYNRANVLEKELETITREQEINRNLLGNAKLELDQLSSEKEQLTIEKAKLEEEKQISDASRQSLRRD
LDASREAKKQVEKDLA

NLTAE LDKVKEDKQISDASRQGLRRDLDA SREAKKQVEKDLANLTAE LDKVKEEKQISDASRQGLRRDLDA SR
EAKKQVEKALEEAN

SKLAALEKLNKELEESKKLTEKEKAELQAKLEAEAKALKEQLAKQAEELAKLRAGKASDSQTPDTKPGNKAVP
GKGQAPQAGTKPNQ

NKAPMKETKRQLPSTGETANPFFTAAAL TVMATAGVAAVVKRKEEN

SPy2025

Seq ID 238

MKKRKLLAVTLLSTILLNSAVPLVVADTSLRNSTSSDQPTTADTDTDDSETPKKDKKSKETASQHDTQKDH
KPSHTHTPPSNDTK

QTDQASSEATDKPNKDNDTKQPDSSDQSTSPKQSSQKESQNKDGRPTPSPDQQKDQTPDKTPEKSADKTP
EKGPEKATDKTP

EPNRDAPKPIQPPLAAAPVFI PWRESKDLSKLPSSSRSSAAYVRHWTGDSAYTHNLLSRRYGITAEQLDGFL
NSLGIHYDKERLNGK

RLLEWEKLTGLDVRAI VAIAMAESSLGTQGVAKEKGANMFGYGAFDFNPNNAKKYSDEVAIRH MVEDTIIANK
NQTFERQDLKAKKW

SLGQLDTLIDGGVYFTDTSGSGQRRADIMTKLDQWIDDHGSTPEIPEHLKITSGTQFSEVPVGYKRSQPQNVL
TYKSETYSFGQCTW

YAYNRVKELGYQVDRYMGNGGDWQRKPGFVTTHKPKVGYVVSFAPGQAGADATYGHVAVVEQIKEDGSILISE
SNVMGLGTISYRT

FTAEQASLLTYVVGDKLPRP

SPy2039

Seq ID 239

MNKKKLGVRLLSLLALGGFVLANPVFADQNFARNEKEAKDSAITFIQKSAAIKAGARSAEDIKLDKVN LGGEL
SGSNMYVYNISTGGFV

IVSGDKRSPEILGYSTSGSFDANGKENIASFMESYVEQIKENKKLDTTYAGTAEIKQPVVKSLLDSKGIHYNQ
GNPYNLLTPVIEKVKPG

EQSFVGQHAATGCVATATAQIMKYHNYPNKGLKDYTYTLSSNNPYFNHPKNLFAAISTRQYNWNNILPTYSGR
ESNVQKMAISELMA

DVGISVDMDYGPSSGSAGSSRVQRALKENFGYNQSVHQINRGDFSKQDWEAQIDKELSQNQPVYYQGVGKVG
HAFVIDGADGRN

FYHVNWGWGVSDGFFRLDALNPSALGTGGGAGGFNGYQSAVVGIKP

SPy2043

Seq ID 240

MNLLGSRRVFSKKCRLVKF SMVALVSATMAVTTVTLENTALARQTQVSNQVVLNDGASKYLNEALAWTFNDS
PNYYKTLGTSQITPA

LFPKAGDILYSKLDELGRTRTARGTLTYANVEGSYGVRSFGKNQNPAGWTGNPNHVKYKIEWLNGLSYVGDF

WNRSHLIADSLGG

DALRVNAVGTGTRTQNVGGRDQKGGMRYTEQRAQEWLEANRDGYLYEAAPIYNADELIPRAVVVSMQSSDNTI
NEKVLVYNTANGY
TINYHNGTPTQK

SPy2059

Seq ID 241

MRFLELLQKKFFPKAYQEKQFLMHQKTRLTPQHNQKQYSPNANHLDSSATKNSEQDPATALQRSRAYEGSPKS
RPAWLQKLEAVLP
SPQRPIRRFWRRYHIGKLLMILIGTLVLLGSLFYLSKTAKVSDLQDALKATTVIYDHKGEYAGSLSGQKGS
YVELNAISDDLENAVIAT
EDRTFYNSNGINLKRFLAVVTAGRFGGGSTITQQLAKNAYLSQDQTIKRKAREFFLALELTKKYSKKDILTM
YLNNSYFGNGVWVGVE
DASQKYFGTTAANLTLDEAATLAGMLKGPEIYNPYHSLKNATHRRDVLGAMVDAKKITQTKAQQARAVGLKN
RLADTYVGKTDDYK
YPSYFDAVISEAIATYGLSEKDIVNNGYKVYTELDQNYQTGMQTTFNDELFPVSAYDGSSAQAAASVALDPKT
GGVRGLIGRVNSSE N
PTFRSFNYATQAKRSPASTIKPLVVYAPAVASGWSIEKELPNTVQDFDGYQPHNYGNYESEDVPMYQALANSY
NIPAVSTLNDIGIDK
AFTYGKTFGLDMSSAKKELGVALGGSVTTNPLEMAQAYAAFANNGVIHPAHLINRIENARGEVLKTFDTKAKR
VVSQSVADKMTAMM
LGTFSNGTAVNANVYGYTLAGKTGTTETNFPDLAGDQWVIGYTPDVVISQWVGFNQTDENHYLTDSSAGTAS
AIFSTQASYILPYTK
GSQFHVDNAYAQNGISAVYGVNETGNQSGVDTQSIIDGLRKSQAEASQSLSKAVDQSGLRDKAQSIWKEIVDY
FR

SPy2110

Seq ID 242

MVSLEEDKVTVPDIKVIKRDGRLVNFDS TKIYSALLKASMKVTRMSPLVEAKLEAISDRIIAEIIERFPTNI
KIYEIQNIVEHKLLAANEYAI
AKEYINYRTQRDFARSQATDINFSDKLINKDQTVVNENANKDSDFNTQRDLTAGIVGKSIGLKMLPSHVAN
AHQKGDIHYHDLDYSP
YTPMTNCCLIDFKGMLANGFKIGNAEVESPKSIQTATAQISQIIANVASSQYGGCTADRIDEFLAPYAELNFK
KHMADAKKWIVETKRE
SYAFEKTQKDIYDAMQSLEYEINTLFTSNGQTPFTSLGFGLGTSWFEREIQKAILTIRINGLGSEHRTAIFPK
LIFTVKRGLNLEPDSPNY
DIKTLALECATKRMPDMLS YDKIIDLTGSFKSPMGCRSFLQGWKDENGQDVTSGRMNLGVVTLNLPRIAMES
NGDMDKFWELFNE
RMLISKDALIYRVERVTEAKPANAPILYQYGAFGKRLEKTGNVNDLFKNRRATVSLGYIGLYEVASVFYGGQW

EGNPDAKAFTLSIVKA

MKQACEDWSDEYGYHFSVYSTPSESLTDRFCRLDTEKFGIVTDITDKEYYTNSFHYDVRKSPTPF EKLD FEKD
YPEAGASGGFIHYC

EYPVLQQNPKALEAVWDYAYDRVGYLGTNTPIDKCYNCCFEGDFTP TERGFTCPNCGNNDPKTVDVVKRTCGY
LGNPQARPMVNG

RHKEISARVKHMNGSTIKYPGL

SPy2127

Seq ID 243

MRRNYSRVIDELRTDYGLNLVAIGQRLGTDPRTVGKWWQ GKHNPNQESRKKLNRLYREVKETMMTQVNI FEEA
NDNTKQVMQVITT

TNFHGQPLDIYGDIQEPLFLARAVAEMIDYTKTSQGYDVQAMLRKVDEDEKLKGMALEGTTKNFRSGQKVWF
LTEHGLYEVL MRSN

KPKAKEFRKAVKNILKEIRLNGYYMQGELVQELAQ PSTQKLPGISDLTYILNKLADLVMDNLADISNGIDRV
QQLVKLISL

SPy2191

Seq ID 244

MFKKENLKQRYFNFG LVALALTILAIIFAFSSKNADTKSYAKKSESKMVTIDKAPKNNHAITKEESKEKAKSI
ASEPIPTVENSVAPT VTE

EVPVVQQEVTQT VQQVSSVAYNPNNVLSNGNTAGIVGSQAAAQMAAATGVPQSTWEHI IARESNGNPNAANA
SGASGLFQTMPG

WGSTATVEDQVNAALKAYSAQGLSAWGY

SPy2211

Seq ID 245

MKNNNKWIIAGLASFLFPLSIIFIILLSMGIYNSDKTILASDAFHQYVIFAQNF RNIMHGSDSFFYTFTSGL
GINFYALMCYYLGSFFSPLL

FFFNLTSMPDAIYLF TLIKFG LIGLAACYSFHRLYPKISAF L MISISVFYSLMSFLTSQMELNSWLDVFILLP
LVILGLNKLITENKTRTYYS

ISLLFIQNYF GYMIALFCILYALVCLLR LNDFNKMFIAFVRFTAVSICAALTSALVILPTYLDLSTYGENLS
PIKQLVTNNAWFLDIPAKLSI

GVYDTTKFNALPMIYVGLFPLMLS VIYFTLES IPLKIKLANACLLTFIIISFY LQPLDLFWQGMHSPNMFLHR
YAWSFSIVILLACETLSRL

KEVTQIKAGFAFIFLIILTSLPYSFSQQYNFLPLTLFLLSVFLLLGYTISLFSFRNSQIPSTFISAFILIFSL
LESGLNTYYQLQGINK EWG FPS

RQIYNSQLKDINN LVNSVSKNSQPFFRMERLLPQTGND SMKFNYG ISQFSSVRNRLSSSLDRLGFQSKGTN
LNLRYQNNTIIMDSL

LGIKYNLSEGPPNKFGFTKLKTS GNTTLYQNHYSPLAILTRNVYKDVNLNVNTLDNQTKLLNQLSGKSLTYF

NLQPAQLISGANQFNG

QISAAQSDYQNSVTLNYQINIPKHSQLYVSIPNIIFSNDPAKEMRIQTDNHNFIYTTDNAYSFFDLGYFADAK
VATFSFVFPKNKQISFKE

PHFYLSLIESYLEAMNSIKQKNVHTYAKSNTVITDYNSTKKGSLIFTLPYDKGWSAQKDGKNLPVKKAAQGGFL
SVTIPKGKGRVILTFIP

NGFKLGLSLSCVGIIAYMLLYKYIDIKSKLL

ARF0450

Seq ID 246

fsrflptrrdysslwsascrnehynsqhhhgvtvsskqnprpl

ARF0569

Seq ID 247

sfiwekmpegs

ARF0694

Seq ID 248

kgeektevtkelkllelarwikdisddtdekteadeayydgdgteettv

ARF0700

Seq ID 249

lyqkkrllkksqrllsklmtsrllnkalmtskie

ARF1007

Seq ID 250

fvlqkyslwg

ARF1145

Seq ID 251

pismqkaiqvaiamveqnvwilskrlllnvlrnfsvphlpmfkpipevkpmqqliwp

ARF1208

Seq ID 252

frnifcdfscssfvscyqklkrkgynrtskrfl

ARF1262

Seq ID 253

wvtgckwwffc

ARF1294

Seq ID 254

lmlkakktnskvltsqptkkfnlqklfnqtnllkplslwvllqttls

ARF1316

Seq ID 255

prngwgrferyerlgrtrhdhvcysrngicspss

ARF1352

Seq ID 256

lmncpslhflqpkhkeqpvlkmlknyeskkqlvfk

ARF1481

Seq ID 257

kttkcnylkrpklvesrlqrtrfrricsrkhgryrrwirrflifltnkskklvkrvrkrllslvtapkes
kkmelsqflig

ARF1557

Seq ID 258

grrlpprlpqekskwllpy

ARF1629

Seq ID 259

fwspgsryfvrndandcqrtgfskcdfswgtkcryflkfagfssvrknvsivntgcwsgprcp

ARF1654

Seq ID 260

cvpsvkcsimiqintplsilfpntlvqagvifrvypigfplfllewqksqq

ARF2027

Seq ID 261

sdyfrhhapflkwlr saknnskdircpyyyangir

ARF2093

Seq ID 262

llllkqtsklnlllkanqkrsgtkssqvkw tasclttlkltkltlflhkftswttakllkltlq

ARF2207

Seq ID 263

hlfvkdvwstlkiwercsvcykkvvkkqelwqprlyqk

CRF0038

Seq ID 264

lyapqslsnpfldsipcdq

CRF0122

Seq ID 265

nrrdfldnvnllirvteplgtslvfdnfinraahinniglgiglneisctfqpfcvttk

CRF0406

Seq ID 266

qaplddhhnkptywsgyl

CRF0416

Seq ID 267

yflkktplkaakswllspfgemaktgfpwaffskinlpsaflkvpsvcrplseivlvdltlvseprvtspvkcl
pt

CRF0507

Seq ID 268

sknkrntdtgcnnkgksksvshnhsknnhakhisenakeaaistrynlpnernhitntsscknss

CRF0549

Seq ID 269

lfihrsrlildflvinflfvqiyyddfing

CRF0569

Seq ID 270

sfiwekrnpegs

CRF0628

Seq ID 271

ikhltqakqrmpvskvlvanplgskgiadsiqlrmkplavkryrsslstr

CRF0727

Seq ID 272

ppppnlppacktvktvstagrpvlawistgiprpssttvmelsasiktli

CRF0742

Seq ID 273

enefnqyyqdakykshkerltintfkrqgylr

CRF0784

Seq ID 274

trvapyfpqalaihsepvfliasfkafpsssaastav

CRF0854

Seq ID 275

spftksvltpsalklssikrpvkppknpvavlamskrekmrltlnplppanlsslamrftdprfiwsiltiks
legfkvtv

CRF0875

Seq ID 276

dhniywhyqlklsqvqtmtppl

CRF0907

Seq ID 277

pnlldhflpnnphqnhkakld

CRF0979

Seq ID 278

qcliiinninfrhnkntpscilkraslhdiifhaetllw

CRF1068

Seq ID 279

eklfrtarqrynfwvskkqimglvliavlksmrnsesklffyhifghidlsnasinnnqvrsllqslfhptv
pskkhfshtgiiwshhcfnlkftiifltfainknnhsmwtcplnvriikgfnahrlvnp
k

CRF1152

Seq ID 280

dnrcgcclstnldnkthsftarfvtncentf

CRF1203

Seq ID 281

igsinqlvsfalvtmdevktlftlitptfeec

CRF1225

Seq ID 282

yqvcqlpegvpqlivnadqtwfykaktkqrtkkqwrnsq

CRF1236

Seq ID 283

ilakieisqktlitiimrs

CRF1362

Seq ID 284

fpdvndkviardsrffgkertcllsnrfwegdkkvetnpndlgd

CRF1524

Seq ID 285

kfnvtllplcqnelvitlflfifchllllnrganissqkvikevr

CRF1525

Seq ID 286

etsallaarflrsadspkiiiclsprsgtkssnslprsrppaikikrllskpdka

CRF1527

Seq ID 287

krlddctliffwcinlntlkrftfdtinffddnfw

CRF1588

Seq ID 288

ewlsqqhshqflhdkrsyldtafdkgkctfhkpvlllrwssylg

CRF1649

Seq ID 289

hdhlshqqslkqlgnlgldsknhqndkyykesaahrg

CRF1749

Seq ID 290

vmlkapvklifktrsksssvilsinlsylvmpalltktsiepklsiasltirlasdasetspemvttvtp

CRF1903

Seq ID 291

relsrfsssaillkplvkvpnnrtmappnppaaplpgivcksmvwwqktflakllai

CRF1964

Seq ID 292

ehfqpnhqigqkwkkerpkptwsksdvahkqtyqsp

CRF2055

Seq ID 293

ksrslrtssitsstvfssptilvlpslwnicsgafcwdsghlahghlkgqpfkktvvripgps

CRF2091

Seq ID 294

rneintssptlsstki

CRF2096

Seq ID 295

cslaafatkfsksalvpksglta

CRF2104

Seq ID 296

eamsqqrccsfladsslierydlakssvviaamifkmawlsqgrrcg

CRF2116

Seq ID 297

nfkarppgvsvgfpnitptfsrnwllkiaivfvllieadnlrma

CRF2153

Seq ID 298

ketchyshqhvhvifqakfdmflphpdksgqhncsidg

NRF0001

Seq ID 299

itllfkqtqqfkrkpslsnnlikhllvkrmslsmlkqkhqltkpkasmlrpqnnkttisfhggdqtikihg
ptlts

NRF0003

Seq ID 300

sgrqdsnlrhlgpkpstlps

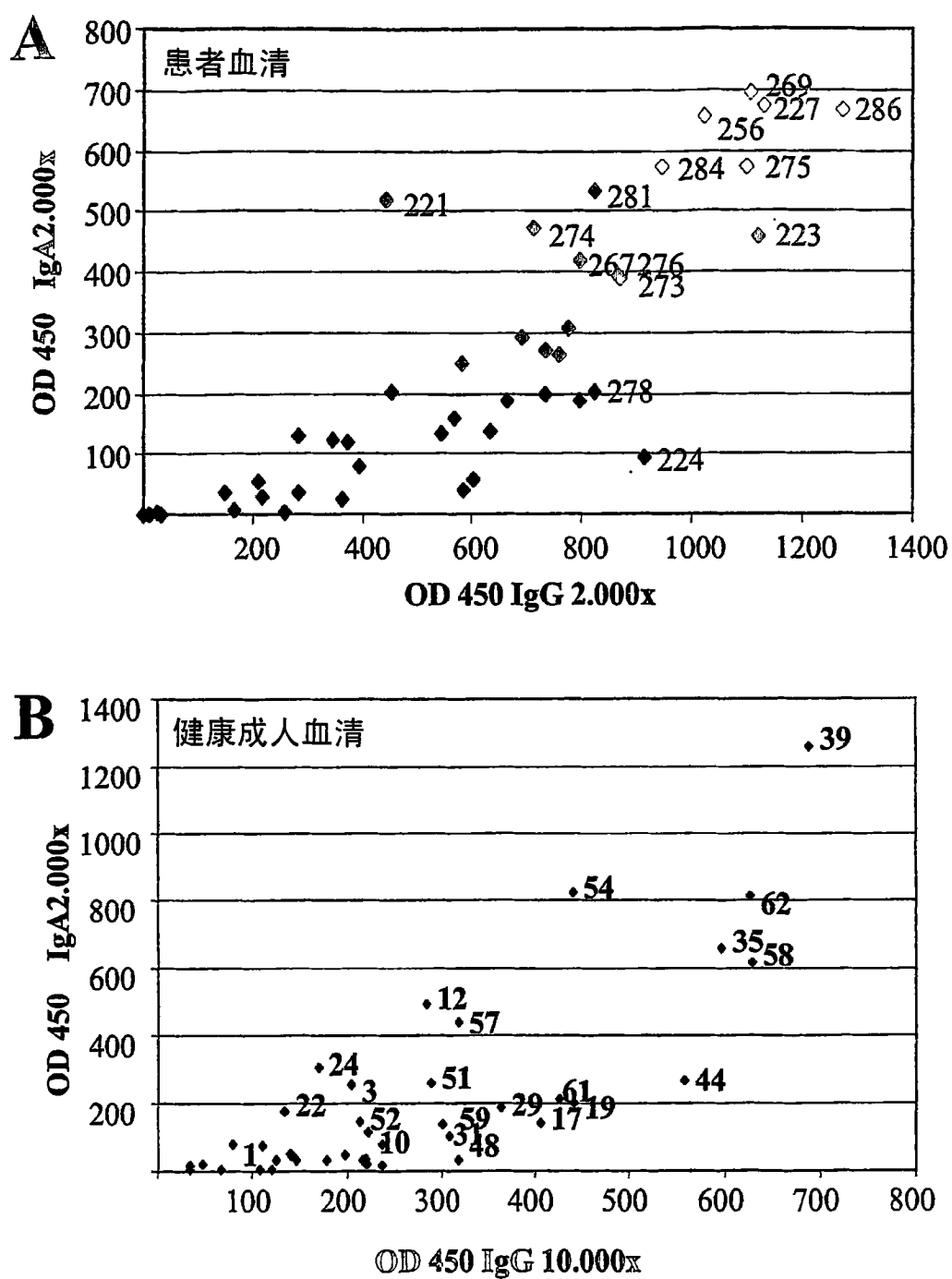


图 1

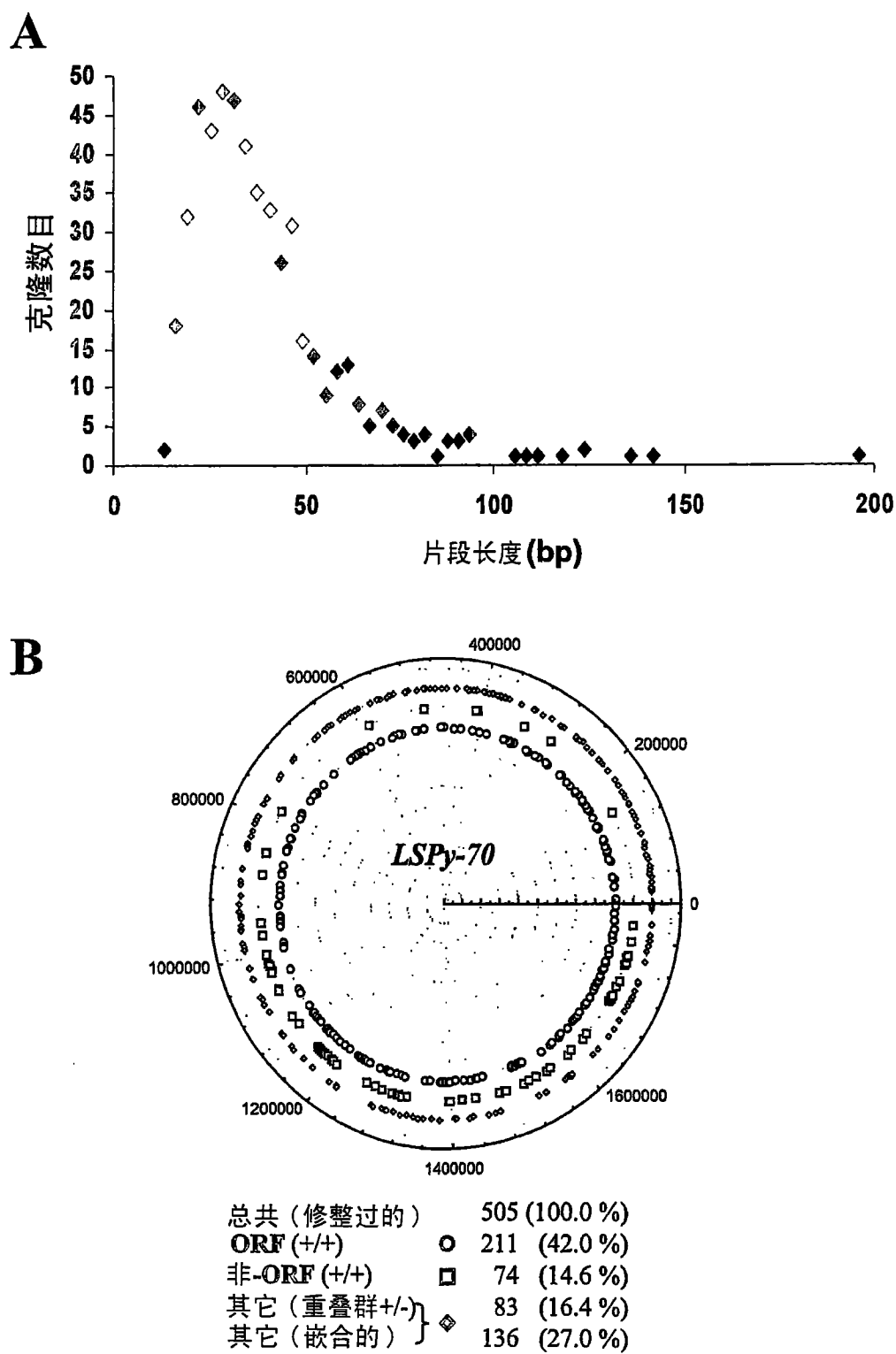


图 2

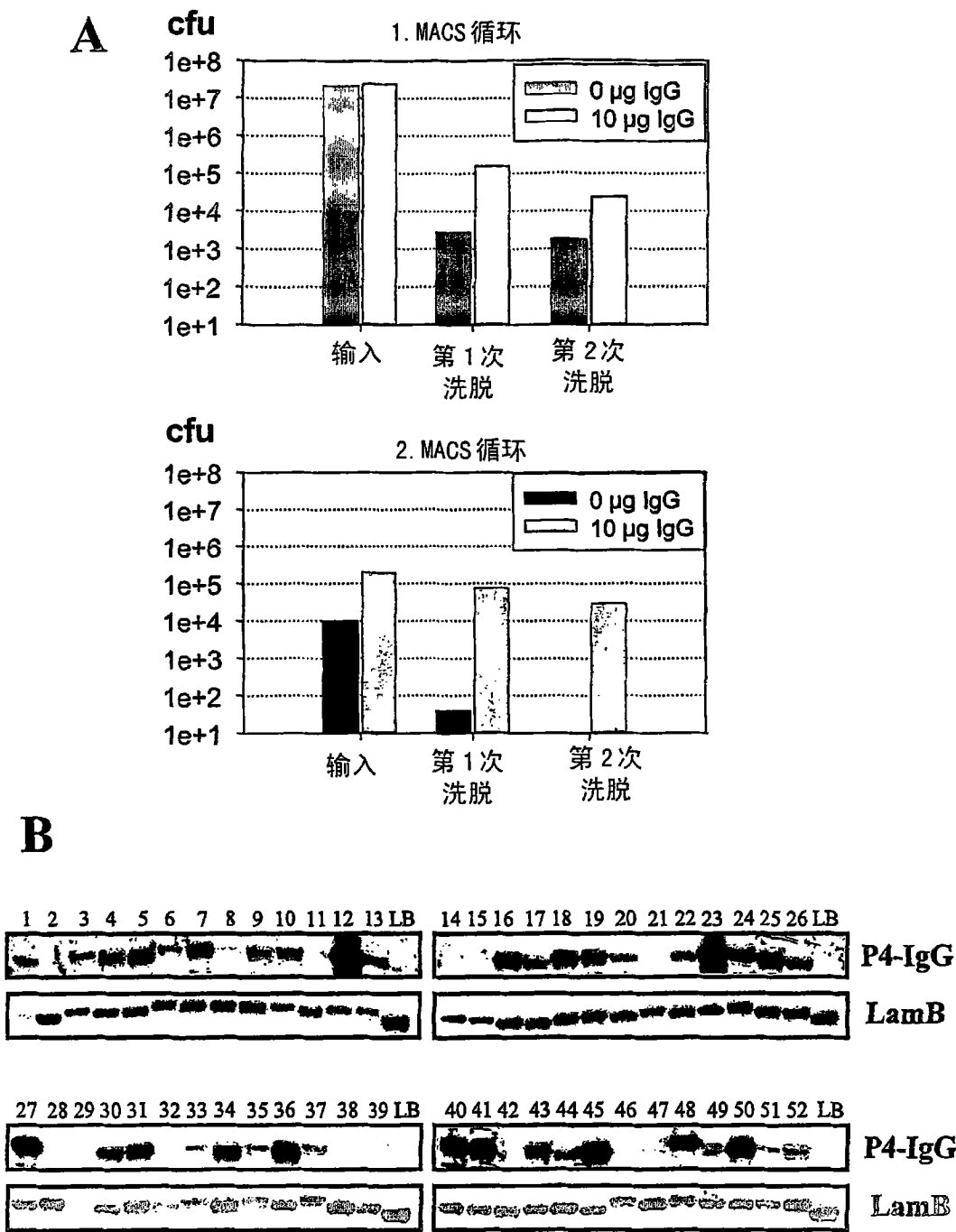


图 3

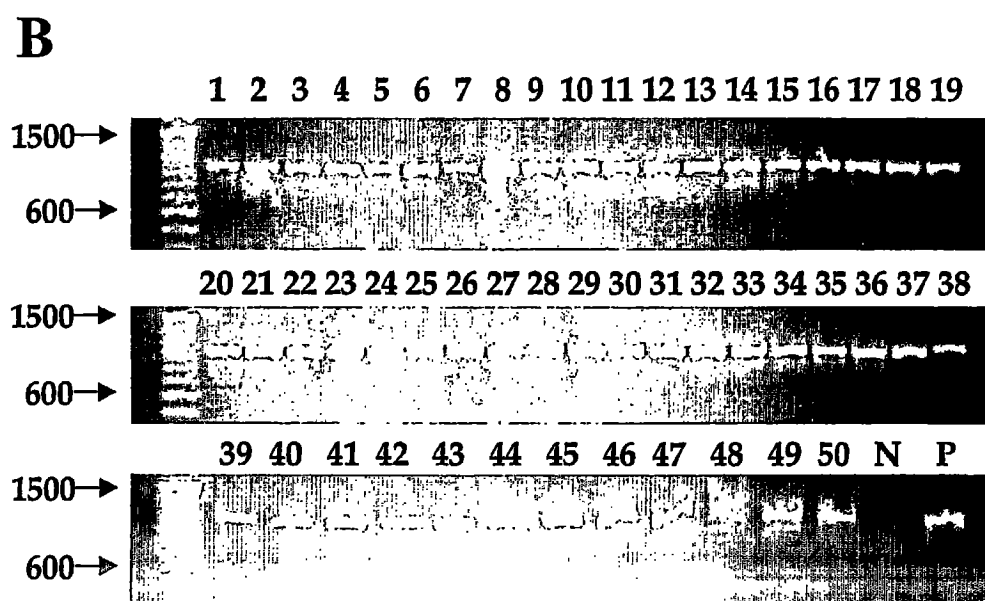
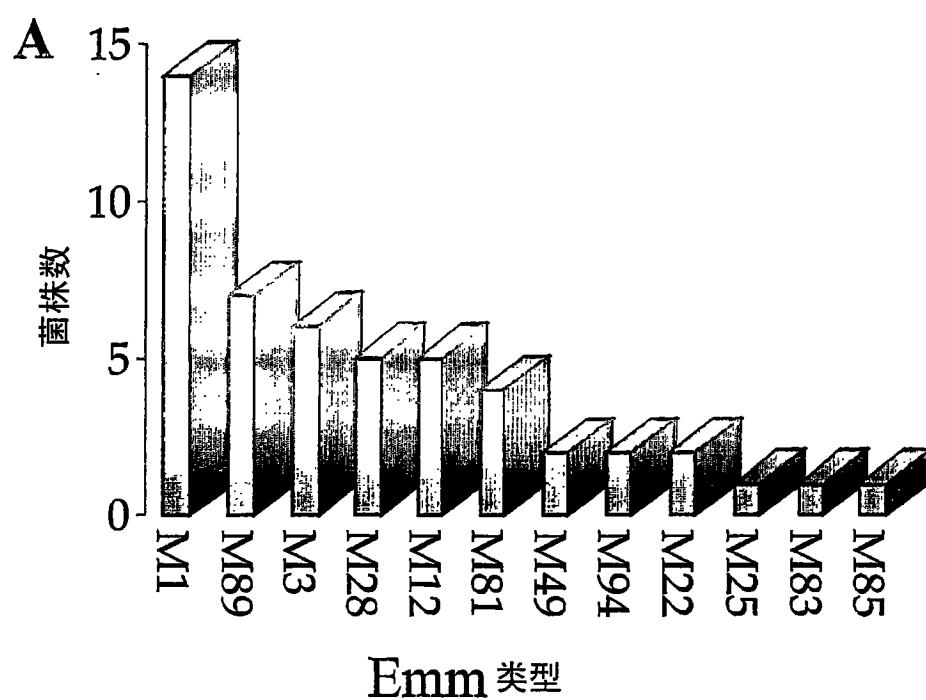


图 4

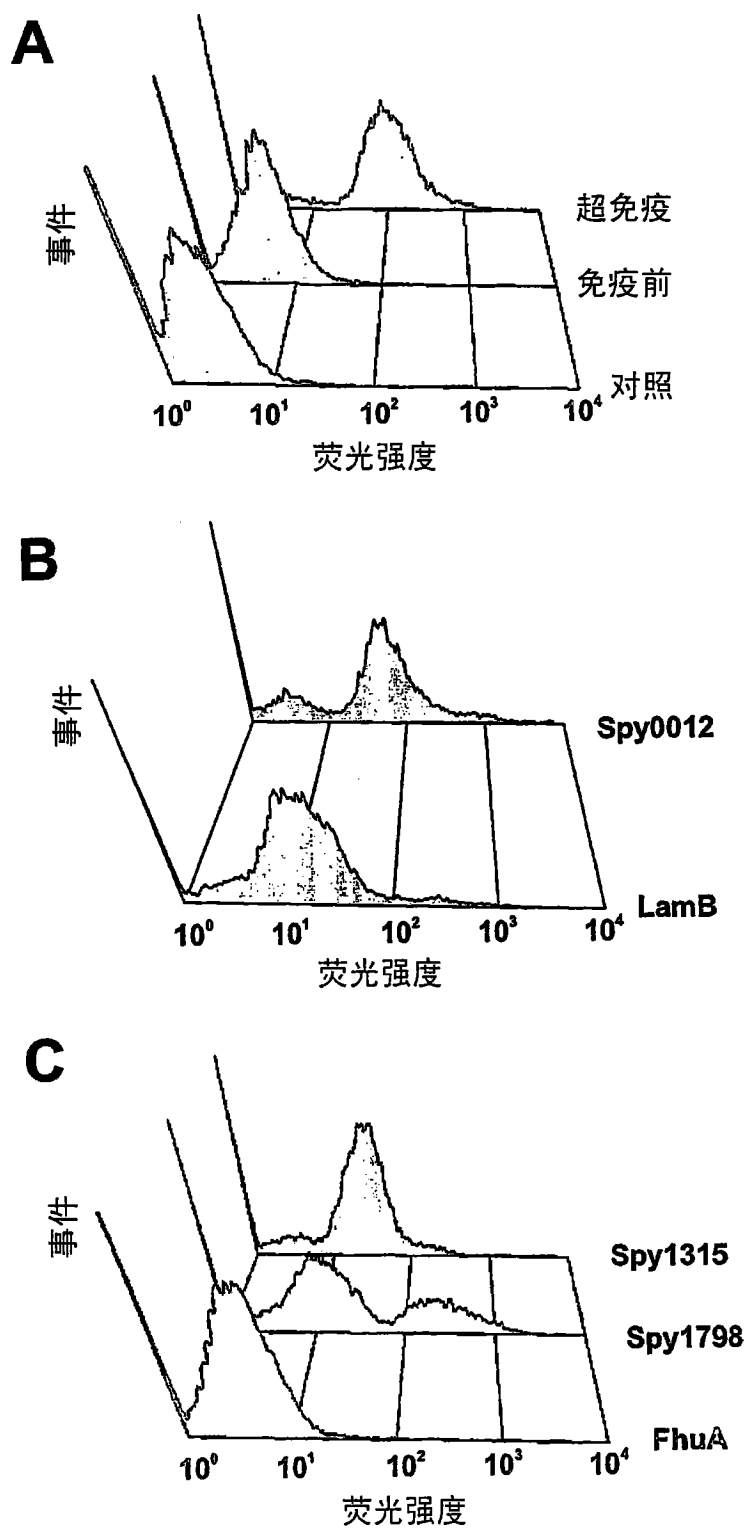


图 5

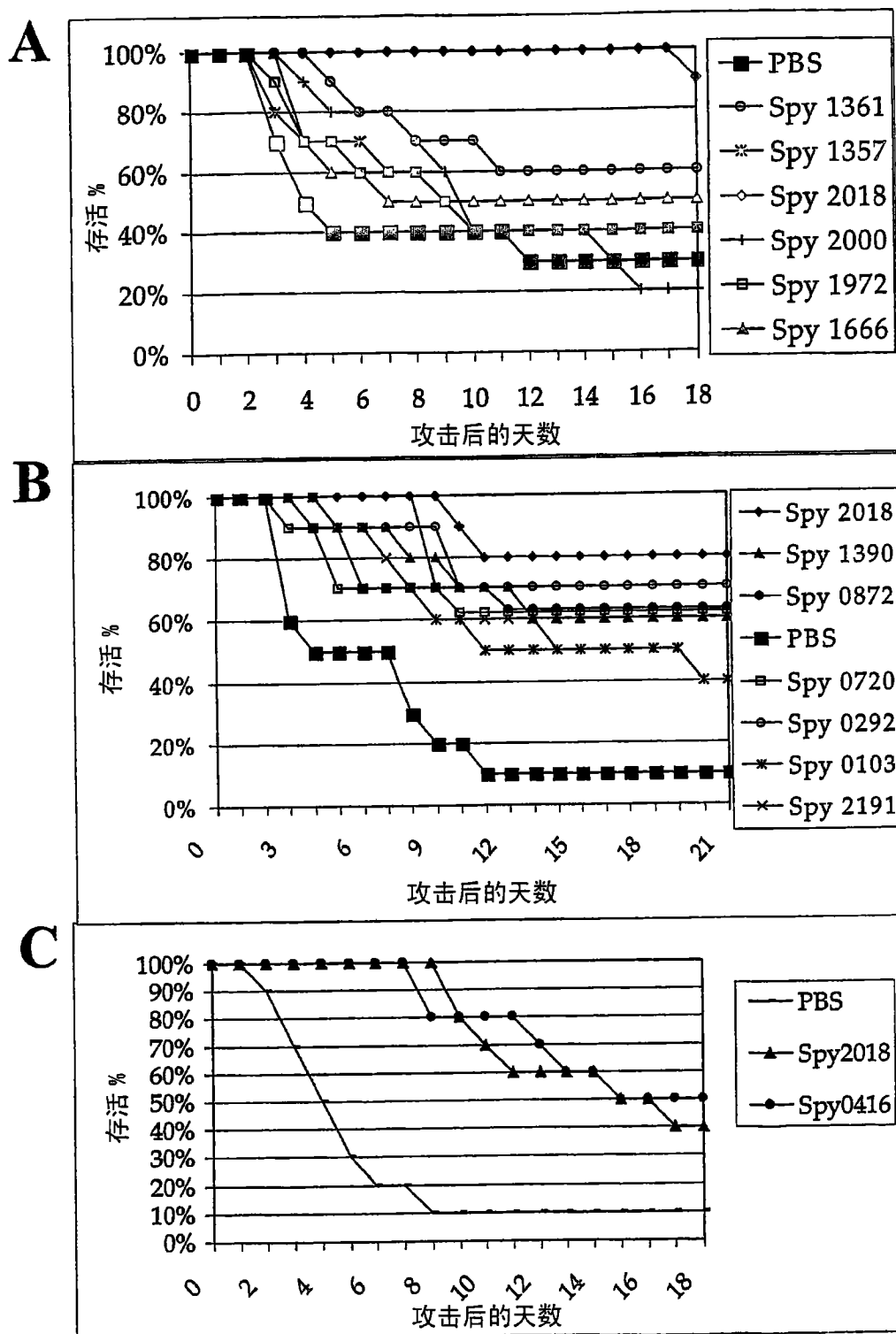


图 6