(19)国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 118510902 A
(43)申请公布日 2024.08.16

(21)申请号 202280084933.6
(22)申请日 2022.12.20
(30)优先权数据
63/292329 2021.12.21 US
(85)PCT国际申请进入国家阶段日
2024.06.21
(86)PCT国际申请的申请数据
PCT/US2022/081988 2022.12.20
(87)PCT国际申请的公布数据
WO2023/122567 EN 2023.06.29
(71)申请人 ABSCI公司
地址 美国华盛顿州
(72)发明人 J·刘
(74)专利代理机构 中国专利代理(香港)有限公司 72001
专利代理师 彭昶
(51)Int.Cl.
C12N 15/64(2006.01)
C12N 15/67(2006.01)
C12Q 1/00(2006.01)
G01N 33/00(2006.01)

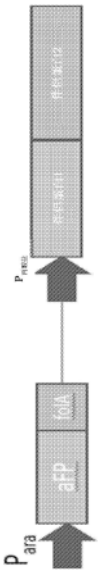
权利要求书3页 说明书26页
序列表（电子公布） 附图11页

(54)发明名称

用于鉴定具有增加的可溶性靶蛋白表达的菌株的FolA选择测定

(57)摘要

本公开提供了用于鉴定产生可溶性蛋白的细胞或菌株的组合物和方法。本公开进一步提供了一种用于在混合群体(文库)中鉴定产生较高可溶性蛋白的大肠杆菌菌株的生长/不生长选择策略。



CN 118510902 A

1.一种鉴定能够产生所关注的可溶性蛋白的宿主细胞的方法,所述方法包括以下步骤:

(a)制备宿主细胞群体,其中所述群体中的每个宿主细胞包括能够表达包括以下的融合蛋白的表达构建体:(i)所关注的蛋白,和(ii)选择蛋白;

(b)在允许表达所述融合蛋白的条件下温育(a)中的所述宿主细胞,其中所述条件包括生长物质,所述生长物质包括至少2种协同选择剂;以及

(c)使能够生长的宿主细胞可视化;

由此鉴定能够产生所关注的可溶性蛋白。

2.一种鉴定在宿主细胞群体中产生最高量的所关注的可溶性蛋白的宿主细胞的方法,所述方法包括以下步骤:

(a)制备宿主细胞群体,其中所述群体中的每个宿主细胞包括能够表达包括以下的融合蛋白的表达构建体:(i)所关注的蛋白,和(ii)选择蛋白;

(b)在允许表达所述融合蛋白的条件下温育(a)中的所述宿主细胞,其中所述条件包括生长物质,所述生长物质包括至少2种协同选择剂;以及

(c)使能够生长的宿主细胞可视化;

由此鉴定在宿主细胞群体中产生最高量的所关注的可溶性蛋白的宿主细胞。

3.根据权利要求1至2中任一项所述的方法,其进一步包括基于所产生的可溶性蛋白的量对所述宿主细胞进行分箱的步骤。

4.根据权利要求1至3中任一项所述的方法,其中所述选择蛋白是抗生素或抗生素抗性蛋白的靶标。

5.根据权利要求4所述的方法,其中所述选择蛋白是FolA。

6.根据权利要求5所述的方法,其中所述FolA来自大肠杆菌(E.coli)。

7.根据权利要求6所述的方法,其中所述FolA示出于SEQ ID NO:1中。

8.根据权利要求1至7中任一项所述的方法,其包括使用2种协同选择剂,其中所述协同选择剂选自由以下组成的组:抗生素、糖、化学剂、酶底物类似物、酶抑制剂、多价螯合(sequester)生物分子的药剂、螯合剂(chelating agent)和损害细胞壁或细胞膜的药剂。

9.根据权利要求8所述的方法,其中所述2种协同选择剂是甲氧苄啶(trimethoprim)和磺胺甲噁唑(sulfamethoxazole)。

10.根据权利要求1至9中任一项所述的方法,其中所述所关注的蛋白是异源蛋白。

11.根据权利要求10所述的方法,其中所述异源蛋白选自由以下组成的组:抗体、Fab、scFv、纳米抗体、T细胞受体和嵌合抗原受体、生长因子、细胞因子、激素、酶或其功能片段。

12.根据权利要求1至11中任一项所述的方法,其中所述宿主细胞群体包括宿主细胞文库,其中所述文库包含具有独特基因型和/或被独特基因工程化的细胞。

13.根据权利要求10所述的方法,其中所述文库包括大约一千个至大约十亿个宿主细胞。

14.根据权利要求1至13中任一项所述的方法,其中所述宿主细胞选自由以下组成的组:真核细胞、原核细胞、细菌细胞、哺乳动物细胞和昆虫细胞。

15.根据权利要求14所述的方法,其中所述细菌细胞是大肠杆菌细胞。

16.根据权利要求15所述的方法,其中所述大肠杆菌细胞包括:(a)编码用于至少一个

诱导型启动子的诱导物的转运蛋白的至少一个基因的基因功能的改变;(b)编码代谢至少一个诱导型启动子的诱导物的蛋白的至少一个基因的降低水平的基因功能;(c)编码参与至少一个诱导型启动子的诱导物的生物合成的蛋白的至少一个基因的降低水平的基因功能;(d)影响宿主细胞质的还原/氧化环境的基因的改变的基因功能;(e)编码还原酶的基因的降低水平的基因功能;(f)编码至少一种二硫键异构酶蛋白的至少一个表达构建体;(g)编码缺乏信号肽的DsbC的形式的至少一种多核苷酸;和/或(h)编码Ervlp的至少一种多核苷酸。

17.根据权利要求1至16中任一项所述的方法,其中所述表达构建体是选自由多核苷酸、质粒和人工染色体组成的组的染色体外构建体。

18.根据权利要求17所述的方法,其中所述表达构建体包括诱导型启动子。

19.根据权利要求17所述的方法,其中所述表达构建体包括两个或更多个诱导型启动子。

20.根据权利要求19所述的方法,其中至少一个诱导型启动子是丙酸盐诱导型启动子,并且至少一个其它诱导型启动子是L-阿拉伯糖诱导型启动子。

21.根据权利要求1至20中任一项所述的方法,其中生长底物选自由选择性培养基组成的组。

22.根据权利要求1至21中任一项所述的方法,其中所述生长底物包括具有2种或更多种协同选择剂的基质。

23.根据权利要求22所述的方法,其中所述2种协同选择剂是甲氧苄啶和磺胺甲噁唑,并且所述药剂分别以1ug/ml至1000ug/ml和1ug/ml至1000ug/ml的浓度范围存在于生长培养基中。

24.根据权利要求1至23中任一项所述的方法,其中所述允许表达所述融合蛋白的条件包含存在表达所述融合蛋白的一个或多个诱导物。

25.根据权利要求1至24中任一项所述的方法,其中所述宿主细胞是使用选自由以下组成的组的技术来鉴定的:目视检查、化学发光、射线照相、荧光和比色分析。

26.根据权利要求1至25中任一项所述的方法,其中步骤(c)中的所述使宿主细胞可视化包括检测琼脂板上的生长。

27.根据权利要求1至26中任一项所述的方法,其进一步包括以下步骤:

(a)将所述宿主细胞平板接种在生长底物上,并且在允许宿主细胞在所述生长底物上生长的条件下温育所述宿主细胞;

(b)任选地制备至少一个复制板,并且在允许宿主细胞生长和产生所述所关注的蛋白的条件下温育所述复制板;

(c)将宿主细胞从所述生长底物或(b)中的所述至少一个复制板转移到膜;

(d)制备已经在(c)中转移到所述膜的所述宿主细胞以进行探测,所述制备包括(i)在允许细胞组分固定化的条件下任选地固定所述宿主细胞;(ii)封闭所述宿主细胞;以及(iii)在允许所述宿主细胞透化的条件下任选地使所述宿主细胞裂解;

(e)在允许至少一个探针与所述所关注的蛋白结合的条件下使经透化的宿主细胞与包括所述至少一个探针的探针溶液接触,并且由此形成所关注的探针-蛋白复合物;以及

(f)在允许鉴定产生所述所关注的蛋白的所述宿主细胞的条件下对所述宿主细胞进行

成像。

用于鉴定具有增加的可溶性靶蛋白表达的菌株的FolA选择测定

[0001] 通过引用并入以电子方式提交的材料

[0002] 作为本公开的一部分的序列表与说明书作为文本文件同时提交。含有序列表的文本文件的名称为“57387_Seqlisting.txt”,所述文件于2022年11月15日创建,并且大小为1,893字节。序列表的主题通过引用整体并入本文。

背景技术

[0003] 抗生素抗性基因常用作用于选择测定的融合配偶体。常用的可选择融合配偶体是抗生素抗性基因。然而,这些方法受到有限的线性应答范围的影响,在所述有限的线性应答范围中,适量的抗生素抗性基因的表达赋予了实际可达到的抗生素浓度的几乎全部程度。旨在克服这一限制的方案采用了如翻译偶联的间接选择方案。到目前为止,尚未提供一种使抗生素抗性基因的表达与可溶性蛋白的表达之间直接相关的测定。本公开解决了此未满足的需求。

发明内容

[0004] 在一个实施例中,本公开提供了一种鉴定能够产生所关注的可溶性蛋白的宿主细胞的方法,所述方法包括以下步骤:(a)制备宿主细胞群体,其中所述群体中的每个宿主细胞包括能够表达包括以下的融合蛋白的表达构建体:(i)所关注的蛋白,和(ii)选择蛋白;(b)在允许表达所述融合蛋白的条件下温育(a)中的所述宿主细胞,其中所述条件包括生长物质,所述生长物质包括至少2种协同选择剂;以及(c)使能够生长的宿主细胞可视化;由此鉴定能够产生所关注的可溶性蛋白的宿主细胞。

[0005] 在另一实施例中,提供了一种鉴定在宿主细胞群体中产生最高量的所关注的可溶性蛋白的宿主细胞的方法,所述方法包括以下步骤:(a)制备宿主细胞群体,其中所述群体中的每个宿主细胞包括能够表达包括以下的融合蛋白的表达构建体:(i)所关注的蛋白,和(ii)选择蛋白;(b)在允许表达所述融合蛋白的条件下温育(a)中的所述宿主细胞,其中所述条件包括生长物质,所述生长物质包括至少2种协同选择剂;以及(c)使能够生长的宿主细胞可视化;由此鉴定在宿主细胞群体中产生最高量的所关注的可溶性蛋白的宿主细胞。

[0006] 在一些实施例中,提供了前述方法,其进一步包括基于所产生的可溶性蛋白的量对所述宿主细胞进行分箱的步骤。在一些实施例中,其中所述选择蛋白是抗生素或抗生素抗性蛋白的靶标。在一个实施例中,所述选择蛋白是FolA。在另一实施例中,所述FolA来自大肠杆菌(E.coli)。在仍另一实施例中,所述FolA示出于SEQ ID NO:1中。

[0007] 在仍其它实施例中,提供了前述方法,所述方法包括使用2种协同选择剂,其中所述协同选择剂选自由以下组成的组:抗生素、糖、化学剂、酶底物类似物、酶抑制剂、多价螯合(sequester)生物分子的药剂、螯合剂(chelating agent)和损害细胞壁或细胞膜的药剂。在一个实施例中,所述2种协同选择剂是甲氧苄啶(trimethoprim)和磺胺甲噁唑(sulfamethoxazole)。在一些实施例中,提供了前述方法,其中所关注的蛋白是异源蛋白。

在另一实施例中,所述异源蛋白选自由以下组成的组:抗体、Fab、scFv、纳米抗体、T细胞受体和嵌合抗原受体、生长因子、细胞因子、激素、酶或其功能片段。

[0008] 在一些实施例中,提供了前述方法,其中所述宿主细胞群体包括宿主细胞文库,其中所述文库包含具有独特基因型和/或被独特基因工程化的细胞。在一个实施例中,所述文库包括大约一千个至大约十亿个宿主细胞。

[0009] 在一些实施例中,提供了前述方法,其中所述宿主细胞选自由以下组成的组:真核细胞、原核细胞、细菌细胞、哺乳动物细胞和昆虫细胞。在另一实施例中,所述细菌细胞是大肠杆菌细胞。在仍另一实施例中,所述大肠杆菌细胞包括:(a)编码用于至少一个诱导型启动子的诱导物的转运蛋白的至少一个基因的基因功能的改变;(b)编码代谢至少一个诱导型启动子的诱导物的蛋白的至少一个基因的降低水平的基因功能;(c)编码参与至少一个诱导型启动子的诱导物的生物合成的蛋白的至少一个基因的降低水平的基因功能;(d)影响宿主细胞质的还原/氧化环境的基因的改变的基因功能;(e)编码还原酶的基因的降低水平的基因功能;(f)编码至少一种二硫键异构酶蛋白的至少一个表达构建体;(g)编码缺乏信号肽的DsbC的形式的至少一种多核苷酸;和/或(h)编码Ervlp的至少一种多核苷酸。

[0010] 在仍其它实施例中,提供了前述方法,其中所述表达构建体是选自由多核苷酸、质粒和人工染色体组成的组的染色体外构建体。在一个实施例中,所述表达构建体包括诱导型启动子。在另一实施例中,所述表达构建体包括两个或更多个诱导型启动子。在仍另一实施例中,至少一个诱导型启动子是丙酸盐诱导型启动子,并且至少一个其它诱导型启动子是L-阿拉伯糖诱导型启动子。

[0011] 在又其它实施例中,提供了前述方法,其中生长底物选自由选择性培养基组成的组。在一些实施例中,提供了前述方法,其中所述生长底物包括具有2种或更多种协同选择剂的基质。在一个实施例中,所述2种协同选择剂是甲氧苄啶和磺胺甲噁唑,并且所述药剂分别以1ug/ml至1000ug/ml和1ug/ml至1000ug/ml的浓度范围存在于生长培养基中。

[0012] 在一些实施例中,提供了前述方法,其中所述允许表达所述融合蛋白的条件包含存在表达所述融合蛋白的一种或多种诱导物。在一些实施例中,提供了前述方法,其中所述宿主细胞是使用选自由以下组成的组的技术来鉴定的:目视检查、化学发光、射线照相、荧光和比色分析。在仍另一实施例中,提供了前述方法,其中步骤(c)中的所述使宿主细胞可视化包括检测琼脂板上的生长。

[0013] 在一些实施例中,本公开提供了前述方法,所述方法进一步包括检测可溶性蛋白表达的固相测定。例如,在一个实施例中,提供了前述方法,所述方法进一步包括以下步骤:(a)将所述宿主细胞平板接种在生长底物上,并且在允许宿主细胞在所述生长底物上生长的条件下温育所述宿主细胞;(b)任选地制备至少一个复制板,并且在允许宿主细胞生长和产生所述所关注的蛋白的条件下温育所述复制板;(c)将宿主细胞从所述生长底物或(b)中的所述至少一个复制板转移到膜;(d)制备已经在(c)中转移到所述膜的所述宿主细胞以进行探测,所述制备包括(i)在允许细胞组分的固定化的条件下任选地固定所述宿主细胞;(ii)封闭所述宿主细胞;以及(iii)在允许所述宿主细胞透化的条件下任选地使所述宿主细胞裂解;(e)在允许至少一个探针与所述所关注的蛋白结合的条件下使经透化的宿主细胞与包括所述至少一个探针的探针溶液接触,并且由此形成所关注的探针-蛋白复合物;以及(f)在允许鉴定产生所述所关注的蛋白的所述宿主细胞的条件下对所述宿主细胞进行成

像。

[0014] 下文还提供了另外的实施例,其中所关注的蛋白由所关注的基因产物编码。在一个实施例中,本公开提供了一种用于鉴定产生所关注的基因产物的宿主细胞的方法,所述方法包括以下步骤:(a)将宿主细胞群体平板接种在生长底物上,其中所述宿主细胞包括编码一个或多个基因产物的表达构建体,并且在允许宿主细胞在所述生长底物上生长的条件下温育所述宿主细胞;(b)制备至少一个复制板,并且在允许宿主细胞生长和产生所述基因产物的条件下温育所述复制板;(c)将宿主细胞从(b)中的所述至少一个复制板转移到膜;(d)制备已经在(c)中转移到所述膜的所述宿主细胞以进行探测,所述制备包括(i)在允许细胞组分的固定化的条件下任选地固定所述宿主细胞;(ii)阻断所述宿主细胞;以及(iii)在允许所述宿主细胞透化的条件下任选地使所述宿主细胞裂解;(e)在允许至少一个探针与所关注的基因产物结合的条件下使经透化的宿主细胞与包括所述至少一个探针的探针溶液接触,并且由此形成所关注的探针-基因产物复合物;以及(f)在允许鉴定产生所述所关注的基因产物的所述宿主细胞的条件下对所述宿主细胞进行成像。在另一实施例中,步骤(c)进一步包括标记所述至少一个复制板和膜,以允许所述至少一个复制板和膜在空间上对准。

[0015] 在仍其它实施例中,提供了前述方法,其中所述宿主细胞选自由以下组成的组:真核细胞、原核细胞、细菌细胞、哺乳动物细胞和昆虫细胞。在仍其它实施例中,提供了前述方法,其中所述所关注的基因产物选自由以下组成的组:治疗性蛋白、抗体、Fab、scFv、纳米抗体、T细胞受体和嵌合抗原受体或其片段。在一个实施例中,所述所关注的基因产物是细胞内蛋白。

[0016] 在又其它实施例中,提供了前述方法,其中所述宿主细胞群体包括大约一千个至大约十亿个宿主细胞。在其它实施例中,所述宿主细胞群体已经修饰以由表达构建体产生基因产物。在一个实施例中,所述表达构建体包括两个或更多个诱导型启动子。在仍另一实施例中,至少一个诱导型启动子是丙酸盐诱导型启动子,并且至少一个其它诱导型启动子是L-阿拉伯糖诱导型启动子。

[0017] 在一些实施例中,本公开还提供了前述方法,其中所述宿主细胞已经基因修饰以包括以下一种或多种:(a)编码用于至少一个诱导型启动子的诱导物的转运蛋白的至少一个基因的基因功能的改变;(b)编码代谢至少一个诱导型启动子的诱导物的蛋白的至少一个基因的降低水平的基因功能;(c)编码参与至少一个诱导型启动子的诱导物的生物合成的蛋白的至少一个基因的降低水平的基因功能;(d)影响宿主细胞质的还原/氧化环境的基因的改变的基因功能;(e)编码还原酶的基因的降低水平的基因功能;(f)编码至少一种二硫键异构酶蛋白的至少一个表达构建体;(g)编码缺乏信号肽的DsbC的形式的至少一种多核苷酸;和/或(h)编码Erv1p的至少一种多核苷酸。

[0018] 在仍其它实施例中,提供了前述方法,其中制备了2个、3个、4个、5个或更多个复制板。在一个实施例中,复制板含有抗生素和使所关注的基因产物表达的至少一种诱导物。在仍其它实施例中,所述膜选自由硝化纤维素膜和PVDF(聚偏二氟乙烯)膜组成的组。在其它实施例中,步骤(c)中的复制板和膜与有色物质接触。在一个实施例中,有色物质选自由丙烯酸颜料和染料组成的组。在另一实施例中,用选自由针和笔组成的组的仪器使有色物质与所述板和膜接触。

[0019] 在其它实施例中,提供了前述方法,其中允许细胞组分的固定化的条件包括使所述膜与包括戊二醛和多聚甲醛中的一种或多种的组合物接触。在其它实施例中,允许宿主细胞透化的条件包括使所述膜与包括溶菌酶和EDTA中的一种或多种的组合物接触。

[0020] 在仍其它实施例中,提供了前述方法,其中所述至少一个探针选自由以下组成的组:抗体或其功能片段、核酸结合蛋白或其功能片段、受体或其功能片段、配体或其功能片段、和抗原或其功能片段以及能够被所关注的基因产物结合的肽或多肽。在一个实施例中,探针包括选自由以下组成的组的报告基因部分:生物素、组氨酸标签、Fc标签、spy标签、Strp标签和Avi标签。

[0021] 在一些实施例中,本公开还提供前述方法,其中当已经制备了2个、3个、4个、5个或更多个复制板时,使由复制板制备的每张膜与不同浓度的探针溶液接触。在仍其它实施例中,提供了前述方法,其中成像包括使所述膜与包括报告基因部分的激活剂的组合物接触。在另一实施例中,所述使所述膜与包括报告基因部分的激活剂的组合物接触重复一次、两次、三次或更多次。在仍另一实施例中,激活剂选自由以下组成的组:碱性磷酸酶、链霉亲和素碱性磷酸酶和链霉亲和素碱性磷酸酶葡聚糖聚合物。

[0022] 在仍其它实施例中,提供了前述方法,其中所述成像包括选自由以下组成的组的方法:化学发光、射线照相、荧光和比色分析。在仍其它实施例中,提供了前述方法,所述方法进一步包括将一个或多个已经被鉴定为能够产生所关注的基因产物的宿主细胞重新平板接种的步骤。在另一实施例中,使重新平板接种的宿主细胞经受根据权利要求1所述的方法以确认宿主细胞产生所关注的基因产物的能力和/或以将宿主细胞菌株与其它被鉴定为能够产生所关注的基因产物的宿主细胞菌株分离。在一些实施例中,所述方法可以进一步包括确定探针-基因产物复合物的亲和力。

[0023] 在一个实施例中,本公开还提供了一种用于从宿主细胞群体中筛选产生所关注的基因产物的宿主细胞的方法,所述方法包括以下步骤:(a)将宿主细胞群体平板接种在生长底物上,其中所述宿主细胞包括编码一个或多个基因产物的表达构建体,并且在允许宿主细胞在所述生长底物上生长的条件下温育所述宿主细胞;(b)制备至少一个复制板,并且在允许宿主细胞生长和产生所述基因产物的条件下温育所述复制板;(c)将宿主细胞从(b)中的所述至少一个复制板转移到膜;(d)制备已经在(c)中转移到所述膜的所述宿主细胞以进行探测,所述制备包括(i)在允许细胞组分的固定化的条件下固定所述宿主细胞;(ii)阻断所述宿主细胞;以及(iii)在允许所述宿主细胞透化的条件下任选地使所述宿主细胞裂解;(e)在允许至少一个探针与所关注的基因产物结合的条件下使经透化的宿主细胞与包括所述至少一个探针的探针溶液接触,并且由此形成所关注的探针-基因产物复合物;以及(f)在允许鉴定产生所述所关注的基因产物的所述宿主细胞的条件下对所述宿主细胞进行成像。

[0024] 在仍另一实施例中,本公开提供了一种确定所关注的探针-基因产物的相对亲和力的方法,所述方法包括以下步骤:(a)将宿主细胞群体平板接种在生长底物上,其中所述宿主细胞包括编码一个或多个基因产物的表达构建体,并且在允许宿主细胞在所述生长底物上生长的条件下温育所述宿主细胞;(b)制备至少三个复制板,并且在允许宿主细胞生长和产生所述基因产物的条件下温育所述复制板;(c)将宿主细胞从(b)中的所述至少三个复制板转移以分离膜;(d)制备已经在(c)中转移到所述膜的所述宿主细胞以进行探测,所述

制备包括(i)在允许细胞组分的固定化的条件下固定所述宿主细胞;(ii)阻断所述宿主细胞;以及(iii)在允许所述宿主细胞透化的条件下任选地使所述宿主细胞裂解;(e)在允许至少一个探针与所述所关注的基因产物结合的条件下使所述至少三个复制膜上的经透化的宿主细胞与包括所述至少一个探针的探针溶液接触,并且由此形成探针-基因产物复合物,并且其中所述探针溶液包括不同的探针浓度;以及(f)在允许确定所关注的探针-基因产物的相对亲和力的条件下对所述宿主细胞进行成像。

附图说明

[0025] 图1示出了携带mCherry-FolA融合物的细胞对甲氧苄啶(trimethoprim)具有很强的抗性。

[0026] 图2示出了与磺胺甲噁唑的协同作用增加了甲氧苄啶动态范围。

[0027] 图3示出了甲氧苄啶-磺胺甲噁唑的基质。

[0028] 图4示出了示例性质粒构建体。

[0029] 图5示出了在温育1周之后具有菌落的板。

[0030] 图6示出了来自固相测定的1轮的结果的图像。

[0031] 图7示出了2轮的结果。

[0032] 图8示出了经处理的膜的图像。

[0033] 图9示出了示例性质粒构建体。

[0034] 图10示出了固相测定的另外的结果。

[0035] 图11示出了与琼脂板相比,液体执行器的结果。

具体实施方式

[0036] 本公开的实施例提供了用于从细胞或菌株文库中鉴定能够产生大量所关注的可溶性蛋白的细胞或菌株的组合物和方法。在一些实施例中,所关注的蛋白与选择蛋白(如抗生素抗性基因或作为抗生素的靶标的蛋白)以基因方式融合。如本文所描述的,在一些实施例中,选择过程包含使用协同作用的2种(或更多种)选择剂以增加测定的动态范围。

[0037] 融合配偶体(即选择蛋白)能够进行选择的考虑因素包含但不限于:(1)能够进行生长/不生长选择;(2)能够通过修改选择条件从低到高鉴定表达菌株的分级;以及(3)清除(低背景噪声)。可选测定应允许在单个10cm的琼脂板上在数十亿变体中测试大量不同的文库。

[0038] 如本文所描述的,每个10cm板可以包括至多100亿个细胞(例如,100亿用于杀生物(-cidal)药剂组合,10亿用于生物抑制(-static)药剂组合)。

[0039] 鉴定能够产生所关注的可溶性蛋白的宿主细胞

[0040] 在本公开的一个实施例中,提供了一种生长/不生长选择策略,以鉴定混合群体(文库)中产生较高可溶性蛋白表达(即相对于文库中的其它菌株的表达)的大肠杆菌菌株。如本文所描述的,选择剂的协同作用增加了测定的动态范围。通过举例的方式,所关注的蛋白被基因工程化,并且与作为抗生素(甲氧苄啶)的靶标的报告基因(folA)融合。在这一示例性实施例中,报告基因(folA)/抗生素(甲氧苄啶)配对的选择允许与第二抗生素剂协同作用,以提高主要抗生素(甲氧苄啶)的效力。因此,成功的生长/不生长选择策略的益处允

许并行测试若干个高复杂度(>10^9)的文库。

[0041] 本文所描述的融合蛋白可以与任何抗生素靶标或抗生素抗性基因进行融合,只要第二协同剂可以用于增加测定的有用动态范围即可。

[0042] 融合蛋白

[0043] 所关注的蛋白

[0044] 如本文所描述的,所关注的蛋白(以及在本文所描述的一些实施例中,编码所关注的蛋白的所关注的基因)可以与本文所提供的方法一起使用。如本文所使用的,在一个实施例中,“可溶性蛋白”或“所关注的可溶性蛋白”是指当所施加的离心力(16,000x g)≥10分钟时将在溶液中保持稳定并且不会随时间的推移而沉降,或者保持处于溶液中的蛋白质。在一些实施例中,蛋白溶解度是热力学参数,所述热力学参数被定义为在给定的一组条件下与固相(结晶相或无定形相)平衡的饱和溶液中的蛋白质的浓度。溶解度可能受到许多包含pH、离子强度、温度和各种溶剂添加剂的存在(Kramer,R.M.等人,《生物物理学杂志(Biphys J.)》,2012,102(8):1907-1915)的外在和内在因素的影响。在一个实施例中,可溶性蛋白是指溶解度大于70%的可溶性蛋白并且不可溶性蛋白是指溶解度小于30%的不可溶性蛋白(Chan,P.等人,《科学报告(Scientific Reports)》,2013,3,3333)。

[0045] 蛋白质可以包含生物活性衍生物或变体或片段。如本文所使用的,“生物活性衍生物”或“生物活性变体”包含分子的任何衍生物或变体,所述衍生物或变体与所述分子具有基本上相同的功能和/或生物特性,如结合特性,和/或相同的结构基础,如肽主链或基本聚合单元。

[0046] “类似物”,如“变体”或“衍生物”是与天然存在的分子在结构上基本上类似并且具有相同生物活性的化合物,尽管在某些情况下具有不同程度的生物活性。例如,多肽变体是指与参考多肽共享基本上类似的结构并且具有相同生物活性的多肽。基于一个或多个突变,变体或类似物的氨基酸序列的组成与源自所述类似物的天然存在的多肽的氨基酸序列的组成不同,所述突变涉及(i)多肽的一个或多个末端处和/或天然存在的多肽序列(例如,片段)的一个或多个内部区缺失一个或多个氨基酸残基;(ii)多肽的一个或多个末端处(通常为“添加”或“融合”)和/或天然存在的多肽序列的一个或多个内部区(通常为“插入”)的一个或多个氨基酸的插入或添加;或(iii)天然存在的多肽序列中的一个或多个氨基酸对其它氨基酸的取代。通过举例的方式,“衍生物”是一种类型的类似物,并且是指与已经修饰的(例如,化学地修饰的)参考多肽共享相同或基本上类似的结构的多肽。

[0047] 变体多肽是一种类型的类似多肽,并且包含插入变体,其中一个或多个氨基酸残基被添加到本公开的生物分子氨基酸序列中。插入可以位于蛋白质的任一或两个末端处,和/或可以定位于治疗性蛋白氨基酸序列的内部区内。在任一或两个末端处具有另外的残基的插入变体包含例如融合蛋白和含有氨基酸标签或其它氨基酸标记的蛋白。一方面,生物分子任选地含有N端Met,特别是当分子在细菌细胞,如大肠杆菌中重组表达时。另一方面,生物分子包含组氨酸标签(His-tag)。

[0048] 在缺失变体中,如本文所描述的生物分子多肽中的一个或多个氨基酸残基被去除。缺失可以在蛋白多肽的一个或两个末端处实现,和/或缺失可以通过去除治疗性蛋白氨基酸序列内的一个或多个残基实现。因此,缺失变体包含蛋白多肽序列的片段。

[0049] 在取代变体中,生物分子的一个或多个氨基酸残基被去除,并且被替代性残基替

代。一方面,取代本质上是保守的,并且这种类型的保守取代在本领域是熟知的。可替代地,本公开涵盖非保守的取代。示例性保守取代在Lehninger,[《生物化学(Biochemistry)》,第2版;纽约沃思出版社有限公司(Worth Publishers,Inc.,New York)(1975),第71-77页]]中进行了描述,并且立即在下文中列出。

[0050] 本文所设想的蛋白质包含全长蛋白、全长蛋白的前体、全长蛋白的生物活性亚基或片段以及任何这些形式的治疗性蛋白中的任一种的生物活性衍生物和变体。因此,蛋白质包含(1)在至少约25个、约50个、约100个、约200个、约300个、约400个或更多个氨基酸的区内,与本文所描述的由参考核酸或氨基酸序列编码的多肽具有大于约60%、约65%、约70%、约75%、约80%、约85%、约90%、约91%、约92%、约93%、约94%、约95%、约96%、约97%、约98%或约99%或更高氨基酸序列同一性的氨基酸序列的蛋白质。根据本公开,术语“重组蛋白”包含通过重组DNA技术获得的任何蛋白。在某些实施例中,所述术语涵盖本文所描述的蛋白质。

[0051] 在一些实施例中,蛋白质是治疗性蛋白,如单克隆或多克隆抗体或其它糖蛋白、生物仿制药、Fc-融合物、酶、疫苗、激素、细胞因子或生长因子。在仍其它实施例中,基因产物是抗凝血剂、血液因子、骨形态发生蛋白、白介素、干扰素、溶栓剂或通过重组方式产生的任何蛋白。在其它实施例中,这些方法提供了所述方法可以与包含小分子的任何生物分子或化学实体一起使用,所小分子由本文所描述的细胞产生,并且具有可以用于结合步骤的探针。

[0052] 如本文所使用的,术语“抗体”是指与靶抗原上的表位相互作用(例如,通过结合、空间位阻、稳定/不稳定、空间分布)的完整抗体。天然存在的“抗体”是包括通过二硫键互连的至少两条重(H)链和两条轻(L)链的糖蛋白。每条重链包含重链可变区(在本文中缩写为VH)和重链恒定区。重链恒定区包含三个结构域CH1、CH2和CH3。每条轻链包含轻链可变区(在本文中缩写为VL)和轻链恒定区。轻链恒定区包含一个结构域CL。VH区和VL区可以进一步细分为被称作互补决定区(CDR)的超变区,所述超变区散布有更保守的被称作框架区(FR)的区。每个VH和VL由按照以下顺序从氨基末端到羧基末端布置的三个CDR和四个FR构成:FR1、CDR1、FR2、CDR2、FR3、CDR3、FR4。重链和轻链的可变区含有与抗原相互作用的结合结构域。抗体的恒定区可以介导免疫球蛋白与宿主组织或因子的结合,包含免疫系统的不同细胞(例如,效应细胞)以及经典补体系统的第一组分(C1q)。术语“抗体”包含例如单克隆抗体、人抗体、人源化抗体、骆驼化抗体、嵌合抗体、单链Fv(scFv)、二硫键连接的Fv(sdFv)、Fab片段、F(ab')片段和抗独特型(抗Id)抗体(包含例如针对本发明的抗体的抗Id抗体)以及上述中的任一种的表位结合片段。抗体可以是任何同种型(例如,IgG、IgE、IgM、IgD、IgA和IgY)、种类(例如,IgG1、IgG2、IgG3、IgG4、IgA1和IgA2)或亚类。抗体或表位结合片段可以是包含双特异性抗体的多特异性分子,或者是所述多特异性分子的组分。

[0053] 轻链和重链两者分为结构同源性区和功能同源性区。术语“恒定”和“可变”在功能上使用。在这方面,应当了解,轻(VL)链部分和重(VH)链部分两者的可变结构域确定抗原识别和特异性。相比之下,轻链(CL)和重链(CH1、CH2或CH3)的恒定结构域赋予重要的生物学特性,如分泌、经胎盘移动性、Fc受体结合、补体结合等。按照惯例,恒定区结构域的编号随着所述恒定区结构域变得更加远离抗体的抗原结合位点或氨基末端而增加。N端是可变区,并且C端是恒定区;CH3和CL结构域实际分别包括重链和轻链的羧基末端。

[0054] 如本文所使用的,短语“抗体片段”是指抗体的保留与靶表位特异性地相互作用(例如,通过结合、空间位阻、稳定/去稳定、空间分布)的能力的一个或多个部分。结合片段的实例包含但不限于Fab片段,由VL、VH、CL和CH1结构域组成的单价片段;F(ab')2片段,包括在铰链区处通过二硫键连接的两个Fab片段的二价片段;由VH和CH1结构域组成的Fd片段;由抗体的单臂的VL和VH结构域组成的Fv片段;由VH结构域组成的dAb片段(Ward等人,(1989)《自然(Nature)》341:544-546);以及分离的互补决定区(CDR)。此外,虽然Fv片段的两个结构域VL和VH由单独的基因编码,但是所述两个结构域可以使用重组方法通过使所述两个结构域能够成为单个蛋白链的合成接头来连接,在所述单个蛋白链中,VL区和VH区配对以形成单价分子(被称为单链Fv(scFv));参见例如Bird等人,(1988),《科学(Science)》242:423-426);以及Huston等人,(1988)《美国国家科学院院刊(Proc.Natl.Acad.Sci.)》85:5879-5883)。此类单链抗体也旨在涵盖在术语“抗体片段”内。这些抗体片段使用本领域的技术人员已知的常规技术获得,并且以与完整抗体相同的方式针对实用性对所述片段进行筛选。

[0055] 在其它实施例中,所关注的蛋白是治疗性蛋白、T细胞受体和嵌合抗原受体、生长因子、细胞因子、激素、酶或其功能片段。在一个实施例中,蛋白质是胰岛素。

[0056] 在一些实施例中,本公开还提供了一种融合蛋白,所述融合蛋白包含两个蛋白组分之间的任选的接头,例如,柔性接头序列。在一些实施例中,接头的长度为至少5个、至少10个、至少15个、至少20个、至少25个、至少30个、至少35个、至少40个、至少45个或至少50个氨基酸。在其它实施例中,接头的长度为少于5个、少于10个、少于15个、少于20个、少于25个、少于30个、少于35个、少于40个、少于45个或少于50个氨基酸。在一个实施例中,接头的长度是23个氨基酸。在各个实施例中,接头可以是柔性的、刚性的或可切割的(例如,二硫化物、蛋白酶敏感序列)(Chen,X.等人,《先进药物递送综述(Advanced Drug Delivery Reviews)》,65(1):1357-1369(2013))。接头可以源自天然存在的多结构域蛋白,或者可以是经验的(参见Argos,P.,《分子生物学杂志(J.Mol.Biol.)》,211:943-958(1990);以及Heringa,G.R.,《蛋白质工程化(Protein Eng.)》,15:871-879(2002),所述文献独立地比较天然接头的若干特性,如长度、疏水性、氨基酸残基和次级结构。在一些实施例中,苏氨酸(Thr)、丝氨酸(Ser)、脯氨酸(Pro)、甘氨酸(Gly)、天冬氨酸(Asp)、赖氨酸(Lys)、谷氨酰胺(Gln)、天冬酰胺(Asn)和丙氨酸(Ala)与精氨酸(Arg)、苯丙氨酸(Phe)和谷氨酸(Glu)都是优选的接头组分。一般来说,优选的氨基酸是极性不带电荷或带电荷残基。天然接头采用各种次级结构,如螺旋、β链、线圈/弯曲和转折,以发挥其功能。

[0057] 选择蛋白和选择剂

[0058] 在一些实施例中,本文所提供的选择蛋白是蛋白质、多肽或如上文所描述的变体或衍生物的类似物。任何选择剂可以用于本公开提供的方法中,只要药剂能协同工作以增加动态范围。“协同选择剂”是指两种或更多种药剂的任何组合,其中药剂组合的效果大于各个药剂单独的效果之和。协同选择剂包含但不限于抗生素、糖、化学剂、酶底物和类似物、酶抑制剂、多价螯合生物分子的药剂、螯合剂和损害细胞壁或细胞膜的药剂。选择剂可以是构成生长底物的组合物的一部分,或者可以单独提供。

[0059] 在一个实例中,选择蛋白是二氢叶酸还原酶(dhfr),并且协同选择剂是甲氧苄啶和磺胺甲噁唑。在某些实施例中,dhfr来自细菌来源。通常,细菌dhfr被指定为FolA。在优选

的实施例中,FolA源自大肠杆菌。在一个实施例中,在根据本文所描述的方法使用的细胞或菌株中,内源性FolA被敲除或以其它方式失活。在一个实施例中,FolA如下源自大肠杆菌(www.uniprot.org/uniprot/P0ABQ4.fasta):

[0060]

SEQ ID NO: 1 FolA	MISLIAALAVDRVIGMENAMPWNLPADLAWFKRNTLNKPVIMGRHT WESIGRPLPGRKNIILSSQPGTDDRVTWVKSVDEAIAACGDVPEIMVI GGGRVYEQFLPKAQKLYLTHIDAEVEGDTHFPDYEPDDWESVFSEFH DADAQNSHSYCFEILERR

[0061] 在其它实施例中,选择蛋白和协同选择剂包含但不限于甲氧苄啶和磺胺甲噁唑、b-内酰胺/b-内酰胺酶抑制剂、细胞壁活性剂/氨基糖苷类和磷霉素/b-内酰胺的协同组合。

[0062] 细胞

[0063] 本公开设想了细胞,包含来自细胞群体的细胞或作为细胞文库的一部分的细胞。如本文所使用的,关于细胞的"文库"是指有意工程化或天然存在的细胞的基因不同的变体的任何集合。例如,设想了具有独特基因型和/或被独特基因工程化的细胞。本文描述了示例性基因修饰,包含允许细胞表达和产生所关注的重组异源蛋白的修饰。在一些实施例中,独特基因型或经独特工程化的细胞是指在基因方面彼此不同的细胞群体,例如存在突变、质粒、质粒中的基因等。

[0064] 在本公开的各个实施例中,设想了包括本文所描述的表达构建体中的一种或多种的细胞。本公开的细胞包含外膜(例如,包含蛋白质和脂质),在所述外膜中,本文所描述的融合蛋白一旦被表达,就可以相互作用或以其它方式驻留。

[0065] 原核宿主细胞。在本公开的一些实施例中,在宿主细胞,如原核宿主细胞中提供了被设计成用于基因产物(包含本文所描述的融合蛋白)的表达的表达构建体。原核宿主细胞可以包含古菌(archaea)(如盐富饶菌(Haloferax volcanii)、硫磺矿硫化叶菌(Sulfolobus solfataricus))、革兰氏阳性菌(如枯草芽孢杆菌(Bacillus subtilis)、地衣芽孢杆菌(Bacillus licheniformis)、桥石短芽孢杆菌(Brevibacillus choshinensis)、短乳杆菌(Lactobacillus brevis)、布氏乳杆菌(Lactobacillus buchneri)、乳酸乳球菌(Lactococcus lactis)和变铅青链霉菌(Streptomyces lividans))或革兰氏阴性菌细菌,包含变形杆菌纲(Alphaproteobacteria)(根瘤农杆菌(Agrobacterium tumefaciens)、新月柄杆菌(Caulobacter crescentus)、类球红杆菌(Rhodobacter sphaeroides)和苜蓿中华根瘤菌(Sinorhizobium meliloti))、β-变形菌(Betaproteobacteria)(真养产碱杆菌(Alcaligenes eutrophus)和丙型变形细菌纲(Gammaproteobacteria)(乙酸钙不动杆菌(Acinetobacter calcoaceticus)、固氮菌(Azotobacter)、棕色固氮菌(vinelandii)、大肠杆菌、铜绿假单胞菌(Pseudomonas aeruginosa)和恶臭假单胞菌(Pseudomonas putida))。优选的宿主细胞包含肠杆菌科的γ变形菌纲,如肠杆菌属(Enterobacter)、欧文氏菌属(Erwinia)、埃希氏杆菌属(Escherichia)(包含大肠杆菌)、克雷伯氏菌属(Klebsiella)、变形杆菌属(Proteus)、沙门氏菌属(Salmonella)(包含鼠伤寒沙门氏菌(Salmonella typhimurium))、沙雷氏菌属(Serratia)(包含粘质沙雷氏菌(Serratia marcescans))和志贺氏杆菌属(Shigella)。

[0066] 真核宿主细胞。许多另外类型的宿主细胞可以用于本公开所提供的表达系统,包含真核细胞,如酵母(休哈塔假丝酵母(Candida shehatae)、乳酸克鲁维酵母

(Kluyveromyces lactis)、脆壁克鲁维酵母(Kluyveromyces fragilis)、其它克鲁维酵母菌属物种(Kluyveromyces species)、毕赤酵母(Pichia pastoris)、酿酒酵母(Saccharomyces cerevisiae)、巴氏酵母(Saccharomyces pastorianus),也称为卡尔斯伯酵母(Saccharomyces carlsbergensis)、粟酒裂殖酵母(Schizosaccharomyces pombe)、德克酵母属/酒香酵母属(Dekkera/Brettanomyces)和解脂耶氏酵母(Yarrowia lipolytica));其它真菌(构巢曲霉(Aspergillus nidulans)、黑曲霉(Aspergillus niger)、粗糙脉孢菌(Neurospora crassa)、青霉菌(Penicillium)、弯颈霉属(Tolypocladium)、木霉属(Trichoderma reesia));昆虫细胞系(黑腹果蝇施耐德(Drosophila melanogaster Schneider)2细胞和草地贪夜蛾(Spodoptera frugiperda)Sf9细胞);以及哺乳动物细胞系,包含无限增殖化细胞系(中国仓鼠卵巢(CHO)细胞、HeLa细胞、幼仓鼠肾(BHK)细胞、猴肾细胞(COS)、人胚肾(HEK、293或HEK-293)细胞)和人肝癌细胞(Hep G2))。上述宿主细胞可从美国型培养物保藏中心(American Type Culture Collection)获得。

[0067] 如通过引用整体并入本文的WO/2017/106583中所描述,通过提供能够在发酵培养物中以高细胞密度生长的适合宿主细胞解决以商业规模和可溶形式产生基因产物,如治疗性蛋白,并且这可以通过高度可控的诱导型基因表达在氧化宿主细胞质中产生可溶性基因产物。通过组合具有以下特性中的一些特性或全部特性来产生本公开的具有这些质量的宿主细胞。(1)通过增加细胞质中的氧化性多肽的表达或功能,和/或通过减少细胞质中的还原性多肽的表达或功能,对宿主细胞进行基因修饰以使其具有氧化性细胞质。本文提供了此类遗传改变的具体实例。任选地,宿主细胞还可以经基因修饰以表达有助于产生期望基因产物的伴侣蛋白和/或辅因子和/或糖基化多肽基因产物。(2)宿主细胞包括一个或多个被设计成用于一个或多个所关注的基因产物的表达的表达构建体;在某些实施例中,至少一个表达构建体包括诱导型启动子和编码由诱导型启动子表达的基因产物的多核苷酸。(3)宿主细胞含有另外的基因修饰,所述基因修饰被设计成改善来自表达构建体的基因产物表达的某些方面。在特定实施例中,宿主细胞(A)具有编码至少一个诱导型启动子的诱导物的转运蛋白的至少一个基因的基因功能的改变,并且作为另一实例,其中编码所述转运蛋白的基因选自由以下组成的组:araE、araE、araG、araH、rhaT、xylF、xylG和xylH,或特别是araE,或其中基因功能的改变更具体地是来自组成型启动子的araE的表达;和/或(B)具有编码代谢至少一个诱导型启动子的诱导物的蛋白质的至少一个基因的降低水平的基因功能,并且作为另外的实例,其中编码代谢所述至少一个诱导型启动子的诱导物的蛋白质的基因选自由以下组成的组:araA、araB、araD、prpB、prpD、rhaA、rhaB、rhaD、xylA和xylB;和/或(C)具有编码涉及至少一个诱导型启动子的诱导物的生物合成的蛋白质的至少一个基因的降低水平的基因功能,在另外的实施例中,所述基因选自由以下组成的组:scpA/sbm、argK/ygfD、scpB/ygfG、scpC/ygfH、rmlA、rmlB、rmlC和rmlD。

[0068] 具有氧化性细胞质的宿主细胞。本公开的表达系统被设计成表达基因产物;在本公开的某些实施例中,基因产物在宿主细胞中表达。提供了宿主细胞的实例,所述宿主细胞允许包含多聚体产物的组分的基因产物的高效和低成本表达。除了培养中的分离细胞之外,宿主细胞还可以包含作为多细胞生物体一部分的细胞,或在不同生物体或生物体的系统内生长的细胞。在本公开的某些实施例中,宿主细胞是微生物细胞,如酵母(酵母属

(Saccharomyces)、裂殖酵母属(Schizosaccharomyces)等)或细菌细胞,或者是革兰式阳性细菌或革兰式阴性细菌,或者是大肠杆菌,或者是大肠杆菌B菌株,或者是大肠杆菌(B菌株)EB0001细胞(也被称为大肠杆菌ASE(DGH)细胞),或者是大肠杆菌(B菌株)EB0002细胞。在用具有氧化性细胞质的大肠杆菌宿主细胞进行的生长实验中,特别是大肠杆菌B菌株SHuffle®表达(NEB目录号C3028H)和SHuffle®T7表达(NEB目录号C3029H)以及大肠杆菌K菌株SHuffle®T7(NEB目录号C3026H),这些具有氧化性细胞质的大肠杆菌B菌株能够生长到远远高于最接近对应的大肠杆菌K菌株的细胞密度(WO/2017/106583)的细胞密度。

[0069] 对宿主细胞基因功能的改变。可以对包括诱导型表达构建体的宿主细胞的基因功能进行某些改变,以通过诱导物促进宿主细胞群体的高效且均匀的诱导。优选地,如通过Khlebnikov等人在WO/2017/106583中的实例9中描述的方法所测量的,表达构建体、宿主细胞基因型和诱导条件的组合会导致培养物中至少75%(更优选地至少85%,并且最优选地至少95%)的细胞表达来自每个诱导型启动子的基因产物。对于除大肠杆菌外的宿主细胞,这些改变可以涉及与大肠杆菌基因结构上类似的基因的功能,或在宿主细胞内执行与大肠杆菌基因的功能类似的功能的基因。对宿主细胞基因功能的改变包含通过使整个基因蛋白编码序列缺失,或使足够大的基因部分缺失,将序列插入到基因中,或以其它方式改变基因序列来消除或减少基因功能,使由所述基因制成的功能性基因产物的水平降低。对宿主细胞基因功能的改变还包含通过例如改变天然启动子以创建引导基因的更高水平的转录的更强的启动子,或者将错义突变引入到导致更高活性的基因产物的蛋白编码序列中来增强基因功能。对宿主细胞基因功能的改变包含以任何方式改变基因功能,所述方式包含例如改变天然诱导型启动子以创建被组成性地激活的启动子。除了改变用于诱导物(如本文所描述地与诱导型启动子有关)的转运和代谢的基因功能和/或改变的伴侣蛋白的表达之外,还可能改变宿主细胞的还原-氧化环境。

[0070] 宿主细胞的还原-氧化环境。在如大肠杆菌等细菌细胞中,需要二硫键的蛋白通常会被输出到周质中,在所述周质中,由包括DsbABCD和DsbG的Dsb系统催化二硫键的形成和异构化。通常全部转运到周质中的半胱氨酸氧化酶DsbA、二硫键异构体酶DsbC或Dsb蛋白的组合的表达增加已经用于需要二硫键的异源蛋白的表达(Makino等人,《微生物细胞工厂(Microb Cell Fact)》2011年5月14日;10:32)。还可能表达这些Dsb蛋白的细胞质形式,如DsbA和/或DsbC(‘cDsbA’或‘cDsbC’)的细胞质版本,其缺乏信号肽并且因此不转运到周质中。细胞质Dsb蛋白,如cDsbA和/或cDsbC用于使宿主细胞的细胞质更具氧化性,并且因此更有利于在细胞质中产生的异源蛋白中形成二硫键。通过直接改变硫氧还蛋白和谷氧还蛋白/谷胱甘肽酶系统也可以使宿主细胞质减少并且因此更具氧化性:谷胱甘肽还原酶(gor)或谷胱甘肽合成酶(gshB)中有缺陷的突变菌株,连同硫氧还蛋白还原酶(trxB)一起使细胞质具有氧化性。这些菌株不能减少核糖核苷酸并且因此不能在不存在外源还原剂如二硫苏糖醇(DTT)的情况下生长。基因ahpC中的编码过氧化物酶AhpC的抑制突变(如ahpC*或ahpCA,Lobstein等人,《微生物细胞工厂》2012年5月8日;11:56;doi:10.1186/1475-2859-11-56),将其转化为产生还原型谷胱甘肽的二硫还原酶,从而允许将电子引导到酶核糖核苷酸还原酶上,并且使gor和trxB中有缺陷的细胞,或gshB和trxB中有缺陷的细胞在不存在DTT的情况下生长。不同种类的AhpC的突变形式可以允许在γ-谷氨酰半胱氨酸合成酶(gshA)的活性中有缺陷和在trxB中有缺陷的菌株在不存在DTT的情况下生长;这些包含

AhpC V164G、AhpC S71F、AhpC E173/S71F、AhpC E171Ter和AhpC dupl62-169(Faulkner等人,《美国国家科学院院刊》2008年5月6日;105(18):6735-6740,电子版2008年5月2日)。在此类具有氧化性细胞质的菌株中,暴露的蛋白半胱氨酸在硫氧还蛋白的催化过程中易被氧化,在逆转其生理功能时,从而形成二硫键。可以有助于对减轻氧化细胞质的宿主细胞中的氧化应激影响的其它蛋白是由大肠杆菌katG编码的HPI(氢过氧化物酶I)过氧化氢酶-过氧化物酶和由大肠杆菌katE编码的HPII(氢过氧化物酶II)过氧化氢酶-过氧化物酶,其使过氧化物歧化成水和O2(Farr和Kogoma,《微生物学评论(Microbiol Rev.)》1991年12月;55(4):561-585;评论)。通过经诱导的共表达或通过提高的组成型表达的水平来增加宿主细胞中的KatG和/或KatE蛋白的水平是本公开的一些实施例的一方面。

[0071] 另一种可以对宿主细胞进行的改变是在宿主细胞质中从酵母线粒体的内膜空间表达巯基氧化酶Ervlp,这已经显示出,即使在不存在gor或trxB中的突变的情况下,增加大肠杆菌的胞质中的各种真核来源的复杂二硫键蛋白的产生(Nguyen等人,《微生物细胞工厂》2011年1月7日;10:1)。

[0072] 包括表达构建体的宿主细胞优选地还表达cDsbA和/或cDsbC和/或Ervlp;缺乏trxB基因功能;并且还缺乏gor、gshB或gshA的基因功能;任选地具有增加水平的katG和/或katE基因功能;并且表达适当的AhpC突变形式,以便宿主细胞可以在不存在DTT的情况下生长。

[0073] 伴侣蛋白。在一些实施例中,期望的基因产物与其它基因产物,如伴侣蛋白共表达,其有益于期望的基因产物的产生。伴侣蛋白是这样一种蛋白,所述蛋白帮助其它基因产物的非共价折叠或展开和/或组装或分解,但当所得单体或多聚体基因产物结构执行其正常生物功能(已完成折叠和/或组装过程)时,所述蛋白不出现在所述结构中。伴侣蛋白可以在表达构建体内由诱导型启动子或组成型启动子表达,或者可以由宿主细胞染色体表达;优选地,宿主细胞中的伴侣蛋白是以足够高的水平来表达以产生共表达的基因产物,所述基因产物适当地折叠和/或组装成期望的产物。大肠杆菌宿主细胞中存在的伴侣蛋白的实例有折叠因子DnaK/DnaJ/GrpE、DsbC/DsbG、GroEL/GroES、IbpA/IbpB、Skp、Tig(触发因子)和FkpA,这些已被用于防止细胞质或周质蛋白质的蛋白质聚集。DnaK/DnaJ/GrpE、GroEL/GroES和ClpB可以协同起作用以帮助蛋白质折叠并且因此,这些伴侣蛋白的组合表达已经显示出有益于蛋白质表达(Makino等人,《微生物细胞工厂》2011年5月14日;10:32)。当在原核宿主细胞中表达真核蛋白时,在本公开的某些实施例中,来自同一或相关真核物种的真核伴侣蛋白,如蛋白二硫键异构酶(PDI)与期望的基因产物共表达或诱导性地共表达。

[0074] 一种可以在宿主细胞中表达的伴侣蛋白是来自特异腐质霉(Humicola insolens)(土壤丝孢菌(软腐真菌))的蛋白二硫键异构酶。特异腐质霉PDI的氨基酸序列如WO/2017/106583中的SEQ ID NO:1所示;其缺少天然蛋白的信号肽,因此其保留在宿主细胞质中。编码PDI的核苷酸序列被优化以在大肠杆菌中表达;PDI的表达构建体示出为WO/2017/106583的SEQ ID NO:2)。SEQ ID NO:2在其5'端处含有GCTAGC Nhel限制性位点,在核苷酸7至12处含有AGGAGG核糖体结合位点,在核苷酸21至1478处含有PDI编码序列,并且在其3'端处含有GTCGAC Sail限制性位点。SEQ ID NO:2的核苷酸序列被设计成紧邻启动子,如诱导型启动子下游插入。SEQ ID NO:2中的Nhel和Sail限制性位点可以用于将其插入到载体多克隆位点中,如pSOL表达载体(WO/2017/106583中的SEQ ID NO:3)的多克隆位点,在公布的美国专

利申请US2015353940A1中有所描述,所述美国专利申请通过引用整体并入本文。其它PDI多肽也可以在宿主细胞中表达,包含来自各种物种的PDI多肽(酿酒酵母(UniProtKB PI 7967)、智人(UniProtKB P07237)、小家鼠(UniProtKB P09103)、秀丽隐杆线虫(Caenorhabditis elegans)(UniProtKB Q 17770和Q 17967)、拟南芥(Arabdopsis thaliana)(UniProtKB 048773、Q9XI01、Q9S G3、Q9LJU2、Q9MAU6、Q94F09和Q9T042)、黑曲霉(UniProtKB Q12730))和此类PDI多肽的经修饰的形式。在本公开的某些实施例中,在本公开的宿主细胞中表达的PDI多肽与WO/2017/106583的SEQ ID NO:1的在长度的至少50%(或至少60%、或至少70%、或至少80%、或至少90%)上共享至少70%、或80%、或90%、或95%的氨基酸序列同一性,其中氨基酸序列同一性根据WO/2017/106583的实例10确定。

[0075] 辅因子的细胞转运。当使用本公开的表达系统来产生需要辅因子发挥功能的酶时,有助于使用能够从可用的前体中合成辅因子或从环境中收取辅因子的宿主细胞。常见的辅因子包含ATP、辅酶A、黄素腺嘌呤二核苷酸(FAD)、NAD+/NADH和血红素。可以将编码辅因子转运多肽和/或辅因子合成多肽的多核苷酸引入到宿主细胞中,并且此类多肽可以被组成型地表达,或者与通过本公开方法将要产生的基因产物诱导性地共表达。

[0076] 多肽基因产物的糖基化。宿主细胞可以具有使多肽糖基化的能力的改变。例如,真核宿主细胞在糖基转移酶和/或寡糖转移酶基因中的基因功能可以被消除或减少,从而损害多肽形成糖蛋白的正常真核糖基化。如通常不会使多肽糖基化的原核宿主细胞,如大肠杆菌可以被改变以表达提供糖基化功能的一组真核和原核基因(DeLisa等人,WO2009089154A2,2009年7月16日)。

[0077] 具有改变的基因功能的可用的宿主细胞菌株。为了创建用于本公开的表达系统和方法的优选的宿主细胞菌株,用已包括期望的遗传改变的菌株开始是有用的(表A;WO/2017/106583)。

[0078] 表A.示例性宿主细胞菌株

[0079]

菌株	基因型	来源
大肠杆菌TOP10	F- *mcrA Δ(mrr-hsdRMS-mcrBC) Φ80lacZ ΔM15 Δlacx74 recA1 araD139 Δ(ara-leu)7697 galU galK rpsL (Str^R) endA1 nupG λ*	英杰生命技术公司(Invitrogen Life Technologies)目录号. C404-10, C4040-03, C4040-06, C4040-50, and C4040-52
大肠杆菌™2	*Δ(ara-leu)7697 Δlacx74 ΔphoA PvuII phoR araD139 ahpC gale galK rpsL F'[lac$^+$ lacI^q pro] gor522::Tn10 trxB (Str^R, Tet^R)*	默克公司(Merck)(EMD 密理博(Millipore)化学品)目录号 71344
大肠杆菌SHuffle®表达	*fhuA2 [lon] ompT ahpC gal λatt::pNEB3-r1-cDsbc (Spec, lacI) ΔtrxB sulA11 R(mcr-73::miniTn10—Tet^S)2 [dcm] R(zgb-210::Tn10—Tet^S) endA1 Δgor Δ(mcrC-mrr)114::IS10*	新英格兰生物实验室(New England Biolabs)目录号. C3028H

[0080] 表达构建体

[0081] 在本公开的一些实施例中,诱导型启动子被设想与表达构建体一起使用。在本文中描述了示例性启动子并且在WO/2016/205570中描述了示例性启动子,所述文献通过引用整体并入本文。如本文所描述的,包括一个或多个表达构建体的细胞可以任选地包含一个或多个诱导型启动子,以表达所关注的基因产物。在一个实施例中,基因产物是如本文所描

述的融合蛋白。在其它实施例中,基因产物是蛋白,例如,治疗性蛋白。

[0082] 表达构建体。表达构建体是被设计成用于表达一种或多种所关注的基因产物的多核苷酸,并且因此不是天然存在的分子。表达构建体可以整合到宿主细胞染色体中,或者作为独立于宿主细胞染色体,如质粒或人工染色体复制的多核苷酸分子维持在宿主细胞内。表达构建体的实例是由将一个或多个多核苷酸序列插入到宿主细胞染色体中产生的多核苷酸,其中所插入的多核苷酸序列改变染色体编码序列的表达。表达载体是特异性地用于表达一种或多种基因产物的质粒表达构建体。一个或多个表达构建体可以整合到宿主细胞染色体中或者维持在染色体外多核苷酸,如质粒或人工染色体上。以下是可以用于基因产物(包含如本文所描述的融合蛋白)的表达或共表达的表达构建体中的特定类型的多核苷酸序列的描述。

[0083] 复制起点。表达构建体必须包括复制起点(也称为复制子),以便作为独立复制的多核苷酸被维持在宿主细胞内。使用相同复制机制的不同复制子无法通过重复的细胞分裂彼此一起维持在单个宿主细胞中。因此,如WO/2016/205570的表2所示,质粒可以取决于其含有的复制起点分类为不相容组。基于不相容组、拷贝数和/或宿主范围等标准,可以选择用于表达构建体中的复制起点。如上文所描述的,如果两种或更多种不同的表达构建体要在用于多基因产物的共表达的同一宿主细胞中使用,最好是如果不同的表达构建体含有来自不同不相容组的复制起点:例如,一个表达构建体中的pMB1复制子和另一个表达构建体中的p15A复制子。相对于宿主染色体分子的数量,表达构建体在细胞中的平均拷贝数由所述表达构建体中所含的复制起点确定。拷贝数可以在每个细胞几个拷贝至数百个的范围内(WO/2016/205570的表2)。在本公开的一个实施例中,使用了不同的表达构建体,其包括由相同诱导物激活的诱导型启动子,但其具有不用的复制起点。通过选择使每个不同的表达构建体以一定的近似拷贝数维持处于细胞中的复制起点,相对于由不同表达构建体表达的另一种基因产物,可能调整由一种表达构建体表达的基因产物的总体产生水平。作为实例,为了共表达多聚体蛋白的亚基A和B,创建包括colE1复制子、am启动子和由am启动子表达的用于亚基A的编码序列的表达构建体:‘colE1-Para-A’。

[0084] 创建包括p1 5A复制子、am启动子和用于亚基B的编码序列的另一个表达构建体:‘p15A-Para-B’。这两个表达构建体可以彼此一起维持在同一宿主细胞中,并且通过向生长培养基添加一种诱导物,即阿拉伯糖来诱导亚基A和亚基B的表达。如果亚基A的表达水平相对于亚基B的表达水平需要显著增加,为了使这两个亚基的表达量的化学计量比更接近期望的比率,例如,可以为亚基A创建新的表达构建体,所述表达构建体具有如存在于pUC9质粒(‘pUC9ori’)的复制起点中的经修饰的pMB 1复制子:pUC9ori-Para-A。相对于由p15A-Para-B表达亚基B,由高拷贝数表达构建体,如pUC9ori-Para-A表达亚基A应增加亚基A的产量。以类似方式,使用使表达构建体维持在较低拷贝数的复制起点,如pSO01(WO/2016/205570),可以降低由所述构建体表达的基因产物的总体水平。选择复制起点还可以确定哪些宿主细胞可以维持包括所述复制子的表达构建体。例如,包括colE1复制起点的表达构建体具有相对较窄范围的可用的宿主,属于肠杆菌科内的物种,虽然包括RK2复制子的表达构建体可以维持处于大肠杆菌、铜绿假单胞菌、恶臭假单胞菌、棕色固氮菌(Azotobacter vinelandii)和真养产碱杆菌内,并且如果表达构建体包括RK2复制子和来自RK2质粒的一些调控基因,所述表达建构体可以维持处于各种宿主细胞中,如苜蓿中华根瘤菌、根瘤农杆

菌、新月柄杆菌、乙酸钙不动杆菌和类球红杆菌(Kiies和Stahl,《微生物学评论》1989年12月;53(4):491-516)。

[0085] 类似的考虑可以用于创建用于真核细胞中的诱导型表达或共表达的表达构建体。例如,酿酒酵母的2微米圆质粒与来自其它酵母菌株的质粒如pSR1(ATCC保藏号48233和66069;Araki等人,《分子生物学杂志(J Mol Biol)》1985年3月20日;182(2):191-203)和pKD1(ATCC保藏号37519;Chen等人;《核酸研究(Nucleic Acids Res)》1986年6月11日;14(11):4471-4481)相容。

[0086] 可选择标志物。表达构建体通常包括选择基因,也称为可选择标志物,其编码宿主细胞在选择性培养基中存活或生长需要的蛋白质。未含有包括选择基因的表达构建体的宿主细胞将不会在培养基中存活。典型的选择基因编码赋予抗生素或其它毒素抗性或补充宿主细胞的营养缺陷型缺陷的蛋白质。选择方案的一个实例利用如抗生素等药物来抑制宿主细胞的生长。含有包括可选择标志物的表达构建体的那些细胞产生赋予药物抗性的蛋白质并且在选择方案中存活。常用于选择可选择标志物(和指示提供抗生素抗性表型的基因的缩写)的抗生素的一些实例为:氨苄西林(AmpR)、氯霉素(CmlR或CmR)、卡那霉素(KanR)、大观霉素(SpcR)、链霉素(StrR)和四环素(TetR)。WO/2016/205570的表2中的许多质粒包括可选择标志物,如pBR322(AmpR、TetR);pMOB45(CmR、TetR);pACYC1W(AmpR、KanR);以及pGBM1(SpcR、StrR)。用于选择基因的天然启动子区通常与其基因产物的编码序列一起包含作为表达构建体的可选择标志物部分的一部分。可替代地,用于选择基因的编码序列可以由组成型启动子表达。

[0087] 在各个方面中,合适的可选择标志物包含但不限于新霉素磷酸转移酶(npt II)、潮霉素磷酸转移酶(hpt)、二氢叶酸还原酶(dhfr)、吉欧霉素、腐草霉素、博来霉素抗性基因(ble)、庆大霉素乙酰转移酶、链霉素磷酸转移酶、乙酰乳酸合酶(als)的突变体形式、溴苯腈腈水解酶、草胺膦乙酰转移酶(bar)、烯醇丙酮莽草酸-3-磷酸(EPSP)合酶(aro A)、肌肉特异性酪氨酸激酶受体分子(MuSK-R)、铜-锌超氧化物歧化酶(sod1)、金属硫蛋白(cup1,MT1)、β-内酰胺酶(BLA)、嘌呤霉素N-乙酰基-转移酶(pac)、杀稻瘟菌素乙酰转移酶(bls)、杀稻瘟菌素脱氨酶(bsr)、组氨醇脱氢酶(HDH)、N-琥珀酰基-5-氨基咪唑-4-甲酰胺核糖酸(SAICAR)合酶(ade1)、精氨琥珀酸裂解酶(arg4)、β-异丙基苹果酸脱氢酶(leu2)、转化酶(suc2)、乳清苷-5'-磷酸(OMP)脱羧酶(ura3)和前述的任何一种的直系同源物。

[0088] 诱导型启动子。如本文所描述的,存在若干不同的诱导型启动子,所述诱导型启动子可以包含在表达构建体中作为本公开的诱导型共表达系统的一部分。优选的诱导型启动子与如WO/2016/205570的表1中所定义的通过参考大肠杆菌K-12亚菌株MG1655基因组序列的启动子多核苷酸序列的至少30个(更优选地至少40个以及最优选地至少50个)连续碱基共享至少80%多核苷酸序列同一性(更优选地至少90%同一性以及最优选地至少95%同一性),其中多核苷酸序列同一性百分比是使用WO/2016/205570的实例11的方法确定的。在'标准'诱导条件下(参见WO/2016/205570的实例5),优选的诱导型启动子具有大肠杆菌K-12亚菌株MG1655的对应"野生型"诱导型启动子的强度的至少75%(更优选地至少100%以及最优选地至少110%),如使用De Mey等人的定量PCR方法确定的(WO/2016/205570的实例6)。在表达构建体内,将诱导型启动子置于要诱导性地表达的基因产物的编码序列的5'(或"上游"),使得诱导型启动子的存在将相对于编码基因产物的多核苷酸的编码链在5'到3'

方向上指导基因产物编码序列的转录。

[0089] 核糖体结合位点。对于多肽基因产物,要诱导性地表达的基因产物的编码序列的转录起始位点与起始密码子之间的核苷酸序列区对应于所述多肽基因产物的mRNA的5'非翻译区("UTR")。优选地,对应于5'UT的表达构建体的区包括与存在于宿主细胞的物种中的共有核糖体结合位点(RBS,也称为Shine-Dalgarno序列)类似的多核苷酸序列。在原核生物(古菌和细菌)中,RBS共有序列是GGAGG或GGAGGU,并且在细菌,如大肠杆菌中,RBS共有序列是AGGAGG或AGGAGGU。RBS与起始密码子通常间隔5个至10个中间核苷酸。在表达构建体中,RBS序列优选地与AGGAGGU共有序列至少55%相同,更优选地至少70%相同,并且最优选地至少85%相同,并且与起始密码子间隔5个至10个中间核苷酸,更优选地相隔6个至9个中间核苷酸,并且最优选地间隔6个或7个中间核苷酸。给定RBS产生期望的翻译启动率的能力可以使用RBS计算器在网站salis.psu.edu/software/RBSLibraryCalculatorSearchMode上计算;同样的工具可以用于优化合成RBS的翻译率至100,000+倍的范围(Salis,《酶学方法(Methods Enzymol)》2011;498:19-42)。

[0090] 多克隆位点。多克隆位点(MCS),也称为多接头,是含有彼此紧邻或重叠的多个限制位点的多核苷酸。MCS中的限制位点通常在MCS序列内发生一次,并且优选地不在质粒或其它多核苷酸构建体的其余部分内发生,从而允许限制酶仅在MCS内切割质粒或其它多核苷酸构建体。MCS序列的实例为pBAD系列的表达载体,包含pBAD18、pBAD18-Cm、pBAD18-Kan、pBAD24、pBAD28、pBAD30和pBAD33(Guzman等人,《细菌学杂志(J Bacteriol)》1995年7月;177(14):4121-4130);或者源自pBAD载体的pPRO系列的表达载体,如pPR018、pPR018-Cm、pPR018-Kan、pPR024、pPR030和pPR033(美国专利第8178338B2号;2012年5月15日;Keasling,Jay)中的那些。多克隆位点可以用于表达构建体的创建:通过将多克隆位点置于启动子序列的3'(或下游),MCS可以用于将要表达或共表达的基因产物的编码序列在相对于启动子的合适位置中插入到所述构建体中,使得编码序列将发生转录。取决于限制酶用于在MCS内切割,在将编码序列或其它多核苷酸序列插入到表达构建体中之后,MCS序列的一些部分可能保留在表达构建体内。任何剩余的MCS序列可以是所插入的序列的上游、或下游或两侧。核糖体结合位点可以置于MCS的上游,优选地紧邻MCS或与MCS仅相隔几个核苷酸,在这种情况下,RBS是插入到MCS中的任何编码序列的上游。另一种替代方案是在MCS内包含核糖体结合位点,在这种情况下,用于在MCS内切割的限制性酶的选择将确定是否保留RBS以及RBS与所插入的序列的关系。另外的替代方案是在多核苷酸序列内包含RBS,所述多核苷酸序列要在MCS处插入到表达构建体内中的,优选地与任何编码序列保持合适的关系,以刺激从经转录的信使RNA开始翻译。

[0091] 来自组成型启动子的表达。本公开的表达构建体还可以包括由组成型启动子表达的编码序列。与诱导型启动子不同,组成型启动子在大多数生长条件下发起持续基因产物产生。组成型启动子的一个实例是Tn3 bla基因的组成型启动子,其编码β-内酰胺酶,并且负责由许多质粒(包含pBR322(ATCC 31344)、pACYC1W(ATCC 37031)和pBAD24(ATCC 87399))赋予在宿主细胞上的氨苄西林抗性(AmpR)表型。另一个可以在表达构建体中使用的组成型启动子是用于大肠杆菌脂蛋白基因的启动子Ipp,其位于大肠杆菌K-12亚菌株MG1655中的位置的1755731-1755406(正链)处(Inouye和Inouye,《核酸研究》1985年5月10日;13(9):3101-3110)。在大肠杆菌中用于异源基因表达的组成型启动子的另外的实例是

trpLEDCBA启动子,位于大肠杆菌K-12亚菌株MG1655中的位置1321169-1321133(负链)处(Windass等人,《核酸研究》1982年11月11日;10(21):6639-6657)。如本文所描述的,组成型启动子可以在表达构建体中使用以用于可选择标志物的表达,并且也可以用于对期望的产物的共表达有用的其它基因产物的组成型表达。例如,诱导型启动子的转录调控因子,如AraC、PrpR、RhaR和XylR,如果不是由双向诱导型启动子表达,可以替代性地在与其调控的诱导型启动子相同的表达构建体,或不同的表达构建体上由组成型启动子表达。类似地,作为几个实例,用于产生或运输诱导物的基因产物,如PrpEC、AraE或Rha,或者修饰细胞的还原氧化环境的蛋白质可以在表达构建体内由组成型启动子表达。用于产生共表达基因产物和所得期望产物的基因产物还包含伴侣蛋白、辅因子转运蛋白等。

[0092] 信号肽。本公开的方法表达或共表达的多肽基因产物可以含有信号肽或者缺乏信号肽,这取决于分别期望此类基因产物是要从宿主细胞质中输出到周质中,或者要保留在细胞质中。信号肽(又称为信号序列、前导序列或前导肽)在结构上以约五至二十个氨基酸长并且通常约十至十五个氨基酸长的疏水性氨基酸链段为特征,所述疏水性氨基酸链段具有形成单α-螺旋的倾向。这一疏水性链段通常刚好在富集带正电氨基酸(特别是赖氨酸)的较短链段之前。从成熟多肽切割的信号肽通常以被信号肽酶识别和切割的氨基酸链段结束。信号肽在功能上的特征可以在于:能够将共翻译或翻译后的多肽通过原核生物的质膜(或革兰氏阴性细菌,如大肠杆菌的内膜)直接转运或转运至真核细胞的内质网中。例如,信号肽使多肽能够转运到宿主细胞,如大肠杆菌的周质空间中的程度可以使用如WO/2016/205570的实例12中所描述的方法,通过将周质蛋白与保留在细胞质中的蛋白质分离来确定。

[0093] 以下将描述可以在用于表达或共表达基因产物的表达构建体中使用的诱导型启动子,以及可以对含有此类表达构建体的宿主细胞作出的一些遗传修饰。除非另有说明,否则这些诱导型启动子和相关基因的实例来自大肠杆菌(Escherichia coli/E.coli)菌株MG1655(美国典型培养物保藏中心(American Type Culture Collection)保藏为ATCC 700926),所述菌株是大肠杆菌K-12(美国典型培养物保藏中心保藏为ATCC 10798)的亚菌株。WO/2016/205570的表1列出了用于诱导型启动子和相关基因的这些实例的核苷酸序列在大肠杆菌MG1655中的基因组位置。WO/2016/205570的表1中由基因组位置引用的核苷酸和其它基因序列通过引用明确并入本文。关于本文所描述的大肠杆菌启动子、基因和菌株的另外的信息可以在许多公共来源,包含位于ecoliwiki.net的在线EcoliWiki资源中找到。

[0094] 阿拉伯糖启动子。(如本文所使用的,“阿拉伯糖”意指L-阿拉伯糖。)涉及阿拉伯糖利用的若干大肠杆菌操纵子由阿拉伯糖araBAD、araC、arciE和araFGH诱导,但术语“阿拉伯糖启动子”和“ara启动子”通常用于指定araBAD启动子。若干另外的术语已经用于指示大肠杆菌araBAD启动子,如Para、ParaB、ParaBAD和PBAD-本文使用的“ara启动子”或上述给定的任何替代性术语意指大肠杆菌araBAD启动子。从另一个术语‘araC-araBAD启动子’的使用可以看出,araBAD启动子被视为双向启动子的一部分,其中araBAD启动子在一个方向上控制araBAD操纵子的表达,并且非常接近araBAD启动子并且在araBAD启动子的相对链上的araC启动子在另一方向上控制araC编码序列的表达。AraC蛋白质是araBAD启动子的正和负转录调控因子。在不存在阿拉伯糖的情况下,AraC蛋白抑制PBAD的转录,但在存在阿拉伯糖

的情况下,在与阿拉伯糖结合后会改变其构象的AraC蛋白成为允许PBAD转录的正向调控元件-araBAD操纵子编码蛋白质,所述蛋白质通过中间体L-核酮糖和L-核酮糖-磷酸盐将L-阿拉伯糖转化为D-木酮糖-5-磷酸盐来代谢L-阿拉伯糖。为了将对由阿拉伯糖诱导型启动子进行的表达的诱导作用最大化,消除或减弱催化L-阿拉伯糖转化成L-核酮糖的AraA的功能以及任选地也消除或减弱AraB和AraD中至少一者的功能极其有用。通过消除或减弱细胞将阿拉伯糖转化成其它糖的能力来消除或减弱宿主细胞减小细胞中阿拉伯糖的有效浓度的能力使得有较多的阿拉伯糖可用于阿拉伯糖诱导型启动子的诱导。编码将阿拉伯糖移动至宿主细胞中的转运蛋白的基因是araE,所述基因编码低亲和力L-阿拉伯糖质子同向转运蛋白;和araFGH操纵子,所述操纵子编码ABC超家族高亲和力L-阿拉伯糖转运蛋白的亚基。可以将L-阿拉伯糖转运到细胞中的其它蛋白是LacY乳糖渗透酶的某些突变体:LacY(AlWC)和LacY(AlWV)蛋白,其在位置177处分别有替代丙氨酸的半胱氨酸或缬氨酸氨基酸,(Morgan-Kiss等人,《美国国家科学院院刊》2002年5月28日;99(11):7373-7377)。为了实现阿拉伯糖诱导型启动子的均匀诱导作用,不依赖于阿拉伯糖的调控将阿拉伯糖转运至细胞中是有用的。这可以通过消除或减弱AraFGH转运蛋白的活性并改变araE的表达,由此使其仅由组成型启动子转录来实现。araE的组成型表达可以通过消除或减弱天然araE基因的功能,并将表达构建体引入细胞中来实现,所述表达构建体包含由组成型启动子表达的AraE蛋白质的编码序列。或者,在缺乏AraFGH功能的细胞中,控制宿主细胞染色体araE基因的表达的启动子可由阿拉伯糖诱导型启动子变为组成型启动子。以类似方式,作为阿拉伯糖诱导型启动子的均匀诱导作用的另外的替代选择,缺乏AraE功能的宿主细胞可以具有由组成型启动子表达的任何功能性AraFGH编码序列存在于细胞中。作为另一个替代选择,可通过用宿主染色体中的组成型启动子替代天然araE和araFGH启动子,由组成型启动子表达araE基因和araFGH操纵子。也可消除或减弱AraE和AraFGH阿拉伯糖转运蛋白的活性,并且在此情况下,可使用LacY乳糖通透酶中的突变,所述突变允许这一蛋白质转运阿拉伯糖。由于lacY基因的表达通常不受阿拉伯糖调控,故在阿拉伯糖诱导型启动子因存在阿拉伯糖而受到诱导时,使用LacY突变体,例如LacY(A177C)或LacY(A177V)将不会导致“全有或全无(all or none)”诱导现象。因为LacY(A177C)蛋白质看来在将阿拉伯糖转运至细胞中方面更为有效,所以使用编码LacY(A177C)蛋白质的多核苷酸相对于使用编码LacY(A177V)蛋白质的多核苷酸是优选的。

[0095] 丙酸盐启动子。‘丙酸盐启动子’或‘prp启动子’是大肠杆菌prpBCDE操纵子的启动子,并且也被称为PrpΦB。与ara启动子类似,prp启动子是双向启动子的一部分,在一个方向上控制prpBCDE操纵子的表达,并且prpR启动子在另一个方向上控制prpR编码序列的表达。PrpR蛋白是prp启动子的转录调控因子,并且当PrpR蛋白与2-甲基柠檬酸(‘2-MC’)结合时激活从prp启动子进行的转录。丙酸盐(Propionate)(也称为丙酸盐(propanoate))是丙酸(propionic acid或‘propanoic acid’)的离子,即$CH_3CH_2COO^-$,并且是具有通式$H(CH_2)_nCOOH$的最小‘脂肪’酸,所述‘脂肪’酸共享此类分子的某些特性:当从水中盐析出来时产生油层以及具有皂样钾盐。可商购的丙酸盐一般是以丙酸的一价阳离子盐形式销售,例如丙酸钠(CH_3CH_2COONa),或以其二价阳离子盐形式销售,例如丙酸钙($Ca(CH_3CH_2COO)_2$)。丙酸盐是可渗透膜的并且通过以下被代谢成2-MC:通过PrpE(丙酰基-CoA合成酶)将丙酸盐转化成丙酰基-CoA,并且然后通过PrpC(2-甲基柠檬酸合成酶)将丙酰基-CoA转化成2-

MC。由prpBCDE操纵子、PrpD(2-甲基柠檬酸脱水酶)和PrpB(2-甲基异柠檬酸裂解酶)编码的其它蛋白质涉及将2-MC进一步分解代解成较小的产物,如丙酮酸盐和琥珀酸盐。为了使添加至细胞生长培养基中的丙酸盐对丙酸盐诱导型启动子的诱导作用最大化,因此期望宿主细胞具有PrpC和PrpE活性以将丙酸盐转化成2-MC,而且还具有消除或减弱的PrpD活性,以及任选地消除或减弱的PrpB活性,以防止2-MC被代谢。编码涉及2-MC生物合成的蛋白的另一操纵子是scpA-argK-scpBC操纵子,又被称为sbm-yg/DGH操纵子。这些基因编码进行以下所需的蛋白:将琥珀酸盐转化成丙酰基-CoA,所述丙酰基-CoA然后可以通过PrpC转化成2-MC,。消除或减弱这些蛋白质的功能将去除2-MC诱导物制造的平行路径,并因此可减小丙酸盐诱导型启动子的背景表达水平,并增加丙酸盐诱导型启动子对外源供应的丙酸盐的敏感性。已发现,在大肠杆菌BL21(DE3)中引入缺失sbm-ygfD-ygfG-ygfH-ygfI以创建菌株JSB(Lee和Keasling,“用于肠道细菌的丙酸盐诱导型表达系统(A propionate-inducible expression system for enteric bacteria)”,《应用与环境微生物学(Appl Environ Microbiol)》2005年11月;71(11):6856-6862)有助于降低在不存在外源供应的诱导物的情况下的背景表达,但这一缺失也降低菌株JSB中的来自prp启动子的总体表达。然而,应注意,缺失sbm-ygfD-ygfG-ygfH-ygfI还明显影响ygfI,其编码功能未知的推定的LysR家族转录调控因子。基因sbm-yg/DGH被转录为一个操纵子,并且ygfI是从相对链转录的。ygfti与ygfI编码序列的3’端重叠几个碱基对,因此去除所有sbm-yg/DGH操纵子的缺失也明显去除ygfI编码功能。消除或减弱sbm-ygfDGH基因产物的子集,如YgfG(又称为ScpB,甲基丙二酰-CoA脱羧酶)的功能,或使大部分sbm-yg/DGH(或scpA-argK-scpBC)操纵子缺失,同时留下足够的ygfIi(或scpC)基因的3’端以使ygfI的表达不受影响,可以充分降低来自丙酸盐诱导型启动子的背景表达,但不降低所诱导的表达的最大水平。

[0096] 鼠李糖启动子。(如本文所使用的,“鼠李糖”意指L-鼠李糖。)“鼠李糖启动子”、或“rha启动子”或PrhaSR是用于大肠杆菌rhaSR操纵子的启动子。与ara和prp启动子类似,rha启动子是双向启动子的一部分,所述启动子在一个方向上控制rhaSR操纵子的表达,并且rhaBAD启动子在另一方向上控制rhaBAD操纵子的表达。然而,rha启动子具有两个转录调控因子涉及到调节表达:RhaR和RhaS。在存在鼠李糖的情况下,RhaR蛋白激活rhaSR操纵子的表达,而RhaS蛋白分别激活L-鼠李糖分解代谢和转运操纵子rhaBAD和rhaT的表达(Wickstrum等人,《细菌学杂志》2010年1月;192(1):225-232)。尽管RhaS蛋白质也可激活rhaSR操纵子的表达,但实际上,RhaS通过干扰环状AMP受体蛋白(CRP)与RhaR将表达共激活至大得多的水平的能力来负面地自调控这一表达。rhaBAD操纵子编码以下鼠李糖分解代谢蛋白:RhaA(L-鼠李糖异构酶),其将L-鼠李糖转化为L-鼠李树胶糖;RhaB(鼠李树胶糖激酶),其使L-鼠李树胶糖磷酸化以形成L-鼠李树胶糖-1-P;以及RhaD(鼠李树胶糖-1-磷酸醛缩酶),其将L-鼠李树胶糖-1-P转化成L-乳醛和DHAP(二羟基丙酮磷酸)。为了将细胞中可用于诱导由鼠李糖诱导型启动子引起的表达的鼠李糖的量最大化,需要通过消除或减弱RhaA,或任选地RhaA以及RhaB与RhaD中的至少一者的功能来减小因催化而分解的鼠李糖的量。大肠杆菌细胞也可以通过由rmlBDACX(或rfbBDACX)操纵子编码的蛋白质RmlA、RmlB、RmlC和RmlD(分别又称为RfbA、RfbB、RfbC和RfbD)的活性由α-D-葡萄糖-1-P合成L-鼠李糖。为了降低来自鼠李糖诱导型启动子的背景表达,并且为了增强外源供应的鼠李糖对鼠李糖诱导型启动子的诱导的敏感性,消除或减弱RmlA、RmlB、RmlC和

[0097] RmlD蛋白中的一种或多种的功能可能有用。L-鼠李糖通过RhaT、鼠李糖渗透酶或L-鼠李糖:质子同向转运蛋白转运到细胞中。如上所述,RhaT的表达是由转录调控因子RhaS激活的。为了使RhaT的表达不依赖于鼠李糖(它诱导RhaS的表达)的诱导,可以对宿主细胞进行改变以使细胞中的所有功能性RhaT编码序列均由组成型启动子表达。此外,RhaS的编码序列可缺失或失活,由此不产生功能性RhaS。通过消除或减弱细胞中RhaS的功能,从rhaSR启动子表达的水平因不存在RhaS的负面自调控而增加,并且鼠李糖催化性操纵子rhaBAD的表达水平降低,由此进一步增加鼠李糖诱导从rha启动子进行表达的能力。

[0098] 木糖启动子。(如本文所使用的,“木糖”意指D-木糖。)木糖启动子、或“xyl启动子”或PxyiA意指用于大肠杆菌xylAB操纵子的启动子。木糖启动子区在组织上类似于其它诱导型启动子,因为xylAB操纵子和xylFGHR操纵子两者均在大肠杆菌染色体上在相反方向上由相邻木糖诱导型启动子表达(Song和Park,《细菌学杂志》1997年11月;179(22):7025-7032)。PxyiA和PxyiF启动子两者的转录调控因子是XylR,其在存在木糖的情况下激活这些启动子的表达。xylR基因是作为xylFGHR操纵子的一步法或从其自身的弱启动子表达的,所述弱启动子位于xylH与xylR蛋白质编码序列之间,不可被木糖诱导。D-木糖是通过XylA(D-木糖异构酶)分解代谢的,XylA将D-木糖转化成D-木酮糖,D-木酮糖接着被XylB(木酮糖激酶)磷酸化而形成D-木酮糖-5-P。为了使细胞中可用于诱导从木糖诱导型启动子表达的木糖的量最大化,需要通过消除或减弱至少XylA,或任选地XylA和XylB两者的功能来减小因催化而分解的木糖的量。xylFGHR操纵子编码XylF、XylG和XylH,这些是ABC超家族高亲和力D-木糖转运蛋白的亚基。xylE基因编码大肠杆菌低亲和力木糖-质子同向转运蛋白,表示单独的操纵子,其表达也由木糖诱导。为了使木糖转运蛋白的表达不依赖于木糖的诱导作用,可对宿主细胞进行改变以使所有功能性木糖转运蛋白都从组成型启动子表达。例如,可以对xylFGHR操纵子进行改变,以使xylFGH编码序列缺失,留下XylR作为由木糖诱导型PxyiF启动子表达的唯一活性蛋白,并且其中xylE编码序列由组成型启动子而非其天然启动子表达。作为另一实例,xylR编码序列在表达构建体中由PxyiA或启动子表达,而xylFGHR操纵子缺失并且xylE是组成型表达的,或者可替代地,xylFGH操纵子(缺乏xylR编码序列,因为所述编码序列存在于表达构建体中)由组成型启动子表达,并且xylE编码序列缺失或被改变,以使得所述编码序列不产生活性蛋白。

[0099] 乳糖启动子。术语‘乳糖启动子’是指用于lacZYA操纵子的乳糖诱导型启动子,所述启动子又称为lacZpl;这一乳糖启动子位于大肠杆菌K-12亚菌株MG1655的基因组序列(NCBI参考序列NC 000913.2,2012年1月11日)中的约365603-365568处(负链,在约365603-365598处具有NA聚合酶结合(‘-35’)位点,在365579-365573处具有Pribnow盒(‘-10’)并且在365567处具有转录起始位点)。在一些实施例中,本公开的诱导型共表达系统可以包括乳糖诱导型启动子,如lacZYA启动子。在其它实施例中,本公开的诱导型共表达系统包括一个或多个非乳糖诱导型启动子的诱导型启动子。

[0100] 碱性磷酸酶启动子。术语‘碱性磷酸酶启动子’和‘phoA启动子’是指用于phoApsiF操纵子的启动子,所述启动子在磷酸盐饥饿的条件下被诱导。phoA启动子区位于大肠杆菌K-12亚菌株MG1655的基因组序列(NCBI参考序列NC 000913.3,2014年12月16日)中的约401647-401746处(正链,在401695-401701处具有Pribnow盒(‘-10’)(Kikuchi等人,《核酸研究》1981年11月11日;9(21):5671-5678))。用于phoA启动子的转录激活因子是PhoB,其是

一种转录调控因子,与传感器蛋白PhoR一起形成大肠杆菌中的双组分信号转导系统。PhoB和PhoR从phoBR操纵子转录,所述操纵子位于大肠杆菌K-12亚菌株MG1655的基因组序列(NCBI参考序列NC 000913.3,2014年12月16日)中的约417050-419300处(正链,在417,142-417,831处具有PhoB编码序列并且在417,889-419,184处具有PhoR编码序列)。phoA启动子与上述诱导型启动子的不同之处在于其是通过缺乏物质,即细胞内磷酸盐而不是通过添加诱导物来诱导的。出于此原因,phoA启动子一般被用于引导将在宿主细胞耗竭磷酸盐的阶段,如在发酵后期产生的基因产物的转录。在一些实施例中,本公开的诱导型共表达系统可以包括phoA启动子。在其它实施例中,本公开的诱导型共表达系统包括一个或多个非phoA启动子的诱导型启动子。

[0101] 如本文所描述的,例如,如果携带表达构建体的细胞系用于或将用于商业目的,去除(例如,通过诱导型或组成型“固化”机制的方式)本文所描述的表达构建体可以是有利的或期望的。因此,在一些实施例中,表达构建体可以包括“杀灭开关”。例如,在实施例中,表达构建体包含温度敏感复制起点。另外的固化方法是本领域已知的,并且包含使用洗涤剂和嵌入剂、药物和抗生素(Buckner,M.M.C.等人,《FEMS微生物学评论(FEMS Microbiology Reviews),fuy031,42,2018,781-804)。

[0102] 在进一步描述本公开之前,应当理解,本公开不限于所描述的特定实施例,因为这些实施例当然可以变化。还应当理解,本文所使用的术语仅仅是出于描述具体实施例的目的,并且不旨在是限制性的,因为本公开的范围将仅由所附权利要求限定。

[0103] 在提供了值范围的情况下,应当理解,介于所述范围的上限与下限之间的每个中间值(到下限的单位的十分之一,除非另外明确说明)以及所述范围中的任何其它所陈述或中间值均涵盖于本公开内。这些较小范围的上限和下限可以独立地包含在较小的范围中,并且也涵盖在本公开内,这受制于所陈述的范围中的任何明确排除的限值。在所陈述范围包含限值中的一个或两个限值的情况下,排除被包含在内的限值中的任一个或两个限值的范围也包含在本公开内。

[0104] 除非另外定义,否则本文所使用的所有技术和科学术语都具有与本公开所属领域的普通技术人员通常所理解的含义相同的含义。虽然与本文所描述的那些方法和材料类似或等同的任何方法和材料也可以用于本公开的实践或测试中,但是现在描述优选的方法和材料。本文所提及的所有公开通过引用并入本文,以结合所引用的所述公开来公开和描述所述方法和/或材料。

[0105] 必须注意,除非上下文另外明确指示,否则如在本文和所附权利要求中所使用的单数形式“一个/一种(a)”、“和(and)”和“所述(the)”包含复数指示物。因此,例如,提及“构象转换探针”包含多个此类构象转换探针并且提及“微流控装置”包含提及一种或多种微流控装置以及本领域技术人员已知的其等价物,等等。应当进一步指出,权利要求书可以撰写成排除任何要素(例如,任何任选的要素)。因此,此陈述旨在用作前置基础,用于结合权利要求要素的叙述使用如“单独地”、“仅”等排他性术语或使用“负面”限制。

[0106] 如对于本领域技术人员将显而易见的是,在阅读本公开时,本文描述和展示的单独实施例中的每一个均具有离散的组成部分和特征,所述组成部分和特征可以在不偏离本公开的范围或精神的情况下易于与任何其它若干实施例的特征分离或组合。可以按照所叙述的事件的顺序或按照逻辑上可行的任何其它顺序来执行任何所叙述的方法。这旨在为所

有此类组合提供支持。

[0107] 实例

[0108] 实例1

[0109] 较高表达的大肠杆菌菌株的生长/不生长选择测定

[0110] 本发明实例描述并执行了简单的生长/不生长选择策略,以鉴定混合群体(文库)中的可溶性蛋白表达较高的大肠杆菌菌株。

[0111] 本文使用了大肠杆菌菌株,所述大肠杆菌菌株携带编码驱动mCherry-FolA融合蛋白的表达的阿拉伯糖启动子的质粒,以示出:(1)甲氧苄啶本身具有非常有限的可选择动态范围并且(2)添加磺胺甲噁唑使mCherry-FolA菌株对甲氧苄啶敏感。

[0112] 菌株:1号菌株。在pBAD启动子的控制下的携带编码抗生素靶基因(folA)融合物的质粒的大肠杆菌。2号菌株。携带编码不同抗生素靶基因(murA)的质粒的大肠杆菌的对照。

[0113] 实验:制备含有阿拉伯糖的甲氧苄啶抗生素梯度板。跨板的抗生素梯度将是相同的。然而,第二抗生素剂(磺胺甲噁唑)在每个板上的浓度将是一致的,但跨板的浓度是不同的。跨抗生素梯度接种菌株。生长长度是相对抗生素抗性的量度。对于携带具有folA基因的pBAD质粒的菌株,生长长度与磺胺甲噁唑的浓度反相关:含有最高磺胺甲噁唑浓度的板应具有最短的生长长度。

[0114] 如图1所示,携带mCherry-FolA融合的细胞对甲氧苄啶具有很强的抗性。图1显示出针对100ug/ml甲氧苄啶(mCherry-FolA)的抗性取决于阿拉伯糖和mCherry-FolA融合物。磺胺甲噁唑达到16ug/ml时不抑制生长。

[0115] 如图2所示,与磺胺甲噁唑的协同作用增加了甲氧苄啶动态范围。图2显示出甲氧苄啶与磺胺甲噁唑有很强的协同作用:当磺胺甲噁唑>2ug/ml时,甲氧苄啶的敏感性会恢复。

[0116] 实例2

[0117] 可溶性蛋白选择测定-胰岛素-FolA融合物

[0118] 本发明实例使用两个胰岛素-FolA融合构建体(胰岛素变体)并且使用两个大肠杆菌伴侣蛋白文库描述和执行了实例1的测定。

[0119] 结合如本文以及实例1和图1-2中所描述的,制备了琼脂板,所述琼脂板覆盖了两种药剂(甲氧苄啶和磺胺甲噁唑)的基质。组合基质是必要的,这是因为文库中表现最好的菌株所能耐受的抗生素的实际浓度是未知的。如图3所示,将两个携带空质粒(不具有伴侣蛋白)的对照菌株和具有伴侣蛋白质粒的实验菌株各自单独地平板接种在被组合基质(例如,甲氧苄啶-磺胺甲噁唑的基质)覆盖的整组板上。将来自六个菌株中的每一个菌株的每个板大约3亿个细胞单独平板接种在所有16个平板上。在最终分析之前,将板温育5天。

[0120] 在温育五天之后,在一些琼脂板中观察到对照菌株与实验菌株之间的差异。这些差异是细微的,即实验板上的菌落比对照板上的菌落大,或者是显著的,即只在实验板中观察到生长,而在对照板上没有观察到生长。对于两个胰岛素-FolA融合物/伴侣蛋白质粒对中的每一者,都随机挑选了七个阳性命中,以用于使用可溶性胰岛素HiPrBind测定进行进一步表征(例如参见WO/2021/163349)。为两种胰岛素-FolA融合物中的每一个挑选的七个菌株中,与对照菌株相比,两个菌株显示出可溶性胰岛素的量增加了。

[0121] 总之,FolA选择测定使得能够对产生更高量的所关注的可溶性蛋白-FolA融合物

的菌株进行生长/不生长选择。另外,还可以通过修饰增加协同抗生素的浓度的选择条件,获得从低到高的表达菌株的分级。最后,可以对含有数十亿细胞的大型文库进行并行测试,以鉴定执行最佳的菌株子集。

[0122] 实例3

[0123] 可溶性蛋白选择测定-甲胎蛋白

[0124] 本发明实例描述并执行了并入了固相测定的上述测定,以进一步分析细胞群体。本文以及例如在PCT/US22/53107中对固相测定进行了描述。

[0125] 在两种不同的靶蛋白(胰岛素和甲胎蛋白)上应用如上文实例中所描述的测定显示出,在选择性板上进行选择之后,所述测定产生几百个生长的菌落(候选者)。理想的目标是测试所有候选者,以确定其是真正的优胜者,还是获得逃逸机制的突变体。在此,采用固相测定以测试整个生长群体,以鉴定比背景对照实际产生更多融合蛋白的优胜者。

[0126] 使用ProteinSimpleJess自动化仪器通过毛细管免疫印迹确定用于可溶性甲胎蛋白的相对表达

[0127] 将来自100μL培养物的湿细胞糊重悬于裂解缓冲液(50mM Tris-HCl,200mM氯化钠,10.0g/L辛基葡糖苷,2400U/mL溶菌酶,2U/mL Benzonase,pH 7.2)中,并且在冰上裂解一小时。裂解后,将样品用0.1×样品缓冲液(Protein Simple)稀释至OD600为0.008,并且在4℃下以3300x g离心30分钟。离心后,将样品与含有100mM碘乙酰胺的荧光主混合物(Protein Simple)混合至OD600为0.0064,涡旋,并且加热至95℃持续5分钟。除了兔单克隆抗α1-胎蛋白初级抗体(艾博抗公司(Abcam)项编号#ab169552)和进行化学发光免疫测定所需的试剂之外,将制备好的样品装载到12-230kDa分离模块板(Protein Simple)中。将分离模块板与25毛细管盒一起装载到Jess中,并且使用默认的化学发光检测方法运行。使用用于Simple Western软件的Protein Simple Compass分析Simple Western测定结果。

[0128] 案例研究:甲胎蛋白

[0129] 构建的质粒文库具有以下特性。(图4)阿拉伯糖启动子控制靶蛋白(甲胎蛋白),所述靶蛋白在C端处与folA在框架内进行了基因融合。另外,质粒被设计成携带两个由丙酸盐启动子驱动的伴侣蛋白,所述伴侣蛋白是从大约1200个独特的伴侣蛋白基因库中随机选择的。理论多样性为144万种独特组合(1200x 1200)。

[0130] 根据下表B,制备了一组琼脂板,所述琼脂板含有10μg/ml卡那霉素、125mM阿拉伯糖、1mM丙酸盐和甲氧苄啶/磺胺甲噁唑的64种不同组合中的一种。

[0131] 表B

[0132]

磺胺甲噁唑(μg/ml)	甲氧苄啶(μg/ml) 1	2	4	8	16	32	64	128
0	TRM1_Sulfa0	TRM2_Sulfa0	TRM4_Sulfa0	TRM8_Sulfa0	TRM16_Sulfa0	TRM32_Sulfa0	TRM64_Sulfa0	TRM128_Sulfa0
1	TRM1_Sulfa1	TRM2_Sulfa1	TRM4_Sulfa1	TRM8_Sulfa1	TRM16_Sulfa1	TRM32_Sulfa1	TRM64_Sulfa1	TRM128_Sulfa1
2	TRM1_Sulfa2	TRM2_Sulfa2	TRM4_Sulfa2	TRM8_Sulfa2	TRM16_Sulfa2	TRM32_Sulfa2	TRM64_Sulfa2	TRM128_Sulfa2
4	TRM1_Sulfa4	TRM2_Sulfa4	TRM4_Sulfa4	TRM8_Sulfa4	TRM16_Sulfa4	TRM32_Sulfa4	TRM64_Sulfa4	TRM128_Sulfa4
8	TRM1_Sulfa8	TRM2_Sulfa8	TRM4_Sulfa8	TRM8_Sulfa8	TRM16_Sulfa8	TRM32_Sulfa8	TRM64_Sulfa8	TRM128_Sulfa8
16	TRM1_Sulfa16	TRM2_Sulfa16	TRM4_Sulfa16	TRM8_Sulfa16	TRM16_Sulfa16	TRM32_Sulfa16	TRM64_Sulfa16	TRM128_Sulfa16
32	TRM1_Sulfa32	TRM2_Sulfa32	TRM4_Sulfa32	TRM8_Sulfa32	TRM16_Sulfa32	TRM32_Sulfa32	TRM64_Sulfa32	TRM128_Sulfa32
64	TRM1_Sulfa64	TRM2_Sulfa64	TRM4_Sulfa64	TRM8_Sulfa64	TRM16_Sulfa64	TRM32_Sulfa64	TRM64_Sulfa64	TRM128_Sulfa64

[0133] 将此文库的大约3亿个细胞平板接种在64个板的每一个板上。还将对照菌株平板接种在64个板中的每一个板上,除了不含伴侣蛋白基因的质粒外,所述对照菌株携带与文库完全相同的质粒。将板的整个集合在30℃下温育1周。在对板进行检查时,仅板的子集(在表C中粗体部分)含有可见菌落。与第0天相比,其它要么没有生长变化,要么不存在具有稳

健的菌落的细菌菌苔。

[0134] 表C

[0135]

磺胺甲噁唑(μg/ml)	甲氧苄啶(μg/ml)							
	1	2	4	8	16	32	64	128
0	TRM1_Sulfa0	TRM2_Sulfa0	TRM4_Sulfa0	TRM8_Sulfa0	TRM16_Sulfa0	TRM32_Sulfa0	TRM64_Sulfa0	TRM128_Sulfa0
1	TRM1_Sulfa1	TRM2_Sulfa1	TRM4_Sulfa1	TRM8_Sulfa1	TRM16_Sulfa1	TRM32_Sulfa1	TRM64_Sulfa1	TRM128_Sulfa1
2	TRM1_Sulfa2	TRM2_Sulfa2	TRM4_Sulfa2	TRM8_Sulfa2	TRM16_Sulfa2	TRM32_Sulfa2	TRM64_Sulfa2	TRM128_Sulfa2
4	TRM1_Sulfa4	TRM2_Sulfa4	TRM4_Sulfa4	TRM8_Sulfa4	TRM16_Sulfa4	TRM32_Sulfa4	TRM64_Sulfa4	TRM128_Sulfa4
8	TRM1_Sulfa8	TRM2_Sulfa8	TRM4_Sulfa8	TRM8_Sulfa8	TRM16_Sulfa8	TRM32_Sulfa8	TRM64_Sulfa8	TRM128_Sulfa8
16	TRM1_Sulfa16	TRM2_Sulfa16	TRM4_Sulfa16	TRM8_Sulfa16	TRM16_Sulfa16	TRM32_Sulfa16	TRM64_Sulfa16	TRM128_Sulfa16
32	TRM1_Sulfa32	TRM2_Sulfa32	TRM4_Sulfa32	TRM8_Sulfa32	TRM16_Sulfa32	TRM32_Sulfa32	TRM64_Sulfa32	TRM128_Sulfa32
64	TRM1_Sulfa64	TRM2_Sulfa64	TRM4_Sulfa64	TRM8_Sulfa64	TRM16_Sulfa64	TRM32_Sulfa64	TRM64_Sulfa64	TRM128_Sulfa64

[0136] 图5示出了在温育1周之后具有菌落的板的实例。

[0137] 通过固相测定分析了9个板中的八个板。

[0138] 第1轮固相测定

[0139] 硝酸纤维素膜(思拓凡公司(Cytiva)Protran BA 83目录号10401316)在琼脂的表面上分层。使用缝合针将乳胶漆施加到琼脂,以用于将琼脂板与由硝酸纤维素膜在处理后得到的图像进行定位对准。这是通过将针浸入到颜色鲜艳的乳胶漆中,然后使用覆盖漆的针尖端穿过琼脂刺穿膜,从而在琼脂中留下一些漆并且在膜上留下刺穿标记来完成的。然后将膜从琼脂表面轻轻提起,并且置于含有结合细胞的一侧面朝上的Whatman纸(思拓凡公司目录号1001-090)上。在所有随后步骤中,膜始终以细胞侧面朝上放置。用Whatman纸从硝酸纤维素膜中去除多余的漆。然后将硝化纤维转移到干净的培养皿。将固定溶液的2ml溶液[2.6%(w/v)多聚甲醛,0.04%(w/v)戊二醛,32.25mM Na_3PO_4 pH 7.4]]施加到膜,并且允许所述溶液在室温下温育3分钟。之后,用1x PBS[135mM NaCl,2.7mM KCl,11mM磷酸盐缓冲液pH 7.4]洗涤膜3次。然后将膜在封闭溶液[0.1M NaHCO3 pH 8.6、1mM EDTA、5mg/ml牛血清白蛋白(V部分)、1μM生物素]中温育过夜。第二天,将膜转移到新培养皿,并用5ml溶菌酶溶液[密理博西格玛公司(Millipore Sigma)目录号71110-4,在1x免疫测定缓冲液(珀金埃尔默公司(PerkinElmer)目录号AL000F)中稀释10,000倍]]处理10分钟。随后,将膜转移到新培养皿,并且用甲胎蛋白的检测试剂[0.25nM抗甲胎蛋白(艾博抗公司目录号ab130748),0.25nM抗兔抗体-碱性磷酸酶(密理博西格玛公司目录号A2556)]温育过夜。第二天,将膜洗涤6-8次,每次都要沥干多余的流体,并且将膜转移到含有20-50ml 1x洗涤缓冲液(Azure生物系统公司目录号(Azure Biosystems)AC2113)的新培养皿。在最后洗涤之后,将膜用2ml碱性磷酸酶底物(SeraCare KPL PhosphaGLO目录号5430-0055)处理约3分钟。将膜沥干并且置于Whatman纸(思拓凡公司目录号1001-090)上。膜上的刺穿标志用黑色钢笔手动突出显示。然后将膜置于Azure 600成像仪(Azure生物系统公司)的成像室内侧,并且进行用于(1)荧光(激发732/发射832)和(2)化学发光的成像。荧光图像捕获了钢笔标记并且使用Azure生物系统公司图像捕获软件将所述荧光图像与化学发光图像合并。荧光图像含有将化学发光图像与琼脂板对准的标记。合并的图像使用激光打印机打印在清晰的透明薄膜上。

[0140] 所获得的图像的实例在图6中示出。

[0141] 对所分析的8个膜进行的审查显示出,8个膜中仅三个含有阳性命中(下表D中的矩

阵中的粗体字)。含有阳性命中的三个膜源自含有甲氧苄啶/磺胺甲噁唑组合的琼脂板,所述组合来自下文基质中的蓝色方框。这一结果表明,只有在甲氧苄啶/磺胺甲噁唑的最高耐受浓度下出现的菌落含有用于此选择策略的真正优胜者。

[0142] 表D

[0143]

		甲氧苄啶(ug/ml)							
		1	2	4	8	16	32	64	128
磺胺甲噁唑 (ug/ml)	0	TRM1_Sulfa0	TRM2_Sulfa0	TRM4_Sulfa0	TRM8_Sulfa0	TRM16_Sulfa0	TRM32_Sulfa0	TRM64_Sulfa0	TRM128_Sulfa0
	1	TRM1_Sulfa1	TRM2_Sulfa1	TRM4_Sulfa1	TRM8_Sulfa1	TRM16_Sulfa1	TRM32_Sulfa1	TRM64_Sulfa1	TRM128_Sulfa1
	2	TRM1_Sulfa2	TRM2_Sulfa2	TRM4_Sulfa2	TRM8_Sulfa2	TRM16_Sulfa2	TRM32_Sulfa2	TRM64_Sulfa2	**TRM128_Sulfa2**
	4	TRM1_Sulfa4	TRM2_Sulfa4	TRM4_Sulfa4	TRM8_Sulfa4	TRM16_Sulfa4	TRM32_Sulfa4	**TRM64_Sulfa4**	TRM128_Sulfa4
	8	TRM1_Sulfa8	TRM2_Sulfa8	TRM4_Sulfa8	TRM8_Sulfa8	TRM16_Sulfa8	**TRM32_Sulfa8**	TRM64_Sulfa8	TRM128_Sulfa8
	16	TRM1_Sulfa16	TRM2_Sulfa16	TRM4_Sulfa16	TRM8_Sulfa16	TRM16_Sulfa16	TRM32_Sulfa16	TRM64_Sulfa16	TRM128_Sulfa16
	32	TRM1_Sulfa32	TRM2_Sulfa32	TRM4_Sulfa32	TRM8_Sulfa32	TRM16_Sulfa32	TRM32_Sulfa32	TRM64_Sulfa32	TRM128_Sulfa32
	64	TRM1_Sulfa64	TRM2_Sulfa64	TRM4_Sulfa64	TRM8_Sulfa64	TRM16_Sulfa64	TRM32_Sulfa64	TRM64_Sulfa64	TRM128_Sulfa64

[0144] 命中的回收和表征

[0145] 从对应板挑选出50个命中,利用所获得的图像指导命中回收。合并的图像(荧光+化学发光图像)含有命中位置和对准信息两者以从来源琼脂板回收活命中。使用琼脂中的漆标志和透明薄膜上的通过荧光成像捕获到的点,将琼脂与打印的透明薄膜对准。使用无菌环的平端,通过用环轻轻接触琼脂的表面,并且然后随后接触锉切成8个扇区的新诱导板[LB/Kan(10μg/ml)/阿拉伯糖(125μM)/丙酸盐(1mM)]]的表面来回收推定的命中。每个推定的命中给定一个扇区。然后,在每个扇区中,使用无菌环将初始平板接种物划线,以通过散播回收的大肠杆菌来分离单一菌落。将诱导板在30℃下温育2天。然后,使用描述用于第1轮固相命中鉴定的确切程序对板进行处理。

[0146] 第2轮结果

[0147] 测试板相较于对照板的示例图像

[0148] 图7中的第2轮图像显示出,来自8个选中命中的甲胎蛋白信号大于从原初文库、无伴侣蛋白对照菌株或阳性对照菌株中获得的甲胎蛋白信号。

[0149] 为确保菌株纯度和稳定的基因型,从第2轮图像中的每个暗扇区挑出单个,并且通过划线将单菌落在LB卡那霉素(50μg/ml)琼脂板上传代。目的是在没有任何选择性压力(即培养基中存在甲氧苄啶/磺胺甲噁唑的情况下)的情况下,鉴定提高甲胎蛋白表达的稳定菌株。在生长过夜之后,在新板上执行在LB卡那霉素(50μg/ml)琼脂上的重复划线。将在第二轮划线纯化之后良好分离的单一菌落粘附在新的分区的诱导板上,其中每个扇区由单一菌落占据。将板温育2天,并且通过如上文所描述的固相处理所述板。图8所示的经处理的膜的图像显示出,扇区中的22个扇区对甲胎蛋白表达仍呈阳性,其中来自许多扇区的信号超过阳性对照菌株的信号。

[0150] 回收和分析来自22个阳性菌株的质粒。鉴定十二个独特的质粒。将folA基因从这十二个质粒中切除,并且重新环化,以创建缺乏folA但仍含有原始伴侣蛋白的新表达质粒。(图9)

[0151] 将新质粒转化回到新宿主大肠杆菌菌株中,并且将新创建的菌株在新的分区的诱

导板上划线,并且在30℃下温育24小时,以通过固相测定分析甲胎蛋白的表达。

[0152] 固相测定显示出,在琼脂板上,两个新创建的菌株(4号和5号)明显比阴性对照(N,无伴侣蛋白)表达更多的甲胎蛋白。(图10)为了确定这些菌株在液体培养中是否为优胜者,将相同的12个菌株在含有50μg/ml卡那霉素、125μM阿拉伯糖和1mM丙酸盐的液体诱导培养基中培养。在30℃下生长24小时之后,通过使细胞沉淀采集培养物,并且通过自动毛细管蛋白质分析定量可溶性单体甲胎蛋白。甲胎蛋白的相对丰度相对于阴性对照菌株(无伴侣蛋白)中存在的甲胎蛋白丰度归一化。

[0153] 在液体培养物中,六个菌株(1号、2号、8号、9号、11号)明显比阴性对照菌株(无伴侣蛋白)表达更多的甲胎蛋白。液体培养物中的优胜者与琼脂板上鉴定的所述两个菌株(5号和6号)并不相同。这一发现并不罕见。

[0154] (图11)

[0155] 这些数据显示出,folA融合测定,特别是与固相测定串联耦合时,在合并的文库中鉴定更高执行的菌株。

[0156] 可以将上述各个实施例进行组合以提供另外的实施例。在本说明书中引用的和/或在申请数据表中列出的所有美国专利、美国专利申请公开、美国专利申请、外国专利、外国专利申请和非专利公开通过引用整体并入本文。如果需要,可以修改实施例的各方面,以采用各种专利、申请和公开的概念来提供又另外的实施例。

[0157] 鉴于以上的详细说明,可以对实施例做出这些和其它改变。通常,在以下权利要求中,所使用的术语不应被解释为将权利要求限制于说明书和权利要求中所公开的具体实施例,而是应被解释为包含所有可能的实施例连同这种权利要求所要求的等效物的全部范围。因此,权利要求不受本公开的限制。

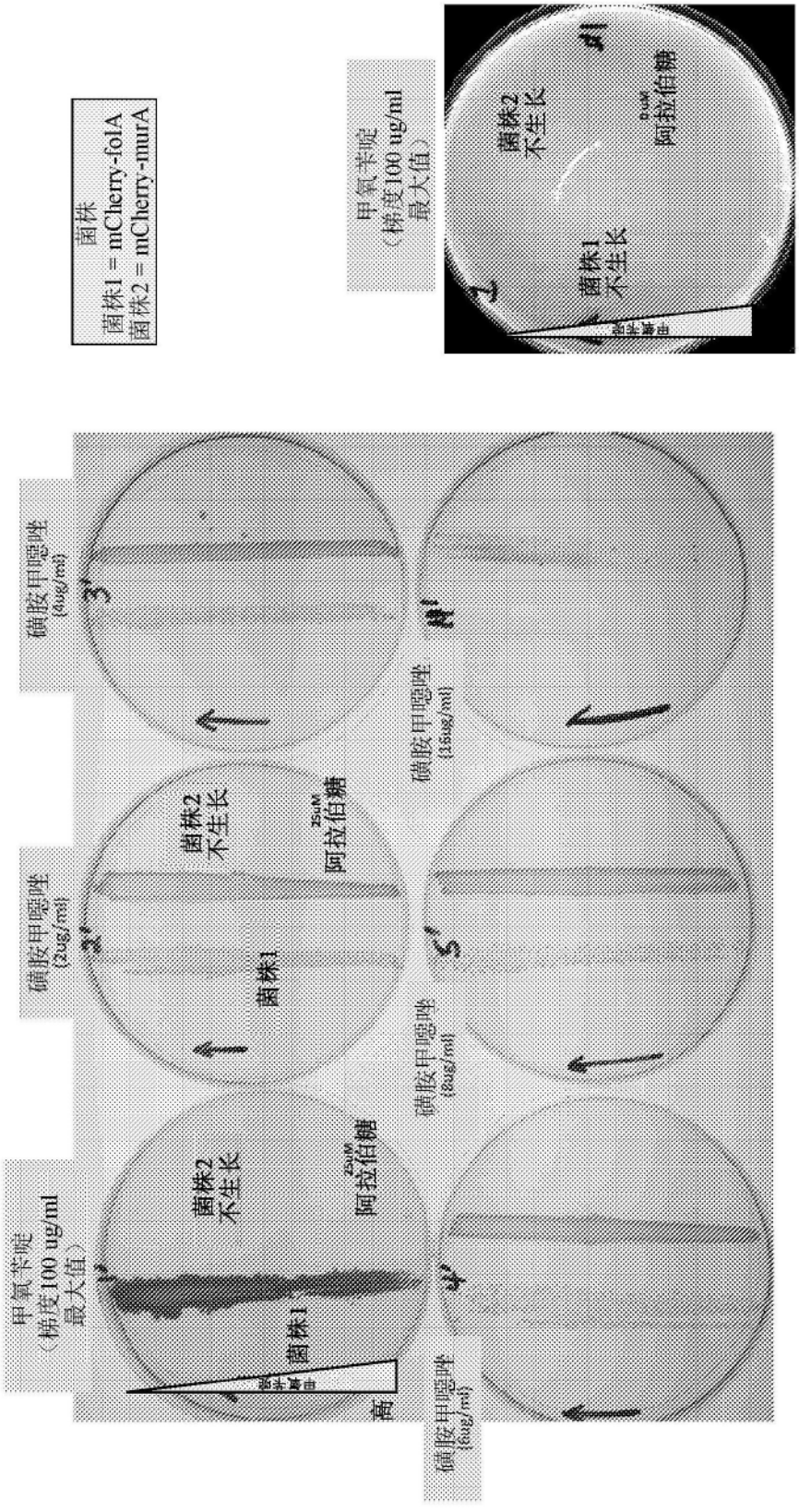

菌株
菌株1 = mCherry-folA
菌株2 = mCherry-murA
甲氧苄啶
(梯度100 ug/ml
最大值)
菌株2
不生长
阿拉伯糖
菌株1
不生长
磺胺甲噁唑
磺胺甲噁唑
磺胺甲噁唑
磺胺甲噁唑
磺胺甲噁唑
菌株2
不生长
阿拉伯糖
菌株1
高

图1

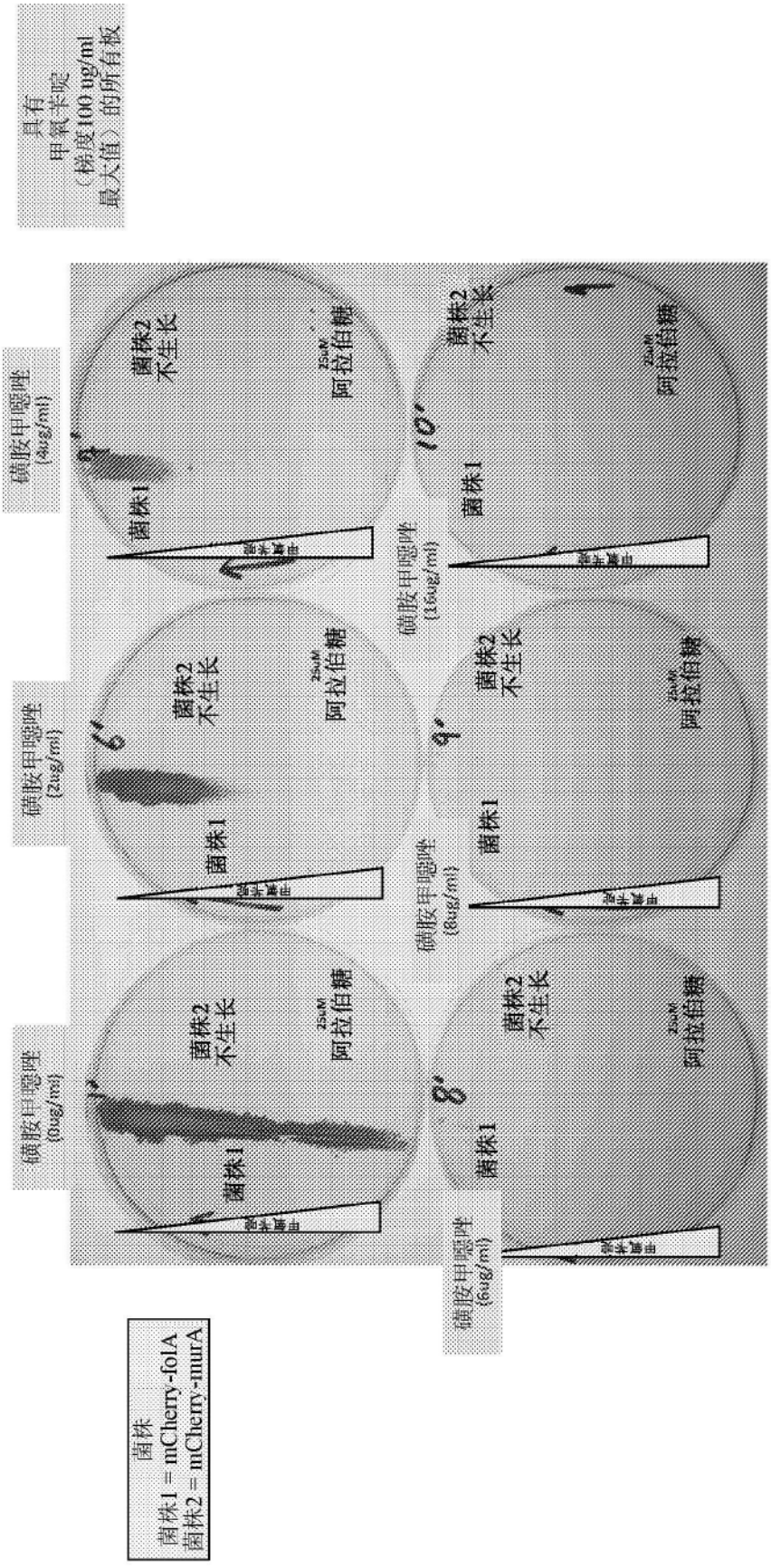

具有
甲氧苄啶
（梯度100 ug/ml
最大值）的所有板
磺胺甲噁唑
{4ug/ml}
磺胺甲噁唑
{2ug/ml}
磺胺甲噁唑
{0ug/ml}
磺胺甲噁唑
{16ug/ml}
磺胺甲噁唑
{8ug/ml}
磺胺甲噁唑
{6ug/ml}
菌株1
菌株2
不生长
阿拉伯糖
25uM
甲氧苄啶
菌株
菌株1 = mCherry-folA
菌株2 = mCherry-murA

图2

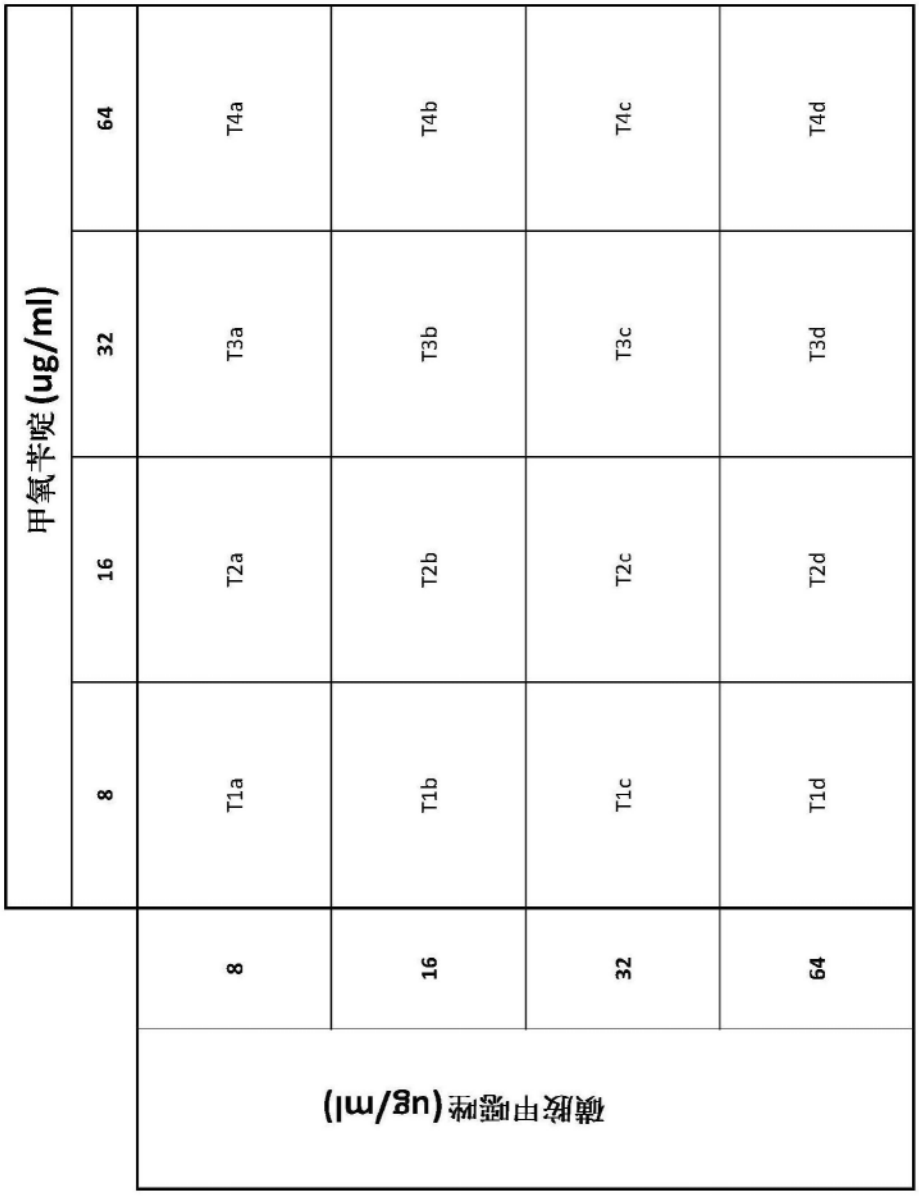

磺胺甲噁唑(ug/ml)
8
16
32
64
甲氧苄啶(ug/ml)
8
16
32
64
T1a
T1b
T1c
T1d
T2a
T2b
T2c
T2d
T3a
T3b
T3c
T3d
T4a
T4b
T4c
T4d

图3

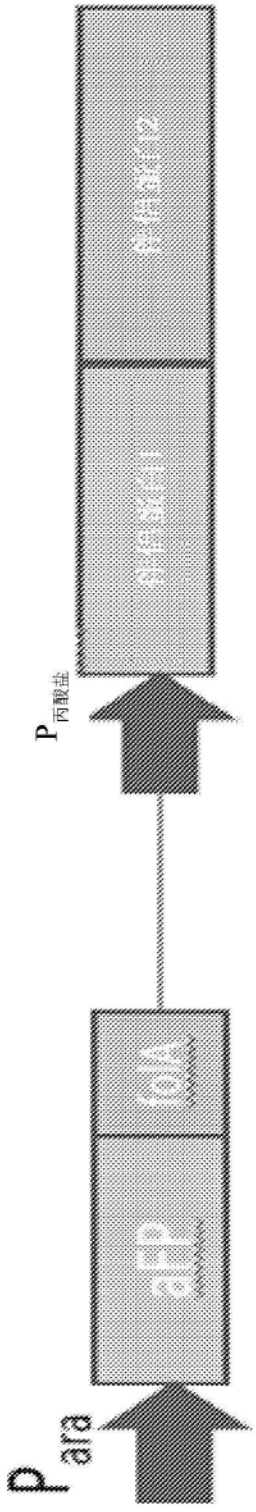


图4

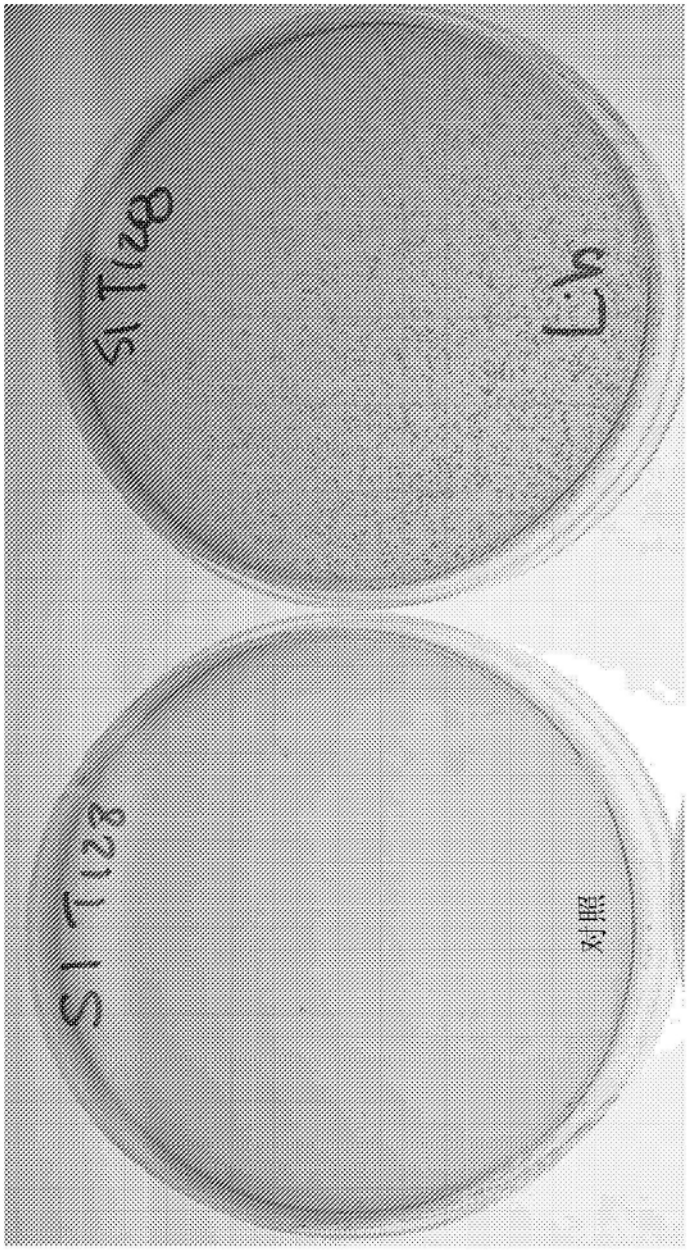


图5

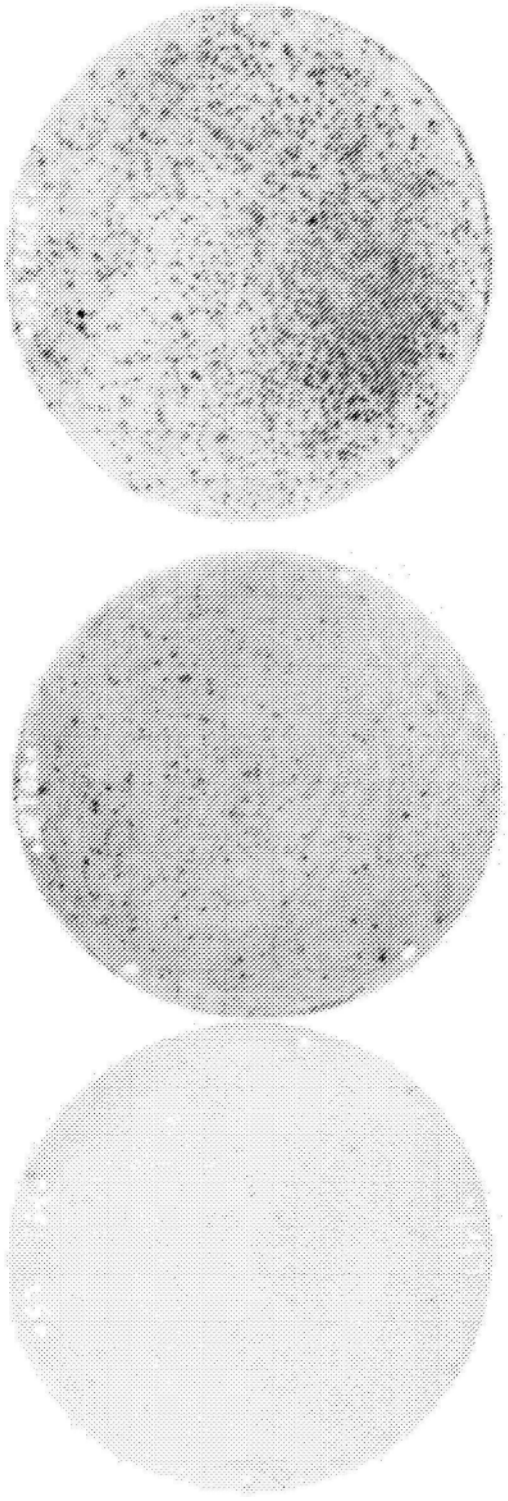


图6

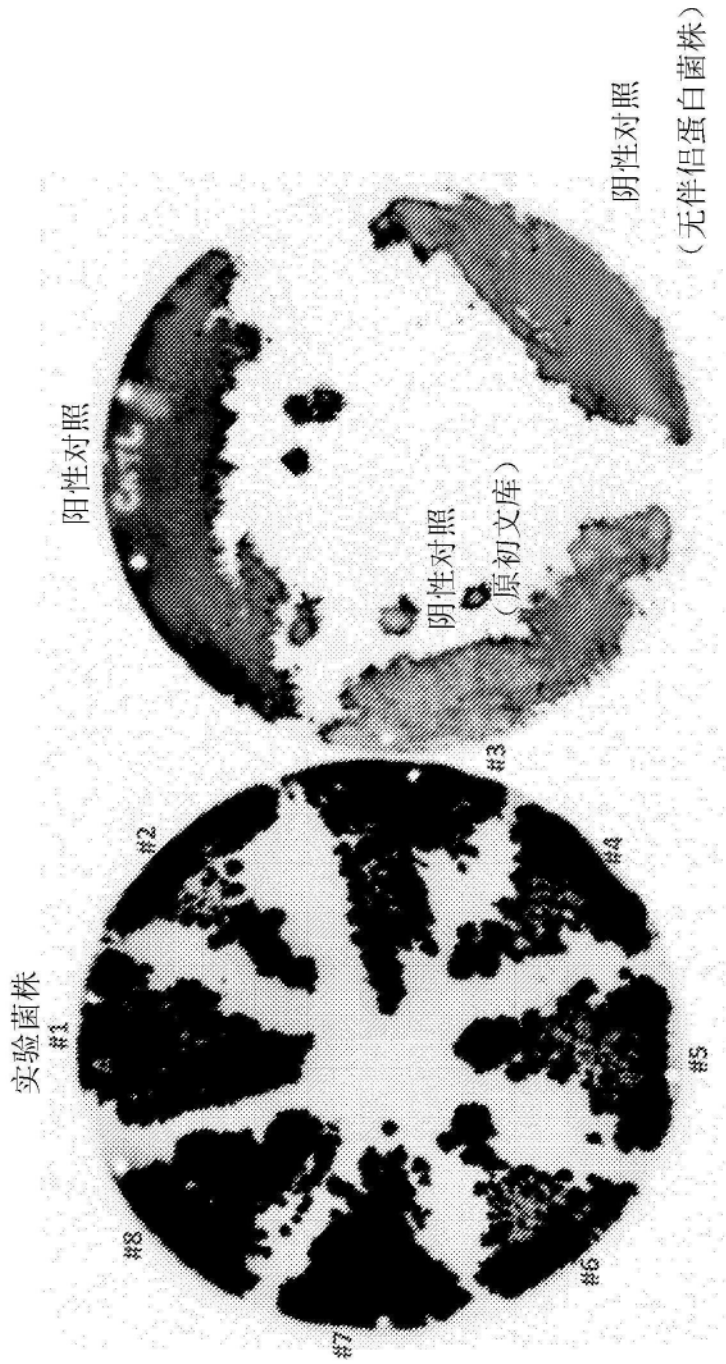


图7

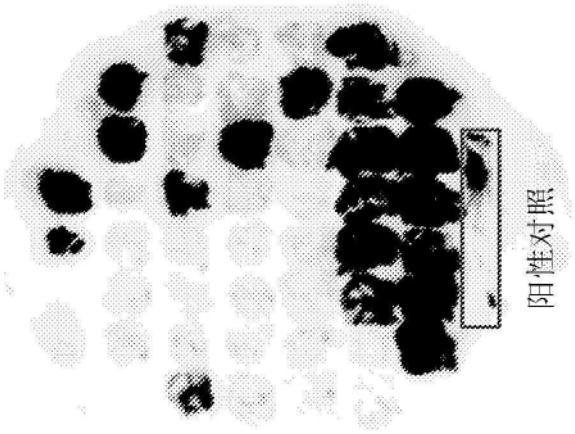


图8

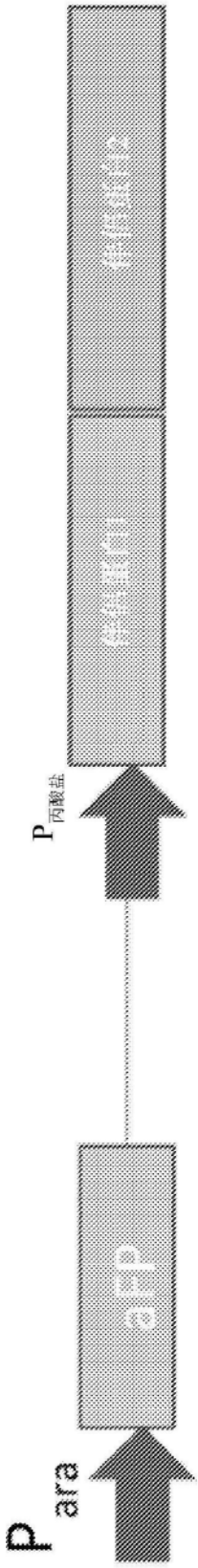


图9

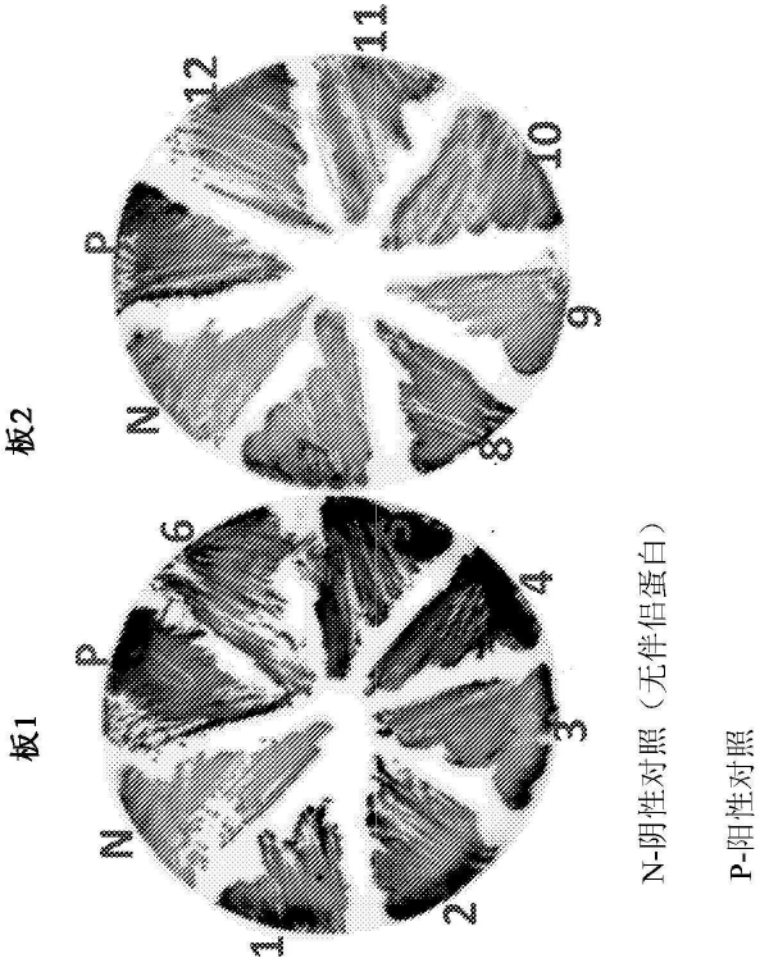


图10

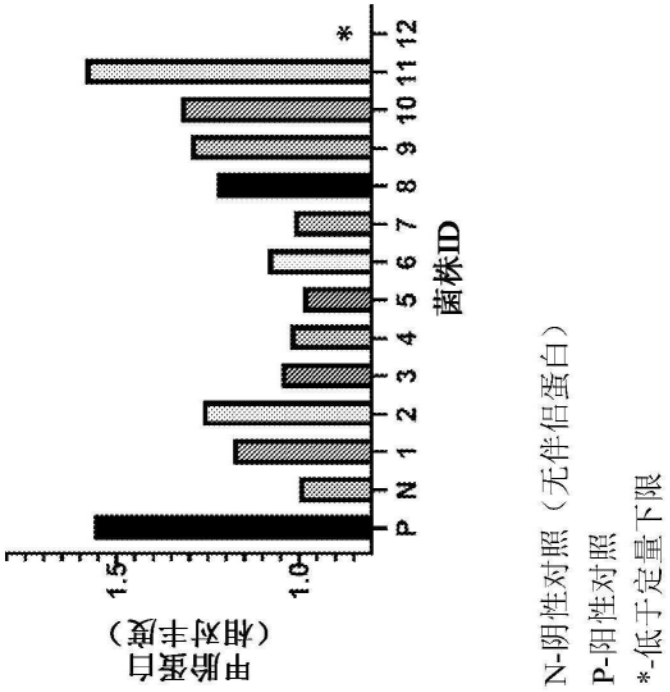

甲胎蛋白
（相对丰度）
1.5
1.0
P N 1 2 3 4 5 6 7 8 9 10 11 12
菌株ID
*
N-阴性对照（无伴侣蛋白）
P-阳性对照
*-低于定量下限

图11