



(12) 发明专利申请

(10) 申请公布号 CN 114450403 A

(43) 申请公布日 2022.05.06

(21) 申请号 202080058809.3

P・鲍尔 D・A・卡林 L・斯通

(22) 申请日 2020.06.19

A・C・塔克

(30) 优先权数据

62/864,875 2019.06.21 US

(74) 专利代理机构 北京嘉和天工知识产权代理
事务所(普通合伙) 11269

62/865,129 2019.06.21 US

专利代理人 缪策 甘玲

(85) PCT国际申请进入国家阶段日

2022.02.18

(51) Int.Cl.

C12N 15/52 (2006.01)

(86) PCT国际申请的申请数据

PCT/US2020/038813 2020.06.19

C12N 15/70 (2006.01)

(87) PCT国际申请的公布数据

W02020/257707 EN 2020.12.24

C12N 9/04 (2006.01)

C12N 9/06 (2006.01)

C12P 7/04 (2006.01)

(71) 申请人 银杏生物制品公司

地址 美国马萨诸塞州

权利要求书5页 说明书55页

申请人 同生运营公司

序列表(电子公布) 附图10页

(72) 发明人 R・杰恩 R・普特曼 K・齐默曼

(54) 发明名称

用于在枫糖尿症 (MSUD) 的治疗中使用的酶
的生物合成

MSUD途径

(57) 摘要

在一些实施方案中,本公开中提供用于治疗
枫糖尿症 (MSUD) 和其他以过多支链氨基酸为特
征的病况的方法和组合物。



酮异己酸

异戊醛

异戊醇

1. 一种宿主细胞,所述宿主细胞包括编码亮氨酸脱氢酶(LeuDH)的异源多核苷酸,其中所述LeuDH酶包括与选自SEQ ID NO:2、SEQ ID NO:4、SEQ ID NO:6、SEQ ID NO:8、SEQ ID NO:10和SEQ ID NO:12的序列至少90%一致的氨基酸序列。

2. 如权利要求1所述的宿主细胞,其中所述LeuDH酶包括与SEQ ID NO:2至少90%一致的氨基酸序列。

3. 如权利要求2所述的宿主细胞,其中所述LeuDH酶包括SEQ ID NO:2。

4. 如权利要求1或2所述的宿主细胞,其中所述LeuDH酶包括:

- a) 与SEQ ID NO:27中的第13位残基对应的残基处的V;
- b) 与SEQ ID NO:27中的第16位残基对应的残基处的W;
- c) 与SEQ ID NO:27中的第42位残基对应的残基处的Q;
- d) 与SEQ ID NO:27中的第43位残基对应的残基处的T、Y、F、E或W;
- e) 与SEQ ID NO:27中的第44位残基对应的残基处的I、H、K或Y;
- f) 与SEQ ID NO:27中的第67位残基对应的残基处的T、E、A、S或K;
- g) 与SEQ ID NO:27中的第71位残基对应的残基处的K;
- h) 与SEQ ID NO:27中的第73位残基对应的残基处的S;
- i) 与SEQ ID NO:27中的第76位残基对应的残基处的R、H、Y、S、K或W;
- j) 与SEQ ID NO:27中的第92位残基对应的残基处的Y;
- k) 与SEQ ID NO:27中的第93位残基对应的残基处的H;
- l) 与SEQ ID NO:27中的第95位残基对应的残基处的G;
- m) 与SEQ ID NO:27中的第100位残基对应的残基处的G;
- n) 与SEQ ID NO:27中的第105位残基对应的残基处的C;
- o) 与SEQ ID NO:27中的第111位残基对应的残基处的G;
- p) 与SEQ ID NO:27中的第113位残基对应的残基处的M;
- q) 与SEQ ID NO:27中的第115位残基对应的残基处的N或V;
- r) 与SEQ ID NO:27中的第116位残基对应的残基处的R、N或W;
- s) 与SEQ ID NO:27中的第120位残基对应的残基处的A;
- t) 与SEQ ID NO:27中的第122位残基对应的残基处的D;
- u) 与SEQ ID NO:27中的第136位残基对应的残基处的E;
- v) 与SEQ ID NO:27中的第140位残基对应的残基处的D;
- w) 与SEQ ID NO:27中的第141位残基对应的残基处的M;
- x) 与SEQ ID NO:27中的第160位残基对应的残基处的S;
- y) 与SEQ ID NO:27中的第185位残基对应的残基处的F;
- z) 与SEQ ID NO:27中的第196位残基对应的残基处的N;
- aa) 与SEQ ID NO:27中的第228位残基对应的残基处的Y;
- bb) 与SEQ ID NO:27中的第248位残基对应的残基处的M;
- cc) 与SEQ ID NO:27中的第256位残基对应的残基处的C;
- dd) 与SEQ ID NO:27中的第293位残基对应的残基处的Q或C;
- ee) 与SEQ ID NO:27中的第296位残基对应的残基处的K或N;
- ff) 与SEQ ID NO:27中的第297位残基对应的残基处的R、Q或K;

- gg) 与SEQ ID NO:27中的第300位残基对应的残基处的C或D;
- hh) 与SEQ ID NO:27中的第302位残基对应的残基处的T或S;
- ii) 与SEQ ID NO:27中的第305位残基对应的残基处的C;
- jj) 与SEQ ID NO:27中的第319位残基对应的残基处的F; 和/或
- kk) 与SEQ ID NO:27中的第330位残基对应的残基处的M。

5. 如权利要求4所述的宿主细胞,其中所述LeuDH酶包括(a) - (kk) 中的全部。

6. 一种宿主细胞,所述宿主细胞包括编码亮氨酸脱氢酶(LeuDH)的异源多核苷酸,其中相对于SEQ ID NO:27,所述LeuDH酶包括氨基酸残基:42、43、44、67、71、76、78、113、115、116、136、293、296、297和/或300处的氨基酸置换。

7. 如权利要求6所述的宿主细胞,其中所述LeuDH酶包括:

- a) 第42位残基处的A、Q或T;
- b) 第43位残基处的E、F、T、W或Y;
- c) 第44位残基处的H、I、K或Y;
- d) 第67位残基处的A、E、K、Q、S或T;
- e) 第71位残基处的C、D、H、K、M或T;
- f) 第76位残基处的E、F、H、I、K、M、R、S、T、W或Y;
- g) 第78位残基处的C、F、H、K、Q、V或Y;
- h) 第113位残基处的F、M、Q、V、W或Y;
- i) 第115位残基处的N、Q、S、T或V;
- j) 第116位残基处的A、L、M、N、R、S、V或W;
- k) 第136位残基处的E、F、L、R、S或Y;
- l) 第293位残基处的A、C、Q、S或T;
- m) 第296位残基处的A、C、E、I、K、L、N、S或T;
- n) 第297位残基处的C、D、E、F、H、K、L、M、N、Q、R、T、W或Y; 和/或
- o) 第300位残基处的A、C、D、F、H、K、M、N、Q、R、S、T、W或Y。

8. 一种非天然存在的LeuDH酶,其中相对于SEQ ID NO:27,所述LeuDH酶包括氨基酸残基:42、43、44、67、71、76、78、113、115、116、136、293、296、297和/或300处的氨基酸置换。

9. 如权利要求8所述的非天然存在的LeuDH酶,其中所述LeuDH酶包括:

- a) 第42位残基处的A、Q或T;
- b) 第43位残基处的E、F、T、W或Y;
- c) 第44位残基处的H、I、K或Y;
- d) 第67位残基处的A、E、K、Q、S或T;
- e) 第71位残基处的C、D、H、K、M或T;
- f) 第76位残基处的E、F、H、I、K、M、R、S、T、W或Y;
- g) 第78位残基处的C、F、H、K、Q、V或Y;
- h) 第113位残基处的F、M、Q、V、W或Y;
- i) 第115位残基处的N、Q、S、T或V;
- j) 第116位残基处的A、L、M、N、R、S、V或W;
- k) 第136位残基处的E、F、L、R、S或Y;

- 1) 第293位残基处的A、C、Q、S或T；
- m) 第296位残基处的A、C、E、I、K、L、N、S或T；
- n) 第297位残基处的C、D、E、F、H、K、L、M、N、Q、R、T、W或Y；和/或
- o) 第300位残基处的A、C、D、F、H、K、M、N、Q、R、S、T、W或Y。

10. 一种宿主细胞，所述宿主细胞包括编码支链 α -酮酸脱羧酶 (KivD) 的异源多核苷酸，其中所述KivD酶包括与选自 SEQ ID NO:14、SEQ ID NO:16 和 SEQ ID NO:18 的序列至少 90% 一致的氨基酸序列。

11. 如权利要求10所述的宿主细胞，其中所述KivD酶包括与 SEQ ID NO:18 至少 90% 一致的氨基酸序列。

12. 如权利要求11所述的宿主细胞，其中所述KivD酶包括 SEQ ID NO:18。

13. 如权利要求10或11所述的宿主细胞，其中所述KivD酶包括：

- a) 与 SEQ ID NO:29 中的第 33 位残基对应的残基处的 Y；
- b) 与 SEQ ID NO:29 中的第 44 位残基对应的残基处的 Q；
- c) 与 SEQ ID NO:29 中的第 117 位残基对应的残基处的 M；
- d) 与 SEQ ID NO:29 中的第 129 位残基对应的残基处的 I；
- e) 与 SEQ ID NO:29 中的第 185 位残基对应的残基处的 W；
- f) 与 SEQ ID NO:29 中的第 190 位残基对应的残基处的 I；
- g) 与 SEQ ID NO:29 中的第 225 位残基对应的残基处的 I；
- h) 与 SEQ ID NO:29 中的第 227 位残基对应的残基处的 Y；
- i) 与 SEQ ID NO:29 中的第 311 位残基对应的残基处的 L；
- j) 与 SEQ ID NO:29 中的第 312 位残基对应的残基处的 G；
- k) 与 SEQ ID NO:29 中的第 313 位残基对应的残基处的 T；
- l) 与 SEQ ID NO:29 中的第 328 位残基对应的残基处的 P；
- m) 与 SEQ ID NO:29 中的第 341 位残基对应的残基处的 W；
- n) 与 SEQ ID NO:29 中的第 345 位残基对应的残基处的 H；
- o) 与 SEQ ID NO:29 中的第 347 位残基对应的残基处的 C；
- p) 与 SEQ ID NO:29 中的第 420 位残基对应的残基处的 R；
- q) 与 SEQ ID NO:29 中的第 494 位残基对应的残基处的 D；
- r) 与 SEQ ID NO:29 中的第 508 位残基对应的残基处的 C；和/或
- s) 与 SEQ ID NO:29 中的第 550 位残基对应的残基处的 F。

14. 如权利要求13所述的宿主细胞，其中所述KivD酶包括 (a) - (s) 中的全部。

15. 一种宿主细胞，所述宿主细胞包括编码醇脱氢酶 (Adh) 的异源多核苷酸，其中所述 Adh 酶包括与选自 SEQ ID NO:20、SEQ ID NO:22 和 SEQ ID NO:24 的序列至少 90% 一致的氨基酸序列。

16. 如权利要求15所述的宿主细胞，其中所述Adh酶包括与 SEQ ID NO:24 至少 90% 一致的氨基酸序列。

17. 如权利要求16所述的宿主细胞，其中所述Adh酶包括 SEQ ID NO:24。

18. 如权利要求15或16所述的宿主细胞，其中所述Adh酶包括：

- a) 与 SEQ ID NO:31 中的第 9 位残基对应的残基处的 P；

- b) 与SEQ ID NO:31中的第16位残基对应的残基处的G;
- c) 与SEQ ID NO:31中的第23位残基对应的残基处的Q;
- d) 与SEQ ID NO:31中的第28位残基对应的残基处的R;
- e) 与SEQ ID NO:31中的第30位残基对应的残基处的A;
- f) 与SEQ ID NO:31中的第93位残基对应的残基处的K;
- g) 与SEQ ID NO:31中的第98位残基对应的残基处的L;
- h) 与SEQ ID NO:31中的第99位残基对应的残基处的R;
- i) 与SEQ ID NO:31中的第114位残基对应的残基处的P;
- j) 与SEQ ID NO:31中的第115位残基对应的残基处的K;
- k) 与SEQ ID NO:31中的第119位残基对应的残基处的Y;
- l) 与SEQ ID NO:31中的第194位残基对应的残基处的Y;
- m) 与SEQ ID NO:31中的第242位残基对应的残基处的P;
- n) 与SEQ ID NO:31中的第249位残基对应的残基处的K;
- o) 与SEQ ID NO:31中的第255位残基对应的残基处的E;
- p) 与SEQ ID NO:31中的第260位残基对应的残基处的D;
- q) 与SEQ ID NO:31中的第269位残基对应的残基处的H;
- r) 与SEQ ID NO:31中的第281位残基对应的残基处的Q;
- s) 与SEQ ID NO:31中的第325位残基对应的残基处的L;
- t) 与SEQ ID NO:31中的第333位残基对应的残基处的M;
- u) 与SEQ ID NO:31中的第334位残基对应的残基处的P; 和/或
- v) 与SEQ ID NO:31中的第348位残基对应的残基处的Q。

19. 如权利要求18所述的宿主细胞,其中所述Adh酶包括(a) - (v) 中的全部。

20. 如权利要求1-7和10-19中任一项所述的宿主细胞,其中所述宿主细胞是植物细胞、藻类细胞、酵母细胞、细菌细胞或动物细胞。

21. 如权利要求20所述的宿主细胞,其中所述宿主细胞是酵母细胞。

22. 如权利要求21所述的宿主细胞,其中所述酵母细胞是酵母属细胞、耶氏酵母属细胞或毕赤酵母属细胞。

23. 如权利要求20所述的宿主细胞,其中所述宿主细胞是细菌细胞。

24. 如权利要求23所述的宿主细胞,其中所述细菌细胞是大肠埃希氏菌细胞或芽孢杆菌细胞。

25. 如权利要求1-7和10-24中任一项所述的宿主细胞,其中所述宿主细胞还包括编码支链氨基酸转运系统2载体蛋白(BrnQ)的异源多核苷酸。

26. 如权利要求25所述的宿主细胞,其中所述BrnQ蛋白与SEQ ID NO:35的氨基酸序列至少90%一致。

27. 如权利要求1-7和10-26中任一项所述的宿主细胞,其中所述异源多核苷酸与诱导型启动子可操作地连接。

28. 如权利要求1-7和10-27中任一项所述的宿主细胞,其中所述异源多核苷酸在操纵子中表达。

29. 如权利要求28所述的宿主细胞,其中所述操纵子表达超过一个异源多核苷酸,并且

其中核糖体结合位点存在于每个异源多核苷酸之间。

30. 如权利要求1-7中任一项所述的宿主细胞,其中所述宿主细胞还包括编码KivD酶的异源多核苷酸和/或编码Adh酶的异源多核苷酸。

31. 如权利要求10-14中任一项所述的宿主细胞,其中所述宿主细胞还包括编码LeuDH酶的异源多核苷酸和/或编码Adh酶的异源多核苷酸。

32. 如权利要求15-19中任一项所述的宿主细胞,其中所述宿主细胞还包括编码LeuDH酶的异源多核苷酸和/或编码KivD酶的异源多核苷酸。

33. 如权利要求1-7和10-32中任一项所述的宿主细胞,其中所述宿主细胞能够从亮氨酸产生异戊醇。

34. 如权利要求33所述的宿主细胞,其中相对于对照宿主细胞,所述宿主细胞消耗多至少两倍的亮氨酸,所述对照宿主细胞包括编码包括SEQ ID NO:27的序列的对照LeuDH酶的异源多核苷酸、编码包括SEQ ID NO:29的序列的对照KivD酶的异源多核苷酸、编码包括SEQ ID NO:31的序列的对照Adh酶的异源多核苷酸、以及编码包括SEQ ID NO:35的序列的对照BrnQ蛋白的异源多核苷酸。

35. 一种包括培养权利要求1-7和10-34中任一项所述的宿主细胞的方法。

36. 一种用于从亮氨酸产生异戊醇的方法,所述方法包括培养权利要求1-7和10-34中任一项所述的宿主细胞。

37. 一种非天然存在的核酸,所述非天然存在的核酸包括与选自SEQ ID NO:1、SEQ ID NO:3、SEQ ID NO:5、SEQ ID NO:7、SEQ ID NO:9和SEQ ID NO:11的核酸序列至少90%一致的序列。

38. 一种非天然存在的核酸,所述非天然存在的核酸包括与选自SEQ ID NO:13、SEQ ID NO:15和SEQ ID NO:17的核酸序列至少90%一致的序列。

39. 一种非天然存在的核酸,所述非天然存在的核酸包括与选自SEQ ID NO:19、SEQ ID NO:21和SEQ ID NO:23的核酸序列至少90%一致的序列。

40. 一种非天然存在的核酸,所述非天然存在的核酸编码与选自SEQ ID NO:2、SEQ ID NO:4、SEQ ID NO:6、SEQ ID NO:8、SEQ ID NO:10和SEQ ID NO:12的序列至少90%一致的序列。

41. 一种非天然存在的核酸,所述非天然存在的核酸编码与选自SEQ ID NO:14、SEQ ID NO:16和SEQ ID NO:18的序列至少90%一致的序列。

42. 一种非天然存在的核酸,所述非天然存在的核酸编码与选自SEQ ID NO:20、SEQ ID NO:22和SEQ ID NO:24的序列至少90%一致的序列。

43. 一种包括权利要求37-42中任一项所述的非天然存在的核酸的运载体。

44. 一种包括权利要求37-42中任一项所述的非天然存在的核酸的表达盒。

用于在枫糖尿症 (MSUD) 的治疗中使用的酶的生物合成

[0001] 相关申请的交叉引用

[0002] 本申请要求于2019年6月21日提交的标题为“用于在枫糖尿症 (MSUD) 的治疗中使用的酶的生物合成”的美国临时申请序列号62/865,129和于2019年6月21日提交的标题为“被工程化以治疗涉及亮氨酸、异亮氨酸和/或缬氨酸的分解代谢的紊乱的优化的细菌”的美国临时申请序列号62/864,875(其各自的公开内容通过引用被整体并入本文)的在35 U.S.C§119(e) 下的权益。

[0003] 对经由EFS-WEB作为文本文件提交的序列表的引用

[0004] 本申请含有已经由EFS-Web以ASCII格式提交的序列表，并且特此通过引用整体并入。创建于2020年6月19日的所述ASCII副本名称为G0919.70033W000-SEQ-0MJ.txt，并且大小为1.76兆字节(MB)。

发明领域

[0005] 本公开涉及对于亮氨酸至异戊醇的转化有用的酶、核酸和细胞。

[0006] 背景

[0007] 枫糖尿症 (MSUD) 是由支链 α -酮酸脱氢酶复合物 (BCKDC) 的缺乏引起的代谢紊乱，从而导致支链氨基酸(亮氨酸、异亮氨酸和缬氨酸)及其毒性副产物(酮酸)在血液和尿液中的累积。MSUD的名字来自于受影响个体的尿液的独特甜味(特别是在诊断之前和在急性病期间)。仍然存在对于用于MSUD和其他以过多支链氨基酸为特征的病况的改进的治疗的需要。

[0008] 概述

[0009] 本公开至少部分地基于含有用于例如通过将亮氨酸转化成异戊醇来消耗亮氨酸的酶的经工程化的细胞的生成。例如，这样的经工程化的细胞对于治疗与亮氨酸的累积有关的疾病(如MSUD)是有用的。

[0010] 本公开的方面涉及宿主细胞，宿主细胞包括编码亮氨酸脱氢酶 (LeuDH) 的异源多核苷酸，其中LeuDH酶包括与选自SEQ ID NO:2、SEQ ID NO:4、SEQ ID NO:6、SEQ ID NO:8、SEQ ID NO:10和SEQ ID NO:12的序列至少90%一致的氨基酸序列。在一些实施方案中，LeuDH酶包括与SEQ ID NO:2至少90%一致的氨基酸序列。在一些实施方案中，LeuDH酶包括SEQ ID NO:2。在一些实施方案中，LeuDH酶包括：与SEQ ID NO:27中的第13位残基对应的残基处的V；与SEQ ID NO:27中的第16位残基对应的残基处的W；与SEQ ID NO:27中的第42位残基对应的残基处的Q；与SEQ ID NO:27中的第43位残基对应的残基处的T、Y、F、E或W；与SEQ ID NO:27中的第44位残基对应的残基处的I、H、K或Y；与SEQ ID NO:27中的第67位残基对应的残基处的T、E、A、S或K；与SEQ ID NO:27中的第71位残基对应的残基处的K；与SEQ ID NO:27中的第73位残基对应的残基处的S；与SEQ ID NO:27中的第76位残基对应的残基处的R、H、Y、S、K或W；与SEQ ID NO:27中的第92位残基对应的残基处的Y；与SEQ ID NO:27中的第93位残基对应的残基处的H；与SEQ ID NO:27中的第95位残基对应的残基处的G；与SEQ ID NO:27中的第100位残基对应的残基处的G；与SEQ ID NO:27中的第105位残基对应的残基处

的C;与SEQ ID NO:27中的第111位残基对应的残基处的G;与SEQ ID NO:27中的第113位残基对应的残基处的M;与SEQ ID NO:27中的第115位残基对应的残基处的N或V;与SEQ ID NO:27中的第116位残基对应的残基处的R、N或W;与SEQ ID NO:27中的第120位残基对应的残基处的A;与SEQ ID NO:27中的第122位残基对应的残基处的D;与SEQ ID NO:27中的第136位残基对应的残基处的E;与SEQ ID NO:27中的第140位残基对应的残基处的D;与SEQ ID NO:27中的第141位残基对应的残基处的M;与SEQ ID NO:27中的第160位残基对应的残基处的S;与SEQ ID NO:27中的第185位残基对应的残基处的F;与SEQ ID NO:27中的第196位残基对应的残基处的N;与SEQ ID NO:27中的第228位残基对应的残基处的Y;与SEQ ID NO:27中的第248位残基对应的残基处的M;与SEQ ID NO:27中的第256位残基对应的残基处的C;与SEQ ID NO:27中的第293位残基对应的残基处的Q或C;与SEQ ID NO:27中的第296位残基对应的残基处的K或N;与SEQ ID NO:27中的第297位残基对应的残基处的R、Q或K;与SEQ ID NO:27中的第300位残基对应的残基处的C或D;与SEQ ID NO:27中的第302位残基对应的残基处的T或S;与SEQ ID NO:27中的第305位残基对应的残基处的C;与SEQ ID NO:27中的第319位残基对应的残基处的F;和/或与SEQ ID NO:27中的第330位残基对应的残基处的M。

[0011] 本公开另外的方面涉及宿主细胞,宿主细胞包括编码亮氨酸脱氢酶(LeuDh)的异源多核苷酸,其中LeuDh酶包括:与SEQ ID NO:27中的第13位残基对应的残基处的V;与SEQ ID NO:27中的第16位残基对应的残基处的W;与SEQ ID NO:27中的第42位残基对应的残基处的Q;与SEQ ID NO:27中的第43位残基对应的残基处的T、Y、F、E或W;与SEQ ID NO:27中的第44位残基对应的残基处的I、H、K或Y;与SEQ ID NO:27中的第67位残基对应的残基处的T、E、A、S或K;与SEQ ID NO:27中的第71位残基对应的残基处的K;与SEQ ID NO:27中的第73位残基对应的残基处的S;与SEQ ID NO:27中的第76位残基对应的残基处的R、H、Y、S、K或W;与SEQ ID NO:27中的第92位残基对应的残基处的Y;与SEQ ID NO:27中的第93位残基对应的残基处的H;与SEQ ID NO:27中的第95位残基对应的残基处的G;与SEQ ID NO:27中的第100位残基对应的残基处的G;与SEQ ID NO:27中的第105位残基对应的残基处的C;与SEQ ID NO:27中的第111位残基对应的残基处的G;与SEQ ID NO:27中的第113位残基对应的残基处的M;与SEQ ID NO:27中的第115位残基对应的残基处的N或V;与SEQ ID NO:27中的第116位残基对应的残基处的R、N或W;与SEQ ID NO:27中的第120位残基对应的残基处的A;与SEQ ID NO:27中的第122位残基对应的残基处的D;与SEQ ID NO:27中的第136位残基对应的残基处的E;与SEQ ID NO:27中的第140位残基对应的残基处的D;与SEQ ID NO:27中的第141位残基对应的残基处的M;与SEQ ID NO:27中的第160位残基对应的残基处的S;与SEQ ID NO:27中的第185位残基对应的残基处的F;与SEQ ID NO:27中的第196位残基对应的残基处的N;与SEQ ID NO:27中的第228位残基对应的残基处的Y;与SEQ ID NO:27中的第248位残基对应的残基处的M;与SEQ ID NO:27中的第256位残基对应的残基处的C;与SEQ ID NO:27中的第293位残基对应的残基处的Q或C;与SEQ ID NO:27中的第296位残基对应的残基处的K或N;与SEQ ID NO:27中的第297位残基对应的残基处的R、Q或K;与SEQ ID NO:27中的第300位残基对应的残基处的C或D;与SEQ ID NO:27中的第302位残基对应的残基处的T或S;与SEQ ID NO:27中的第305位残基对应的残基处的C;与SEQ ID NO:27中的第319位残基对应的残基处的F;以及与SEQ ID NO:27中的第330位残基对应的残基处的M。

[0012] 本公开另外的方面涉及宿主细胞,宿主细胞包括编码亮氨酸脱氢酶(LeuD)的异源多核苷酸,其中相对于SEQ ID NO:27,LeuD酶包括氨基酸残基:42、43、44、67、71、76、78、113、115、116、136、293、296、297和/或300处的氨基酸置换。在一些实施方案中,LeuD酶包括:第42位残基处的A、Q或T;第43位残基处的E、F、T、W或Y;第44位残基处的H、I、K或Y;第67位残基处的A、E、K、Q、S或T;第71位残基处的C、D、H、K、M或T;第76位残基处的E、F、H、I、K、M、R、S、T、W或Y;第78位残基处的C、F、H、K、Q、V或Y;第113位残基处的F、M、Q、V、W或Y;第115位残基处的N、Q、S、T或V;第116位残基处的A、L、M、N、R、S、V或W;第136位残基处的E、F、L、R、S或Y;第293位残基处的A、C、Q、S或T;第296位残基处的A、C、E、I、K、L、N、S或T;第297位残基处的C、D、E、F、H、K、L、M、N、Q、R、T、W或Y;和/或第300位残基处的A、C、D、F、H、K、M、N、Q、R、S、T、W或Y。

[0013] 本公开另外的方面涉及非天然存在的LeuD酶,其中相对于SEQ ID NO:27,LeuD酶包括氨基酸残基:42、43、44、67、71、76、78、113、115、116、136、293、296、297和/或300处的氨基酸置换。在一些实施方案中,LeuD酶包括:第42位残基处的A、Q或T;第43位残基处的E、F、T、W或Y;第44位残基处的H、I、K或Y;第67位残基处的A、E、K、Q、S或T;第71位残基处的C、D、H、K、M或T;第76位残基处的E、F、H、I、K、M、R、S、T、W或Y;第78位残基处的C、F、H、K、Q、V或Y;第113位残基处的F、M、Q、V、W或Y;第115位残基处的N、Q、S、T或V;第116位残基处的A、L、M、N、R、S、V或W;第136位残基处的E、F、L、R、S或Y;第293位残基处的A、C、Q、S或T;第296位残基处的A、C、E、I、K、L、N、S或T;第297位残基处的C、D、E、F、H、K、L、M、N、Q、R、T、W或Y;和/或第300位残基处的A、C、D、F、H、K、M、N、Q、R、S、T、W或Y。

[0014] 本公开另外的方面涉及宿主细胞,宿主细胞包括编码支链 α -酮酸脱羧酶(KivD)的异源多核苷酸,其中KivD酶包括与选自SEQ ID NO:14、SEQ ID NO:16和SEQ ID NO:18的序列至少90%一致的氨基酸序列。在一些实施方案中,KivD酶包括与SEQ ID NO:18至少90%一致的氨基酸序列。在一些实施方案中,KivD酶包括SEQ ID NO:18。在一些实施方案中,KivD酶包括:与SEQ ID NO:29中的第33位残基对应的残基处的Y;与SEQ ID NO:29中的第44位残基对应的残基处的Q;与SEQ ID NO:29中的第117位残基对应的残基处的M;与SEQ ID NO:29中的第129位残基对应的残基处的I;与SEQ ID NO:29中的第185位残基对应的残基处的W;与SEQ ID NO:29中的第190位残基对应的残基处的I;与SEQ ID NO:29中的第225位残基对应的残基处的I;与SEQ ID NO:29中的第227位残基对应的残基处的Y;与SEQ ID NO:29中的第311位残基对应的残基处的L;与SEQ ID NO:29中的第312位残基对应的残基处的G;与SEQ ID NO:29中的第313位残基对应的残基处的T;与SEQ ID NO:29中的第328位残基对应的残基处的P;与SEQ ID NO:29中的第341位残基对应的残基处的W;与SEQ ID NO:29中的第345位残基对应的残基处的H;与SEQ ID NO:29中的第347位残基对应的残基处的C;与SEQ ID NO:29中的第420位残基对应的残基处的R;与SEQ ID NO:29中的第494位残基对应的残基处的D;与SEQ ID NO:29中的第508位残基对应的残基处的C;和/或与SEQ ID NO:29中的第550位残基对应的残基处的F。

[0015] 本公开另外的方面涉及宿主细胞,宿主细胞包括编码支链 α -酮酸脱羧酶(KivD)的异源多核苷酸,其中KivD酶包括:与SEQ ID NO:29中的第33位残基对应的残基处的Y;与SEQ ID NO:29中的第44位残基对应的残基处的Q;与SEQ ID NO:29中的第117位残基对应的残基处的M;与SEQ ID NO:29中的第129位残基对应的残基处的I;与SEQ ID NO:29中的第185位残基对应的残基处的W;与SEQ ID NO:29中的第190位残基对应的残基处的I;与SEQ ID NO:

29中的第225位残基对应的残基处的I;与SEQ ID NO:29中的第227位残基对应的残基处的Y;与SEQ ID NO:29中的第311位残基对应的残基处的L;与SEQ ID NO:29中的第312位残基对应的残基处的G;与SEQ ID NO:29中的第313位残基对应的残基处的T;与SEQ ID NO:29中的第328位残基对应的残基处的P;与SEQ ID NO:29中的第341位残基对应的残基处的W;与SEQ ID NO:29中的第345位残基对应的残基处的H;与SEQ ID NO:29中的第347位残基对应的残基处的C;与SEQ ID NO:29中的第420位残基对应的残基处的R;与SEQ ID NO:29中的第494位残基对应的残基处的D;与SEQ ID NO:29中的第508位残基对应的残基处的C;以及与SEQ ID NO:29中的第550位残基对应的残基处的F。

[0016] 本公开另外的方面涉及宿主细胞,宿主细胞包括编码醇脱氢酶(Adh)的异源多核苷酸,其中Adh酶包括与选自SEQ ID NO:20、SEQ ID NO:22和SEQ ID NO:24的序列至少90%一致的氨基酸序列。在一些实施方案中,Adh酶包括与SEQ ID NO:24至少90%一致的氨基酸序列。在一些实施方案中,Adh酶包括SEQ ID NO:24。在一些实施方案中,Adh酶包括:与SEQ ID NO:31中的第9位残基对应的残基处的P;与SEQ ID NO:31中的第16位残基对应的残基处的G;与SEQ ID NO:31中的第23位残基对应的残基处的Q;与SEQ ID NO:31中的第28位残基对应的残基处的R;与SEQ ID NO:31中的第30位残基对应的残基处的A;与SEQ ID NO:31中的第93位残基对应的残基处的K;与SEQ ID NO:31中的第98位残基对应的残基处的L;与SEQ ID NO:31中的第99位残基对应的残基处的R;与SEQ ID NO:31中的第114位残基对应的残基处的P;与SEQ ID NO:31中的第115位残基对应的残基处的K;与SEQ ID NO:31中的第119位残基对应的残基处的Y;与SEQ ID NO:31中的第194位残基对应的残基处的Y;与SEQ ID NO:31中的第242位残基对应的残基处的P;与SEQ ID NO:31中的第249位残基对应的残基处的K;与SEQ ID NO:31中的第255位残基对应的残基处的E;与SEQ ID NO:31中的第260位残基对应的残基处的D;与SEQ ID NO:31中的第269位残基对应的残基处的H;与SEQ ID NO:31中的第281位残基对应的残基处的Q;与SEQ ID NO:31中的第325位残基对应的残基处的L;与SEQ ID NO:31中的第333位残基对应的残基处的M;与SEQ ID NO:31中的第334位残基对应的残基处的P;和/或与SEQ ID NO:31中的第348位残基对应的残基处的Q。

[0017] 本公开另外的方面涉及宿主细胞,宿主细胞包括编码醇脱氢酶(Adh)的异源多核苷酸,其中Adh酶包括:与SEQ ID NO:31中的第9位残基对应的残基处的P;与SEQ ID NO:31中的第16位残基对应的残基处的G;与SEQ ID NO:31中的第23位残基对应的残基处的Q;与SEQ ID NO:31中的第28位残基对应的残基处的R;与SEQ ID NO:31中的第30位残基对应的残基处的A;与SEQ ID NO:31中的第93位残基对应的残基处的K;与SEQ ID NO:31中的第98位残基对应的残基处的L;与SEQ ID NO:31中的第99位残基对应的残基处的R;与SEQ ID NO:31中的第114位残基对应的残基处的P;与SEQ ID NO:31中的第115位残基对应的残基处的K;与SEQ ID NO:31中的第119位残基对应的残基处的Y;与SEQ ID NO:31中的第194位残基对应的残基处的Y;与SEQ ID NO:31中的第242位残基对应的残基处的P;与SEQ ID NO:31中的第249位残基对应的残基处的K;与SEQ ID NO:31中的第255位残基对应的残基处的E;与SEQ ID NO:31中的第260位残基对应的残基处的D;与SEQ ID NO:31中的第269位残基对应的残基处的H;与SEQ ID NO:31中的第281位残基对应的残基处的Q;与SEQ ID NO:31中的第325位残基对应的残基处的L;与SEQ ID NO:31中的第333位残基对应的残基处的M;与SEQ ID NO:31中的第334位残基对应的残基处的P;以及与SEQ ID NO:31中的第348位残基对应

的残基处的Q。

[0018] 在一些实施方案中,宿主细胞是植物细胞、藻类细胞、酵母细胞、细菌细胞或动物细胞。在一些实施方案中,宿主细胞是酵母细胞。在一些实施方案中,酵母细胞是酵母属细胞、耶氏酵母属细胞或毕赤酵母属细胞。在一些实施方案中,宿主细胞是细菌细胞。在一些实施方案中,细菌细胞是大肠埃希氏菌细胞或芽孢杆菌细胞。

[0019] 在一些实施方案中,宿主细胞还包括编码支链氨基酸转运系统2载体蛋白(BrnQ)的异源多核苷酸。在一些实施方案中,BrnQ蛋白与SEQ ID NO:35的氨基酸序列至少90%一致。在一些实施方案中,BrnQ蛋白包括SEQ ID NO:35的氨基酸序列。

[0020] 在一些实施方案中,异源多核苷酸与诱导型启动子可操作地连接。在一些实施方案中,异源多核苷酸在操纵子中表达。在一些实施方案中,操纵子表达超过一个异源多核苷酸,并且核糖体结合位点可以存在于每个异源多核苷酸之间。

[0021] 在一些实施方案中,宿主细胞还包括编码KivD酶的异源多核苷酸和/或编码Adh酶的异源多核苷酸。

[0022] 在一些实施方案中,宿主细胞还包括编码LeuDH酶的异源多核苷酸和/或编码Adh酶的异源多核苷酸。

[0023] 在一些实施方案中,宿主细胞还包括编码LeuDH酶的异源多核苷酸和/或编码KivD酶的异源多核苷酸。

[0024] 在一些实施方案中,宿主细胞能够从亮氨酸产生异戊醇。在一些实施方案中,相对于对照宿主细胞,宿主细胞消耗多至少两倍的亮氨酸,对照宿主细胞包括编码包括SEQ ID NO:27的序列的对照LeuDH酶的异源多核苷酸、编码包括SEQ ID NO:29的序列的对照KivD酶的异源多核苷酸、编码包括SEQ ID NO:31的序列的对照Adh酶的异源多核苷酸、以及编码包括SEQ ID NO:35的序列的对照BrnQ蛋白的异源多核苷酸。

[0025] 本公开另外的方面涉及包括培养本申请中公开的宿主细胞中的任何一种的方法。

[0026] 本公开另外的方面涉及用于从亮氨酸产生异戊醇的方法,方法包括培养本申请中公开的宿主细胞中的任何一种。

[0027] 本公开另外的方面涉及非天然存在的核酸,非天然存在的核酸包括与选自SEQ ID NO:1、SEQ ID NO:3、SEQ ID NO:5、SEQ ID NO:7、SEQ ID NO:9和SEQ ID NO:11的核酸序列至少90%一致的序列。

[0028] 本公开另外的方面涉及非天然存在的核酸,非天然存在的核酸包括与选自SEQ ID NO:13、SEQ ID NO:15和SEQ ID NO:17的核酸序列至少90%一致的序列。

[0029] 本公开另外的方面涉及非天然存在的核酸,非天然存在的核酸包括与选自SEQ ID NO:19、SEQ ID NO:21和SEQ ID NO:23的核酸序列至少90%一致的序列。

[0030] 本公开另外的方面涉及非天然存在的核酸,非天然存在的核酸编码与选自SEQ ID NO:2、SEQ ID NO:4、SEQ ID NO:6、SEQ ID NO:8、SEQ ID NO:10和SEQ ID NO:12的序列至少90%一致的序列。

[0031] 本公开另外的方面涉及非天然存在的核酸,非天然存在的核酸编码与选自SEQ ID NO:14、SEQ ID NO:16和SEQ ID NO:18的序列至少90%一致的序列。

[0032] 本公开另外的方面涉及非天然存在的核酸,非天然存在的核酸编码与选自SEQ ID NO:20、SEQ ID NO:22和SEQ ID NO:24的序列至少90%一致的序列。

[0033] 本公开另外的方面涉及包括本申请中公开的非天然存在的核酸中的任何一种的运载体。

[0034] 本公开另外的方面涉及包括本申请中公开的非天然存在的核酸中的任何一种的表达盒。

[0035] 本发明的限制中的每个可以涵盖本发明的各种实施方案。因此，预期涉及任何一个要素或要素的组合的本发明的限制中的每个都可以包含在本发明的各个方面中。本发明的应用不限于在以下描述中所示或者在附图中所图示说明的构造细节和组分的布置。本发明能够具有其他实施方案并且能够以各种方式被实践或进行。

[0036] 附图的简要说明

[0037] 附图并非旨在按比例绘制。在附图中，在各种图中图示说明的每个相同或几乎相同的组成部分由类似的数字表示。为清楚起见，并非每个组成部分都可以在每个图中标记。在附图中：

[0038] 图1A-图1C描绘了序列相似性网络。每个点表示序列数据库中可用的单个氨基酸序列。氨基酸序列越紧密相关，点彼此之间越接近。每个序列相似性网络具有相应的簇键，簇键具有关于酶的注释或来源的信息。图1A示出亮氨酸脱氢酶(LeuDH)的序列相似性网络。簇键表示酶的注释。图1B示出酮异戊酸脱羧酶(KivD)的序列相似性网络。每个点的注释表示酶来源的系统发育进化枝。图1C示出醇脱氢酶(Adh)的序列相似性网络。每个点的注释表示酶来源的系统发育进化枝。

[0039] 图2描绘了示出来自LeuDH酶的数据的图表。用生物学复制($n=4$)筛选220种LeuDH酶，以验证酶活性和排序。相对于蜡样芽孢杆菌LeuDH活性报告活性。

[0040] 图3描绘了示出来自LeuDH酶的活性和特异性的比较的数据的图表。针对对Leu、Val和Ile的活性筛选前200种LeuDH酶。相对于蜡样芽孢杆菌LeuDH活性报告LeuDH酶对Leu的活性。将特异性测量为对Leu的活性相对于Val/Leu的比率。在左侧图片中，相对于Leu/Val特异性报告对Leu的酶活性。在右侧图片中，相对于Leu/Ile特异性报告酶活性。将经合理工程化的活性位点变体示出为未填充的圆形。以实心填充的圆形示出有来源的LeuDH酶。还示出阴性对照和阳性对照蜡样芽孢杆菌LeuDH。

[0041] 图4示出来自LeuDH酶的特异性的比较的数据。针对对Leu、Val和Ile的活性筛选前200种LeuDH酶。将特异性测量为对Leu的活性相对于Val/Leu的比率。将经合理工程化的活性位点变体示出为未填充的圆形。以填充的圆形示出有来源的LeuDH酶。示出阴性对照和阳性对照蜡样芽孢杆菌LeuDH。

[0042] 图5描绘了示出来自KivD酶的数据的图。用生物学复制($n=4$)针对活性筛选55种KivD酶。相对于含有异源表达的金黄色葡萄球菌KivD的裂解物(其活性与裂解物的可测量的背景活性难以区分，并且因此等同于背景)的活性报告活性。

[0043] 图6示出来自Adh酶的数据。用生物学复制($n=4$)筛选55种Adh酶。相对于含有异源表达的酿酒酵母ADH2的裂解物(其活性与裂解物的可测量的背景活性难以区分，并且因此等同于背景)的活性报告活性。

[0044] 图7示出LeuDH酶的选择性的数据。测试了总计21种候选LeuDH酶。每组条形从左至右示出消耗的Leu、消耗的Ile和消耗的Val。

[0045] 图8示出消耗Leu最高的菌株(5941、5942和5943)与原型菌株(1980)之间随着时间

的推移Leu消耗的速率的比较。向基础培养基中添加8mM亮氨酸，并且在厌氧培养之后的0小时、2小时和4小时时间点处采集样品。

[0046] 图9示出用于亮氨酸向异戊醇的转化的MSUD途径。

[0047] 图10示出在Ambr15生物反应器(n=2)中测定的菌株5941的异戊醇途径中间产物的细胞外图谱。误差条反映了重复生物反应器的标准偏差。与“加和”对应的数据表示所示中间产物的总浓度。Leu=亮氨酸，酸=2-氧化异己酸，醛=异戊醛，醇=异戊醇。

[0048] 发明的详细说明

[0049] 在一些方面，本公开提供被工程化用于亮氨酸消耗的支链氨基酸(BCAA)途径的细胞、以及用于亮氨酸消耗的支链氨基酸(BCAA)途径的酶的组合。这些BCAA途径酶包含亮氨酸脱氢酶(LeuDH)、酮异戊酸脱羧酶(KivD)和醇脱氢酶(Adh)。所公开的酶和包括这样的酶的宿主细胞可以用于促进亮氨酸消耗(例如，在患有与BCAA(例如，亮氨酸)累积相关的紊乱(如枫糖尿症(MSUD))的受试者中以及在其他医学环境和工业环境中)。

[0050] 亮氨酸脱氢酶(LeuDH)

[0051] 如本公开中所使用的，“亮氨酸脱氢酶(LeuDH)”指催化支链L-氨基酸(例如，L-亮氨酸、L-缬氨酸、L-异亮氨酸)向其2-氧化类似物的可逆脱氨基的酶。LeuDH酶可以使用L-亮氨酸作为底物。在一些实施方案中，与L-缬氨酸和/或L-异亮氨酸相比，LeuDH展现出对L-亮氨酸的特异性。在一些实施方案中，LeuDH从L-亮氨酸产生酮异己酸(也称为2-氧化异己酸)。

[0052] 在一些实施方案中，宿主细胞包括LeuDH酶和/或编码这样的酶的异源多核苷酸。在一些实施方案中，宿主细胞包括编码下列的异源多核苷酸：包括与SEQ ID NO:2、SEQ ID NO:4、SEQ ID NO:6、SEQ ID NO:8、SEQ ID NO:10、SEQ ID NO:12或SEQ ID NO:257-475中的任何一种至少80%(例如，至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%或100%)一致的氨基酸序列的LeuDH酶、表3或表4中的LeuDH酶、或者本公开中另外描述的LeuDH酶。在一些实施方案中，宿主细胞包括与SEQ ID NO:1、SEQ ID NO:3、SEQ ID NO:5、SEQ ID NO:7、SEQ ID NO:9、SEQ ID NO:11或SEQ ID NO:37-255中的任何一种至少90%(例如，至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%或100%)一致的异源多核苷酸、编码表3或表4中的LeuDH酶或者本公开中另外描述的LeuDH酶的多核苷酸。

[0053] 在一些实施方案中，宿主细胞包括来自蜡样芽孢杆菌的LeuDH。在其他实施方案中，宿主细胞不包括来自蜡样芽孢杆菌的LeuDH。

[0054] 来自蜡样芽孢杆菌的LeuDH可以包括UniProtKB-P0A392的氨基酸序列(SEQ ID NO:27)：

MTLEIFEYLEKYDYEQVVFCQDKESGLKAIIAIHDTTLGPALGGTRMWTDSEAAIEDALRLAKGMY
KNAAAGLNLLGGAKTVIIGDPRKDKSEAMFRALGRYIQGLNNGRYITAEDVGTTVDDMDIIHEETDFVTGISPSFGS
SGNPSPVTAYGVYRGMKAAAKEAFGTDNLEGKVIAVQGVGNVAYHLCKHLHAEGAKLIVTDINKEAVQRAVEEFG
[0055] ASAVEPNEIYGVECDIYAPCALGATVNDETIPQLKAKVIAGSANNQLKEDRHGDIHEMGIVYAPDYVINAGGV
NVADELYGYNRERALKRVESIYDTIAKVIEISKRDGIATYVAADRLAEERIASLKNSRSTYLRNIGHDIISRR
(SEQ ID NO: 27)。

[0056] 在一些实施方案中，SEQ ID NO:27的氨基酸序列由核酸序列：

ATGACCCTTGAGATTTTGAATACTCGAAAAATATGATTATGAGCAGGTCTTTCTGTCAAGACAAG
 GAATCAGGACTGAAAGCGATCATTGCTATCCATGATACTACACTGGGCCAGCCTAGGTGGCACCCGTATGTGG
 ACGTACGACTCGGAAGAACGGCAATTGAGGATGCCCTGAGGTTAGCTAAGGGCATGACGTATAAAACGCGCA
 GCCGGTTGAATCTGGCGGTGCGAAAACCGTATTATCGGGATCCCCCAAAGACAAATCTGAAGCAATGTT
 CGGGCGCTGGCGATACATACAGGGACTAAATGGTCGCTATATCACCGCTGAAGATGTAGGAACCTACCGTGGAT
 GATATGGACATAATTACGAAGAACGGACTTCGTCACGGCATTAGCCCTAGTTTGGTAGCTCCGGAACCCG
 TCTCCGGTTACCGCTATGGCGTGTACCGTGGCATGAAGGCAGCGAAAGAGGCCCTTGGTACAGACAACCTG
 [0057] GAGGGGAAAGTATCGCGGTTCAAGGGTAGGTAATGTGGCGTATCATCTGTGCAAACACCTACATGCCGAGGGC
 GCCAAGCTGATTGTCACGGATATCAACAAAGAACGGTACAGCGTGCAGTCGAAGAATTGGCGCTTCCGCCGTT
 GAGCCGAATGAAATCTACGGCGTGGATGCGATATTACCGCCGTGCTTGGTGCACAGTCACAGATGAA
 ACGATCCCTCAGCTGAAAGCAAAGGTAATTGGGGTTCGGCTAATAACCAGTTAAAAGAACAGACATGGAGAC
 ATAATTACCGAGATGGGTATTGTTATGCACCAGATTATGTAATCAATGCCGGCGCGTTATTACGTCGAGAT
 GAACTGTATGGCTACAACCGCGAACCGCCCTCAAACGTGGAGTCATTTATGACACCATTGCCAAAGTGATC
 GAAATCAGCAAGCGCGATGGAATGCCACTTATGTGGCTGCCGATCGTCTGGCGGAAGAACGCATTGCAAGTCTC
 AAAAATAGCCGTTCCACCTACCTCGCAATGCCATGATATTATAAGTCGGCGTTGA (SEQ ID NO: 28)

[0058] 编码。

[0059] 在一些实施方案中,相对于对照,表达编码LeuDH酶的异源多核苷酸的宿主细胞可以将亮氨酸至酮异己酸的转化增加多0.5倍、1倍、1.5倍、2倍、2.5倍、3倍、3.5倍、4倍、4.5倍、5倍、5.5倍或6倍(例如,多2倍至6倍)。在一些实施方案中,对照是表达编码SEQ ID NO: 27的异源多核苷酸的宿主细胞。在一些实施方案中,对照是大肠埃希氏菌Nissle菌株SYN1980 Δ leuE、Δ ilvC、lacZ:tetR-Ptet-livKHMGF、tetR-Ptet-leuDH(Bc)-kivD-adh2-brnQ-rrnB ter(pSC101)(如在美国专利申请公开号US20170232043中描述的)。

[0060] 在一些实施方案中,相对于缬氨酸,表达编码LeuDH酶的异源多核苷酸的宿主细胞可以展现出高至少0.5倍、1倍、1.5倍、2倍、2.5倍、3倍、3.5倍、4倍、4.5倍、5倍、5.5倍或6倍(例如,高2倍至6倍)的对亮氨酸的活性。在一些实施方案中,相对于异亮氨酸,表达编码LeuDH酶的异源多核苷酸的宿主细胞可以展现出高至少0.5倍、1倍、1.5倍、2倍、2.5倍、3倍、3.5倍、4倍、4.5倍、5倍、5.5倍或6倍(例如,高2倍至6倍)的对亮氨酸的活性。

[0061] 在一些实施方案中,LeuDH包括与SEQ ID NO:27、SEQ ID NO:2、SEQ ID NO:4、SEQ ID NO:6、SEQ ID NO:8、SEQ ID NO:10、SEQ ID NO:12或SEQ ID NO:257-475中的任何一种、SEQ ID NO:1、SEQ ID NO:3、SEQ ID NO:5、SEQ ID NO:7、SEQ ID NO:9、SEQ ID NO:11或SEQ ID NO:37-255中的任何一种、表3或表4中的LeuDH酶或本公开中另外描述的LeuDH酶的氨基酸序列或多核苷酸序列至少5%、至少10%、至少15%、至少20%、至少25%、至少30%、至少35%、至少40%、至少45%、至少50%、至少55%、至少60%、至少65%、至少70%、至少71%、至少72%、至少73%、至少74%、至少75%、至少76%、至少77%、至少78%、至少79%、至少80%、至少81%、至少82%、至少83%、至少84%、至少85%、至少86%、至少87%、至少88%、至少89%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%、或100%一致的序列。

[0062] 在一些实施方案中,这样的LeuDH酶包括:与SEQ ID NO:27中的第13位残基对应的残基处的V;与SEQ ID NO:27中的第16位残基对应的残基处的W;与SEQ ID NO:27中的第42位残基对应的残基处的Q;与SEQ ID NO:27中的第43位残基对应的残基处的T、Y、F、E或W;与SEQ ID NO:27中的第44位残基对应的残基处的I、H、K或Y;与SEQ ID NO:27中的第67位残基

对应的残基处的T、E、A、S或K;与SEQ ID NO:27中的第71位残基对应的残基处的K;与SEQ ID NO:27中的第73位残基对应的残基处的S;与SEQ ID NO:27中的第76位残基对应的残基处的R、H、Y、S、K或W;与SEQ ID NO:27中的第92位残基对应的残基处的Y;与SEQ ID NO:27中的第93位残基对应的残基处的H;与SEQ ID NO:27中的第95位残基对应的残基处的G;与SEQ ID NO:27中的第100位残基对应的残基处的G;与SEQ ID NO:27中的第105位残基对应的残基处的C;与SEQ ID NO:27中的第111位残基对应的残基处的G;与SEQ ID NO:27中的第113位残基对应的残基处的M;与SEQ ID NO:27中的第115位残基对应的残基处的N或V;与SEQ ID NO:27中的第116位残基对应的残基处的R、N或W;与SEQ ID NO:27中的第120位残基对应的残基处的A;与SEQ ID NO:27中的第122位残基对应的残基处的D;与SEQ ID NO:27中的第136位残基对应的残基处的E;与SEQ ID NO:27中的第140位残基对应的残基处的D;与SEQ ID NO:27中的第141位残基对应的残基处的M;与SEQ ID NO:27中的第160位残基对应的残基处的S;与SEQ ID NO:27中的第185位残基对应的残基处的F;与SEQ ID NO:27中的第196位残基对应的残基处的N;与SEQ ID NO:27中的第228位残基对应的残基处的Y;与SEQ ID NO:27中的第248位残基对应的残基处的M;与SEQ ID NO:27中的第256位残基对应的残基处的C;与SEQ ID NO:27中的第293位残基对应的残基处的Q或C;与SEQ ID NO:27中的第296位残基对应的残基处的K或N;与SEQ ID NO:27中的第297位残基对应的残基处的R、Q或K;与SEQ ID NO:27中的第300位残基对应的残基处的C或D;与SEQ ID NO:27中的第302位残基对应的残基处的T或S;与SEQ ID NO:27中的第305位残基对应的残基处的C;与SEQ ID NO:27中的第319位残基对应的残基处的F;和/或与SEQ ID NO:27中的第330位残基对应的残基处的M。

[0063] 在一些实施方案中,LeuDH酶包括:与SEQ ID NO:27中的第13位残基对应的残基处的V;与SEQ ID NO:27中的第16位残基对应的残基处的W;与SEQ ID NO:27中的第42位残基对应的残基处的Q;与SEQ ID NO:27中的第43位残基对应的残基处的T、Y、F、E或W;与SEQ ID NO:27中的第44位残基对应的残基处的I、H、K或Y;与SEQ ID NO:27中的第67位残基对应的残基处的T、E、A、S或K;与SEQ ID NO:27中的第71位残基对应的残基处的K;与SEQ ID NO:27中的第73位残基对应的残基处的S;与SEQ ID NO:27中的第76位残基对应的残基处的R、H、Y、S、K或W;与SEQ ID NO:27中的第92位残基对应的残基处的Y;与SEQ ID NO:27中的第93位残基对应的残基处的H;与SEQ ID NO:27中的第95位残基对应的残基处的G;与SEQ ID NO:27中的第100位残基对应的残基处的G;与SEQ ID NO:27中的第105位残基对应的残基处的C;与SEQ ID NO:27中的第111位残基对应的残基处的G;与SEQ ID NO:27中的第113位残基对应的残基处的M;与SEQ ID NO:27中的第115位残基对应的残基处的N或V;与SEQ ID NO:27中的第116位残基对应的残基处的R、N或W;与SEQ ID NO:27中的第120位残基对应的残基处的A;与SEQ ID NO:27中的第122位残基对应的残基处的D;与SEQ ID NO:27中的第136位残基对应的残基处的E;与SEQ ID NO:27中的第140位残基对应的残基处的D;与SEQ ID NO:27中的第141位残基对应的残基处的M;与SEQ ID NO:27中的第160位残基对应的残基处的S;与SEQ ID NO:27中的第185位残基对应的残基处的F;与SEQ ID NO:27中的第196位残基对应的残基处的N;与SEQ ID NO:27中的第228位残基对应的残基处的Y;与SEQ ID NO:27中的第248位残基对应的残基处的M;与SEQ ID NO:27中的第256位残基对应的残基处的C;与SEQ ID NO:27中的第293位残基对应的残基处的Q或C;与SEQ ID NO:27中的第296位残基对

应的残基处的K或N；与SEQ ID N0:27中的第297位残基对应的残基处的R、Q或K；与SEQ ID N0:27中的第300位残基对应的残基处的C或D；与SEQ ID N0:27中的第302位残基对应的残基处的T或S；与SEQ ID N0:27中的第305位残基对应的残基处的C；与SEQ ID N0:27中的第319位残基对应的残基处的F；以及与SEQ ID N0:27中的第330位残基对应的残基处的M。

[0064] 在一些实施方案中，相对于SEQ ID N0:27、SEQ ID N0:2、SEQ ID N0:4、SEQ ID N0:6、SEQ ID N0:8、SEQ ID N0:10、SEQ ID N0:12或SEQ ID N0:257-475中的任何一种、表3或表4中的LeuDH酶、或本公开中另外描述的LeuDH酶，LeuDH酶包括至少1个、至少2个、至少3个、至少4个、至少5个、至少6个、至少7个、至少8个、至少9个、至少10个、至少11个、至少12个、至少13个、至少、至少15个、至少16个、至少17个、至少18个、至少19个、至少20个、至少21个、至少22个、至少23个、至少24个、至少25个、至少26个、至少27个、至少28个、至少29个、至少30个、至少31个、至少32个、至少33个、至少34个、至少35个、至少36个、至少37个、至少38个、至少39个、至少40个、至少41个、至少42个、至少43个、至少44个、至少45个、至少46个、至少47个、至少48个、至少49个、至少50个、至少60个、至少70个、至少80个、至少90个、或至少100个氨基酸置换、氨基酸缺失、氨基酸插入或氨基酸添加。

[0065] 在一些实施方案中，相对于SEQ ID N0:27，LeuDH酶包括一个或更多个残基处的氨基酸置换。在一些实施方案中，LeuDH酶包括与SEQ ID N0:27中的第42位对应的残基处的氨基酸置换、与SEQ ID N0:27中的第43位对应的残基处的氨基酸置换、与SEQ ID N0:27中的第44位对应的残基处的氨基酸置换、与SEQ ID N0:27中的第67位对应的残基处的氨基酸置换、与SEQ ID N0:27中的第71位对应的残基处的氨基酸置换、与SEQ ID N0:27中的第76位对应的残基处的氨基酸置换、与SEQ ID N0:27中的第78位对应的残基处的氨基酸置换、与SEQ ID N0:27中的第113位对应的残基处的氨基酸置换、与SEQ ID N0:27中的第115位对应的残基处的氨基酸置换、与SEQ ID N0:27中的第136位对应的残基处的氨基酸置换、与SEQ ID N0:27中的第293位对应的残基处的氨基酸置换、与SEQ ID N0:27中的第296位对应的残基处的氨基酸置换、与SEQ ID N0:27中的第297位对应的残基处的氨基酸置换、和/或与SEQ ID N0:27中的第300位对应的残基处的氨基酸置换。在一些实施方案中，LeuDH酶包括：与SEQ ID N0:27中的第42位对应的残基处的A、Q或T；与SEQ ID N0:27中的第43位对应的残基处的E、F、T、W或Y；与SEQ ID N0:27中的第44位对应的残基处的H、I、K或Y；与SEQ ID N0:27中的第67位对应的残基处的A、E、K、Q、S或T；与SEQ ID N0:27中的第71位对应的残基处的C、D、H、K、M或T；与SEQ ID N0:27中的第76位对应的残基处的E、F、H、I、K、M、R、S、T、W或Y；与SEQ ID N0:27中的第78位对应的残基处的C、F、H、K、Q、V或Y；与SEQ ID N0:27中的第113位对应的残基处的F、M、Q、V、W或Y；与SEQ ID N0:27中的第115位对应的残基处的N、Q、S、T或V；与SEQ ID N0:27中的第116位对应的残基处的A、L、M、N、R、S、V或W；与SEQ ID N0:27中的第136位对应的残基处的E、F、L、R、S或Y；与SEQ ID N0:27中的第293位对应的残基处的A、C、Q、S或T；与SEQ ID N0:27中的第296位对应的残基处的A、C、E、I、K、L、N、S或T；与SEQ ID N0:27中的第297位对应的残基处的C、D、E、F、H、K、L、M、N、Q、R、T、W或Y；和/或与SEQ ID N0:27中的第300位对应的残基处的A、C、D、F、H、K、M、N、Q、R、S、T、W或Y。

[0066] 在一些实施方案中，相对于SEQ ID N0:27，LeuDH酶包括氨基酸残基：42、43、44、67、71、76、78、113、115、116、136、293、296、297和/或300处的氨基酸置换。在一些实施方案

中,LeuDH酶包括第42位残基处的A、Q或T;第43位残基处的E、F、T、W或Y;第44位残基处的H、I、K或Y;第67位残基处的A、E、K、Q、S或T;第71位残基处的C、D、H、K、M或T;第76位残基处的E、F、H、I、K、M、R、S、T、W或Y;第78位残基处的C、F、H、K、Q、V或Y;第113位残基处的F、M、Q、V、W或Y;第115位残基处的N、Q、S、T或V;第116位残基处的A、L、M、N、R、S、V或W;第136位残基处的E、F、L、R、S或Y;第293位残基处的A、C、Q、S或T;第296位残基处的A、C、E、I、K、L、N、S或T;第297位残基处的C、D、E、F、H、K、L、M、N、Q、R、T、W或Y;和/或第300位残基处的A、C、D、F、H、K、M、N、Q、R、S、T、W或Y。

[0067] 酮异戊酸脱羧酶 (KivD)

[0068] 如本公开中所使用的,“酮异戊酸脱羧酶 (KivD)”指催化来源于氨基酸转胺反应的 α -酮酸向醛的脱羧的酶。KivD可以使用酮异己酸作为底物。在一些实施方案中,KivD从酮异己酸产生异戊醛。

[0069] 在一些实施方案中,宿主细胞包括KivD酶和/或编码这样的酶的异源多核苷酸。在一些实施方案中,宿主细胞包括编码KivD酶的异源多核苷酸,KivD酶包括与SEQ ID NO:14、SEQ ID NO:16、SEQ ID NO:18或SEQ ID NO:533-588中的任何一种、表3或表5中的KivD酶、或本公开中另外描述的KivD酶至少80% (例如,至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%、或100%)一致的氨基酸序列。在一些实施方案中,宿主细胞包括与SEQ ID NO:13、SEQ ID NO:15、SEQ ID NO:17或SEQ ID NO:477-532中的任何一种、编码表3或表5中的KivD酶的多核苷酸、或编码本公开中另外描述的KivD酶的多核苷酸至少90% (例如,至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%、或100%)一致的异源多核苷酸。

[0070] 在一些实施方案中,宿主细胞包括来自乳酸乳球菌的KivD。在其他实施方案中,宿主细胞不包括来自乳酸乳球菌的KivD。

[0071] 来自乳酸乳球菌的KivD可以包括UniProtKB-Q684J7的氨基酸序列 (SEQ ID NO:29) :

MYTVG DYLLDRLHELGIEE IFGVPGDYNLQFLDQII SHKDMKWVGNANELNASY MADGYARTK KAAAFL
TTFGVGELSAVNGLAGSYAENLPVVEIVGSPTSKVQNEGKFVHHTLADGDFKHF MKMHEPVTAARTLLTAENATV
EIDRVLSALLKERKPVYINLPVDVAAA KAEKPSLPLKKENSTSNTSDQEILNKIQESLKNAKKPIVITGHEIISF
GLEKTVTQFISKTKLPITT LNFGKSSVDEALPSFLGIYNGTLSEPNLKEFVESADFILMLGVKL TDSSSTGAFTHH
[0072] LNENKMISLNIDE GKIFNERIQNFDFESLISSL DLSEIEYKGKYIDKKQEDFVPSNALLSQDRLWQAVENLTQS
NETIVAEQGTSFFGASSI FLKS KSHFIGQPLWGSIGYTFPAALGSQIADKESRHLFIGDGSIQLTVQELGLAIR
EKINPICFIINNDGYTVEREIHGPNQSYNDIPMWNYSKLPESFGATEDRVVSKIVRTENE FVSVMKEAQADPNRM
YWIELLILA KE GAPKVLKKMGKLFAEQNKS (SEQ ID NO: 29)。

[0073] 在一些实施方案中,SEQ ID NO:29的氨基酸序列由核酸序列:

ATGTACACAGTCGGTGATTATCTTTAGACCGACTGCACGAACCTCGGAATCGAGGAAATTTTGGCGTG
 CCCGGGATTATAACTGCAGTCCGGACCAAATAATTCCCATAAGGATATGAAATGGGTAGGCAATGCTAAC
 GAACTGAATGCGTCTTACATGGCCGATGGTTATGCACGGACAAAAAGCGGCAGCCTTCTGACGACTTCGGC
 GTTGGTGAGTTAACGCAGGTGAACGGCCTGGCGGGTCATACGCCAAAATCTACAGTTGCAAATCGTGGC
 TCGCCGACCAGCAAAGTCAGAACGAGGGTAAGTTGTGCATCACACCCCTGCTGACGGAGATTAAACATTTC
 ATGAAAATGCACGAACCTGTAACGGCAGCGCACACTGTTGACTGCGGAGAACGCCACCGTCGAAATTGATCGC
 GTCCTGAGTGCTCTTCTGAAGGAACGTAAACCGGTGTATATCAATCTCCGGTTGACGTGGCGCAGCTAAAGCC
 GAAAAACCGAGTTGCCCTAAAGAAAGAGAATAGCACGCTAACACGTCTGACCAAGAAATTCTGAACAAAATT
 CAGGAATCCCTAAAAATGCAAAAAACCTATCGTCATCACCGTCATGAAATAATTTCATTGGACTGGAGAAA
 ACCGTTACACAGTTACCTCAAAGACGAAACTGCCAATTACCAACCTAAATTGGCAAATCGTCCGTAGACGAA
 GCCCTGCCGAGCTCTGGGGATCTAACGGCACTTAAAGCGAACCGAATTAAAGGAATTGTGGAGAGCGCC
 [0074] GATTTCATTCTCATGCTGGTGTAAAGCTGACAGATTCCAGTACGGCGCGTCACTCATCACCTGAACGAGAAC
 AAAATGATCTCGTTAACATTGATGAAGGAAAAATTAAATGAACGTATTCAAACCTTCGATTTGAATCGCTG
 ATTCTCCCTACTGGACCTCAGCGAGATCGAATAACAAAGGTAAATATATTGATAAAAACAGGAAGACTTGTG
 CCGAGTAACGCACTGTTGTCAGGATCGCCTGTGGCAAGCTGTGGAAAATCTGACCCAGAGTAACGAAACGATT
 GTCGCGGAACAGGGGACCTCTTCTTGGTGCCTCGTCAATCTTTAAAGTCAAAATCACATTATTGCCAA
 CCACTTGGGTAGTATCGGCTACACTTCCCTCGGCACGGTAGTCAGATTGCCATAAAGAGTCGCGTCAC
 CTTTGTTATTGGGATGGCTCGTACAATTGACCGTTAGGTAGGTCTTGTATACCGAAAAATCAAT
 CCGATCTGTTCAATTATCAATAATGACGGCTAACCGTGGAGCGCAAATCCATGGTCCGAATCAGAGCTATAAC
 GATATACCGATGTGGAAATTACAGCAAACCCCCGAGAGCTTGGCGAACAGAAGATAGGGTTGTCTCCAAGATC
 GTGCGTACGGAAAACGAATTGTAAAGTGTAAAGAAGCGAACGGACCCTAATCGAATGTACTGGATTGAA
 CTATTCTGGCAAAAGAAGGGGCCCTAAAGTCTCAAGAAAATGGGAAGTTGTGCGCAACAAAACAAAAGC
 TGA (SEQ ID NO: 30)

[0075] 编码。

[0076] 在一些实施方案中,相对于对照,表达编码KivD酶的异源多核苷酸的宿主细胞可以将酮异己酸向异戊醛的转化增加多0.5倍、1倍、1.5倍、2倍、2.5倍、3倍、3.5倍、4倍、4.5倍、5倍、5.5倍或6倍(例如,多2倍至6倍)。在一些实施方案中,对照是表达编码SEQ ID NO: 29的异源多核苷酸的宿主细胞。在一些实施方案中,对照是大肠埃希氏菌Nissle菌株SYN1980 Δ leuE、Δ ilvC、lacZ:tetR-Ptet-livKHMGF、tetR-Ptet-leuDH(Bc)-kivD-adh2-brnQ-rrnB ter(pSC101)(如在美国专利申请公开号US20170232043中描述的)。

[0077] 在一些实施方案中,KivD酶包括与SEQ ID NO:29、SEQ ID NO:14、SEQ ID NO:16、SEQ ID NO:18或SEQ ID NO:533-588中的任何一种、SEQ ID NO:13、SEQ ID NO:15、SEQ ID NO:17或SEQ ID NO:477-532中的任何一种、编码表3或表5中的KivD酶或本公开中另外描述的KivD酶的氨基酸序列或多核苷酸序列至少5%、至少10%、至少15%、至少20%、至少25%、至少30%、至少35%、至少40%、至少45%、至少50%、至少55%、至少60%、至少65%、至少70%、至少71%、至少72%、至少73%、至少74%、至少75%、至少76%、至少77%、至少78%、至少79%、至少80%、至少81%、至少82%、至少83%、至少84%、至少85%、至少86%、至少87%、至少88%、至少89%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%、或100%一致的序列。

[0078] 在一些实施方案中,相对于SEQ ID NO:29、SEQ ID NO:14、SEQ ID NO:16、SEQ ID NO:18或SEQ ID NO:533-588中的任何一种、表3或表5中的KivD酶、或本公开中另外描述的KivD酶,KivD酶包括至少1个、至少2个、至少3个、至少4个、至少5个、至少6个、至少7个、至少

8个、至少9个、至少10个、至少11个、至少12个、至少13个、至少、至少15个、至少16个、至少17个、至少18个、至少19个、至少20个、至少21个、至少22个、至少23个、至少24个、至少25个、至少26个、至少27个、至少28个、至少29个、至少30个、至少31个、至少32个、至少33个、至少34个、至少35个、至少36个、至少37个、至少38个、至少39个、至少40个、至少41个、至少42个、至少43个、至少44个、至少45个、至少46个、至少47个、至少48个、至少49个、至少50个、至少60个、至少70个、至少80个、至少90个、或至少100个氨基酸置换、氨基酸缺失、氨基酸插入或氨基酸添加。

[0079] 在一些实施方案中,KivD酶包括:与SEQ ID N0:29中的第33位残基对应的残基处的Y;与SEQ ID N0:29中的第44位残基对应的残基处的Q;与SEQ ID N0:29中的第117位残基对应的残基处的M;与SEQ ID N0:29中的第129位残基对应的残基处的I;与SEQ ID N0:29中的第185位残基对应的残基处的W;与SEQ ID N0:29中的第190位残基对应的残基处的I;与SEQ ID N0:29中的第225位残基对应的残基处的I;与SEQ ID N0:29中的第227位残基对应的残基处的Y;与SEQ ID N0:29中的第311位残基对应的残基处的L;与SEQ ID N0:29中的第312位残基对应的残基处的G;与SEQ ID N0:29中的第313位残基对应的残基处的T;与SEQ ID N0:29中的第328位残基对应的残基处的P;与SEQ ID N0:29中的第341位残基对应的残基处的W;与SEQ ID N0:29中的第345位残基对应的残基处的H;与SEQ ID N0:29中的第347位残基对应的残基处的C;与SEQ ID N0:29中的第420位残基对应的残基处的R;与SEQ ID N0:29中的第494位残基对应的残基处的D;与SEQ ID N0:29中的第508位残基对应的残基处的C;和/或与SEQ ID N0:29中的第550位残基对应的残基处的F。

[0080] 在一些实施方案中,KivD酶包括:与SEQ ID N0:29中的第33位残基对应的残基处的Y;与SEQ ID N0:29中的第44位残基对应的残基处的Q;与SEQ ID N0:29中的第117位残基对应的残基处的M;与SEQ ID N0:29中的第129位残基对应的残基处的I;与SEQ ID N0:29中的第185位残基对应的残基处的W;与SEQ ID N0:29中的第190位残基对应的残基处的I;与SEQ ID N0:29中的第225位残基对应的残基处的I;与SEQ ID N0:29中的第227位残基对应的残基处的Y;与SEQ ID N0:29中的第311位残基对应的残基处的L;与SEQ ID N0:29中的第312位残基对应的残基处的G;与SEQ ID N0:29中的第313位残基对应的残基处的T;与SEQ ID N0:29中的第328位残基对应的残基处的P;与SEQ ID N0:29中的第341位残基对应的残基处的W;与SEQ ID N0:29中的第345位残基对应的残基处的H;与SEQ ID N0:29中的第347位残基对应的残基处的C;与SEQ ID N0:29中的第420位残基对应的残基处的R;与SEQ ID N0:29中的第494位残基对应的残基处的D;与SEQ ID N0:29中的第508位残基对应的残基处的C;以及与SEQ ID N0:29中的第550位残基对应的残基处的F。

[0081] 醇脱氢酶(Adh)

[0082] 如本公开中所使用的,“醇脱氢酶(Adh)”指催化乙醇向乙醛的转化的酶。Adh可以使用异戊醛作为底物。在一些实施方案中,Adh从异戊醛产生异戊醇。

[0083] 在一些实施方案中,宿主细胞包括Adh酶和/或编码这样的酶的异源多核苷酸。在一些实施方案中,宿主细胞包括编码Adh酶的异源多核苷酸,Adh酶包括与SEQ ID N0:20、SEQ ID N0:22、SEQ ID N0:24或SEQ ID N0:645-700中的任何一种、表3或表6中的Adh酶、或本公开中另外描述的Adh酶至少80% (例如,至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%、或100%)

一致的氨基酸序列。在一些实施方案中,宿主细胞包括与SEQ ID NO:19、SEQ ID NO:21、SEQ ID NO:23或SEQ ID NO:589-644中的任何一种、编码表3或表6中的Adh酶或本公开中另外描述的Adh酶的多核苷酸至少90% (例如,至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%、或100%)一致的异源多核苷酸。

[0084] 在一些实施方案中,宿主细胞包括来自酿酒酵母的Adh。在其他实施方案中,宿主细胞不包括来自酿酒酵母的Adh。

[0085] 来自酿酒酵母的Adh可以包括UniProtKB-P00331的氨基酸序列(SEQ ID NO:31) :

```
MSIPETQKAIIFYESNGKLEHKDIPVPKPKPNELLINVKYSGVCHDLHAWHGDWPLPTKLPLVGGHEG  
AGVVVGMGENVKGKWIGDYAGIKWLNGSCMACEYCELGNESNCPHADLSGYTHDGFSQEYATADAVQAAHIPQGT  
DLAEVAPILCAGITVYKALKSANLRAGHWAAISGAAGGLGSSLAVQYAKAMGYRVLGIDGGPGKEELFTSLGGEVF  
IDFTKEKDIVERSAVVKATNGGAHGIINVSVSEAAIEASTTRYCRANGTVVLVGLPAGAKCSSDVFNHVVKSISIVGS  
YVGNRADTREALDFFARGLKVSPKIVVGLSSLPEIYEKMEKGQIAGRYYVVDTSK (SEQ ID NO: 31).
```

[0086] 在一些实施方案中,SEQ ID NO:31的氨基酸序列由核酸序列:

```
ATGTCGATCCCAGAACCTCAGAAGGCTATTATTTATGAGTCAAACGGCAAACCTCGAACATAAAGACATTCCC  
GTGCCTAAACCGAAACCGAATGAACCTCTGATTAACGTAAAGTACAGCGGAGTCTGCCACACGGATTGCATGCC  
TGGCACGGGATTGGCCGTTACCGACCAAACCTGCCTCTGGTGGGTGCTATGAGGGCGCGGGCGTTGTTGGGT  
ATGGGAGAAAATGTCAAAGGCTGGAAAATCGCGACTATGCAGGGATCAAGTGGCTGAACGGGTCTGTATGGCG  
TGCAGACTGTGAATTAGGTAATGAATCCAACTGCCCACACGCACTGAGTGGTTATACCCATGACGGCAGC  
TTCCAAGAATACGCCACAGCGGATGCCGTGCAGGCAGCTCACATTCCGAAGGAACGTGATCTGCGGAAGTAGCC  
CCAATTCTGTGCGCGGGCATCAGGTATATAAGCTCTCAAAGTGCACATTGCGCCGGTCATTGGCTGCG  
ATTTCGGGTGCCGCGGGCGGGCTGGGATCATTAGCTGTTCACTGCGGAAGGCAATGGGTTATCGAGTTCTGGC  
ATCGACGGCGGGCCCGGTTAAAGAAGAGCTATTACCGCCTCGGCGGTGAGGTCTTCATGATTTACCAAAGAA  
AAAGATATCGTGTCCGAGTCGTGAAAGCAACCAATGGCGCGCTACGGAATTATAATGTGTCTGTATCAGAA  
GCGCGATTGAAGCCAGCACCGTTATTGTCGCGCAACGGCACAGTGGTCTGGTAGGCCTGCCCGCCGGTGC  
AAATGTAGCTCGGACGTGTTCAATCATGTGGTGAAGAGTATTCATTGTTGGATCTTACGTAGGGAACCGTGC  
GATACCGGGAGGCACTGGATTTTGCAAGGGCTTGGTAAAGCCGATCAAAGTCGTGGTCTGTCGTCT  
CTACCTGAAATATATGAGAAAATGGAAAAGGGACAGATGCCGGACGCTACGTCGACACCTCAAAGTGA  
(SEQ ID NO: 32)
```

[0088] 编码。

[0089] 在一些实施方案中,相对于对照,表达编码Adh酶的异源多核苷酸的宿主细胞可以将异戊醛向异戊醇的转化增加多0.5倍、1倍、1.5倍、2倍、2.5倍、3倍、3.5倍、4倍、4.5倍、5倍、5.5倍或6倍(例如,多2倍至6倍)。在一些实施方案中,对照是表达编码SEQ ID NO:31的异源多核苷酸的宿主细胞。在一些实施方案中,对照是表达编码SEQ ID NO:31的异源多核苷酸的宿主细胞。在一些实施方案中,对照是大肠埃希氏菌Nissle菌株SYN1980 Δ leuE、Δ ilvC、lacZ:tetR-Ptet-livKHMGF、tetR-Ptet-leuDH(Bc)-kivD-adh2-brnQ-rrnB ter(pSC101) (如在美国专利申请公开号US20170232043中描述的)。

[0090] 在一些实施方案中,Adh包括与SEQ ID NO:31、SEQ ID NO:20、SEQ ID NO:22、SEQ ID NO:24或SEQ ID NO:645-700中的任何一种、SEQ ID NO:19、SEQ ID NO:21、SEQ ID NO:23或SEQ ID NO:589-644中的任何一种、编码表3或表6中的Adh酶或本公开中另外公开的Adh酶的氨基酸序列或多核苷酸序列至少5%、至少10%、至少15%、至少20%、至少25%、至少30%、至少35%、至少40%、至少45%、至少50%、至少55%、至少60%、至少65%、至少

70%、至少71%、至少72%、至少73%、至少74%、至少75%、至少76%、至少77%、至少78%、至少79%、至少80%、至少81%、至少82%、至少83%、至少84%、至少85%、至少86%、至少87%、至少88%、至少89%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%、或100%一致的序列。

[0091] 在一些实施方案中,相对于SEQ ID N0:31、SEQ ID N0:20、SEQ ID N0:22、SEQ ID N0:24或SEQ ID N0:645-700中的任何一种、表3或表6中的Adh酶、或本公开中另外公开的Adh酶,Adh包括至少1个、至少2个、至少3个、至少4个、至少5个、至少6个、至少7个、至少8个、至少9个、至少10个、至少11个、至少12个、至少13个、至少、至少15个、至少16个、至少17个、至少18个、至少19个、至少20个、至少21个、至少22个、至少23个、至少24个、至少25个、至少26个、至少27个、至少28个、至少29个、至少30个、至少31个、至少32个、至少33个、至少34个、至少35个、至少36个、至少37个、至少38个、至少39个、至少40个、至少41个、至少42个、至少43个、至少44个、至少45个、至少46个、至少47个、至少48个、至少49个、至少50个、至少60个、至少70个、至少80个、至少90个、或至少100个氨基酸置换、氨基酸缺失、氨基酸插入或氨基酸添加。

[0092] 在一些实施方案中,Adh包括与SEQ ID N0:31中的第9位残基对应的残基处的P;与SEQ ID N0:31中的第16位残基对应的残基处的G;与SEQ ID N0:31中的第23位残基对应的残基处的Q;与SEQ ID N0:31中的第28位残基对应的残基处的R;与SEQ ID N0:31中的第30位残基对应的残基处的A;与SEQ ID N0:31中的第93位残基对应的残基处的K;与SEQ ID N0:31中的第98位残基对应的残基处的L;与SEQ ID N0:31中的第99位残基对应的残基处的R;与SEQ ID N0:31中的第114位残基对应的残基处的P;与SEQ ID N0:31中的第115位残基对应的残基处的K;与SEQ ID N0:31中的第119位残基对应的残基处的Y;与SEQ ID N0:31中的第194位残基对应的残基处的Y;与SEQ ID N0:31中的第242位残基对应的残基处的P;与SEQ ID N0:31中的第249位残基对应的残基处的K;与SEQ ID N0:31中的第255位残基对应的残基处的E;与SEQ ID N0:31中的第260位残基对应的残基处的D;与SEQ ID N0:31中的第269位残基对应的残基处的H;与SEQ ID N0:31中的第281位残基对应的残基处的Q;与SEQ ID N0:31中的第325位残基对应的残基处的L;与SEQ ID N0:31中的第333位残基对应的残基处的M;与SEQ ID N0:31中的第334位残基对应的残基处的P;和/或与SEQ ID N0:31中的第348位残基对应的残基处的Q。

[0093] 在一些实施方案中,Adh包括与SEQ ID N0:31中的第9位残基对应的残基处的P;与SEQ ID N0:31中的第16位残基对应的残基处的G;与SEQ ID N0:31中的第23位残基对应的残基处的Q;与SEQ ID N0:31中的第28位残基对应的残基处的R;与SEQ ID N0:31中的第30位残基对应的残基处的A;与SEQ ID N0:31中的第93位残基对应的残基处的K;与SEQ ID N0:31中的第98位残基对应的残基处的L;与SEQ ID N0:31中的第99位残基对应的残基处的R;与SEQ ID N0:31中的第114位残基对应的残基处的P;与SEQ ID N0:31中的第115位残基对应的残基处的K;与SEQ ID N0:31中的第119位残基对应的残基处的Y;与SEQ ID N0:31中的第194位残基对应的残基处的Y;与SEQ ID N0:31中的第242位残基对应的残基处的P;与SEQ ID N0:31中的第249位残基对应的残基处的K;与SEQ ID N0:31中的第255位残基对应的残基处的E;与SEQ ID N0:31中的第260位残基对应的残基处的D;与SEQ ID N0:31中的第269位残基对应的残基处的H;与SEQ ID N0:31中的第281位残基对应的残基处的Q;与SEQ

ID NO:31中的第325位残基对应的残基处的L;与SEQ ID NO:31中的第333位残基对应的残基处的M;与SEQ ID NO:31中的第334位残基对应的残基处的P;以及与SEQ ID NO:31中的第348位残基对应的残基处的Q。

[0094] 支链氨基酸转运系统2载体蛋白 (BrnQ)

[0095] 如本公开中所使用的,“支链氨基酸转运系统2载体蛋白 (BrnQ)”指用于支链氨基酸的LIV-II转运系统的组分。BrnQ可以用于将支链氨基酸(例如,亮氨酸) 转运到细胞(如宿主细胞) 中。

[0096] 在一些实施方案中,宿主细胞包括BrnQ蛋白和/或编码这样的蛋白质的异源多核苷酸。在一些实施方案中,宿主细胞包括编码BrnQ蛋白的异源多核苷酸,BrnQ蛋白包括与本申请中描述的BrnQ蛋白(例如,SEQ ID NO:35) 至少80% (例如,至少80%、至少85%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%、或100%) 一致的氨基酸序列。在一些实施方案中,BrnQ蛋白包括UniProtKB-B7MD59中所示的氨基酸序列。

[0097] UniProtKB-B7MD59具有氨基酸序列:

```
MTHQLRSRDIIALGFMFTALFVGAGNIIFPPMVGLQAGEHVWTAAFGFLITAVGLPVLTVALAKVGGGVDSLST
PIGVAVGVLLATVCYLVAGPLFATPRTATVSFEVGIAPLTGDSALPLFIYSLVYFAIVILVSLYPGKLLDTVGNF
LAPLKIIIALVILSVAIIWPGSISTATEAYQNAAFSNGFVNGLTMDTLGAMVFGIVIVNAARSRGVTEARLLT
[0098] RYTVWAGLMAGVGLTLLYLALFRLGSDSASLVDQSANGAAILHAYVQHTFGGGGSFLLAALIFIACLVTA
VGLTC
ACAEFFAQYVPLSYRTLVLFGFSMVSNLQLIQISVPVLTAIYPPCIALVVLSFTRSWWHNSSRVIA
APP
FISLLFGILDGIKASAFSDILPSWAQRLPLAEQGLAWLMPTVVMVVLAIIWDRAGRQVTSSAH (SEQ ID NO:
35)。
```

[0099] 在一些实施方案中,SEQ ID NO:35由核酸序列:

```
ATGACCCATCAATTAAAGATCGCGCATATCATCGCTCTGGCTTTATGACATTGCGTTGTTCGTCGGCGCAGGT
AACATTATTTCCCTCCAATGGTCGGCTTGCAAGGCAGGCAGACACGTCTGGACTGCGGCATTGGCTTCCCTCATT
ACTGCCGTTGGCCTACCGGTATTAACGGTAGTGGCGTGGCAAAAGTTGGCGCGGTGTTGACAGTCAGCAGC
CCAATTGGTAAAGTCGCTGGCGTACTGCTGGCAACAGTTGTTACCTGGCGGTGGGCCGCTTTGCTACGCCG
CGTACAGCTACCGTTCTTGAAAGTGGCATTGCGCGCTGACGGGTGATTCCCGCCTGCCGCTGTTTACCGC
AGCAGCTGGCTATTCGCTATCGTTATTCTGGTTCGCTCTATCCGGCAAGCTGCTGGATACCGTGGCAACTTC
CTTGCGCCGCTGAAAATTATCGCGCTGGCATCCTGTCTGTTGCCGAATTATCTGGCGGCCGGTTCTATCAGT
ACGGCGACTGAGGCTTATCAAACGCTGCGTTCTAACGGCTTCGTCAACGGCTATCTGACCATGGATA
CGCTGGCGCAATGGTGGTATCGTTATTGTTAACGCGCGCTCTCGTGGCGTTACCGAAGCGCGTCTGCTGACC
CGTTATACCGTCTGGCCTGGCGATGGCGGTGTTGGCTGACTCTGCTGTACCTGGCGCTGTTCCGTCGGGT
TCAGACAGCGCGTCGCTGGTCGATCAGTCTGCAAACGGTGCAGCGATCCTGCATGCTTACGTTACGCA
GCTTACCGTCTGGCGCTGGCGTTGGCGTTACCGAAGCGCGTCTGCTGACC
GGCGCGGGCGGTAGCTTCTGCTGGCGCGTTAATCTTCATCGCCTGCCGCTGACGGCGTTGGCGCTGACCGT
GCTTGTGAGAATTCTCGCCAGTACGTACCGCTCTCTTATCGTACGCTGGTGTATCCTCGCGCTCTCG
ATGGTGGTGTCTAACCTCGGCTTGAGCCAGCTGATTGAGATCTCTGTACCGGTGCTGACCGCCATTATCCGCG
TGTATCGCACTGGTTGTTAACGCTCATGGTGGCATAATTGCTCCCGGTGATTGCTCCGCCGATG
TTTATCAGCCTGCTTTGGTATTCTCGACGGGATCAAGGCATCTGCATCGCAGTATCTTACCGTCCGCGCTGGCG
CAGCGTTACCGCTGGCGAACAGGTCTGGCGTTAATGCCAACAGTGGTGATGGTGGTCTGGCCTATTATC
TGGGATCGTGCAGGCGACGGTCGTCAGGTGACCTCCAGCGCTCACTAA (SEQ ID NO: 36)
```

[0101] 编码。

[0102] 变体

[0103] 本公开也涵盖本公开中描述的酶和蛋白质的变体(例如,LeuDh、KivD或Adh(并且包含核酸序列和氨基酸序列的变体))。变体可以与参考序列共有至少5%、至少10%、至少15%、至少20%、至少25%、至少30%、至少35%、至少40%、至少45%、至少50%、至少55%、至少60%、至少65%、至少70%、至少71%、至少72%、至少73%、至少74%、至少75%、至少76%、至少77%、至少78%、至少79%、至少80%、至少81%、至少82%、至少83%、至少84%、至少85%、至少86%、至少87%、至少88%、至少89%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%、或100%(包含之间的全部值)的序列一致性。

[0104] 除非另外指出,否则本领域已知的术语“序列一致性”指通过序列比较(比对)确定的两个多肽或多核苷酸的序列之间的关系。在一些实施方案中,在序列(例如,LeuDh序列、KivD序列或Adh序列)的整个长度上确定序列一致性。在一些实施方案中,在序列(例如,LeuDh序列、KivD序列或Adh序列)的区(例如,氨基酸或核酸的段,例如,跨越活性位点的序列)上确定序列一致性。

[0105] 一致性还可以指由两个或更多个残基(例如,核酸残基或氨基酸残基)的串之间的匹配数确定的两个序列之间的序列相关性程度。一致性测量具有由特定数学模型或计算机程序(例如,算法)解决的空位比对(如果有的话)的两个或更多个序列之间一致匹配的百分比。

[0106] 可以通过本领域普通技术人员已知的方法中的任何一种容易地计算相关多肽或核酸序列的一致性。可以例如使用Karlin and Altschul Proc.Natl.Acad.Sci.USA 87:2264-68,1990的算法,Karlin and Altschul Proc.Natl.Acad.Sci.USA 90:5873-77,1993中修改的算法来确定两个序列(例如,核酸序列或氨基酸序列)的“百分比一致性”。这样的算法被并入Altschul et al.,J.Mol.Biol.215:403-10,1990的NBLAST[®]程序和XBLAST[®]程序(版本2.0)。例如,可以用XBLAST程序(评分=50,字长=3)进行BLAST[®]蛋白质搜索,以获得与本申请中描述的蛋白质同源的氨基酸序列。在两个序列之间存在空位的情况下,例如,如Altschul et al.,Nucleic Acids Res.25(17):3389-3402,1997中所描述的,可以利用Gapped BLAST[®]。当利用BLAST[®]程序和Gapped BLAST[®]程序时,如本领域普通技术人员将理解的,可以使用各自程序(例如,XBLAST[®]和NBLAST[®])的系统内定参数,或者可以适当地调整参数。

[0107] 例如,可以使用的另外的局部比对技术基于史密斯-沃特曼算法(Smith, T.F.& Waterman, M.S. (1981) “Identification of common molecular subsequences.” J.Mol.Biol.147:195-197)。例如,可以使用的通用全局比对技术是基于动态编程的尼德曼-翁施算法(Needleman, S.B.& Wunsch, C.D. (1970) “A general method applicable to the search for similarities in the amino acid sequences of two proteins.” J.Mol.Biol.48:443-453)。

[0108] 最近,开发了一种快速最优全局序列比对算法(FOGSAA),据称该算法比其他最优全局比对方法(包含尼德曼-翁施算法)更快地产生核酸序列和氨基酸序列的全局比对。在一些实施方案中,通过比对两个氨基酸序列、计算相同氨基酸的数量、并且除以氨基酸序列之一的长度来确定两个多肽的一致性。在一些实施方案中,通过比对两个核苷酸序列并且

计算相同核苷酸的数量并且除以核酸之一的长度来确定两个核酸的一致性。

[0109] 对于多序列比对,可以使用计算机程序(包含Clustal Omega(Sievers et al., Mol Syst Biol.2011 Oct 11;7:539))。

[0110] 在优选实施方案中,当使用Karlin and Altschul Proc.Natl.Acad.Sci.USA 87: 2264-68,1990(如在Karlin and Altschul Proc.Natl.Acad.Sci.USA 90:5873-77,1993中修改的)的算法(例如,BLAST[®]程序、NBLAST[®]程序、XBLAST[®]程序或Gapped BLAST[®]程序,使用各程序的默认参数)确定序列一致性时,发现序列(包含核酸序列或氨基酸序列)(如本申请中公开的和/或权利要求中限定的序列)与参考序列具有特定的百分比一致性。

[0111] 在一些实施方案中,当使用史密斯-沃特曼算法(Smith, T.F.&Waterman, M.S. (1981) "Identification of common molecular subsequences." J.Mol.Biol.147:195-197)或尼德曼-翁施算法(Needleman, S.B.&Wunsch, C.D. (1970) "A general method applicable to the search for similarities in the amino acid sequences of two proteins." J.Mol.Biol.48:443-453)使用默认参数确定序列一致性时,发现序列(包含核酸序列或氨基酸序列)(如本申请中公开的和/或权利要求中限定的序列)与参考序列具有特定的百分比一致性。

[0112] 在一些实施方案中,当使用快速最优全局序列比对算法(FOGSAA)使用默认参数确定序列一致性时,发现序列(包含核酸序列或氨基酸序列)(如本申请中公开的和/或权利要求中限定的序列)与参考序列具有特定的百分比一致性。

[0113] 在一些实施方案中,当使用Clustal Omega(Sievers et al., Mol Syst Biol.2011 Oct 11;7:539)使用默认参数确定序列一致性时,发现序列(包含核酸序列或氨基酸序列)(如本申请中公开的和/或权利要求中限定的序列)与参考序列具有特定的百分比一致性。

[0114] 如本公开中所使用的,当使用本领域已知的氨基酸序列比对工具(如,例如,Clustal Omega或BLAST[®])比对序列X和序列Y且当序列“X”中的残基在序列“Y”中的“Z”的对应位置处时,序列“X”中的残基(如核酸残基或氨基酸残基)被称为对应于不同序列“Y”中的位置或残基(如核酸残基或氨基酸残基)“Z”。

[0115] 如本公开中所使用的,变体序列可以是同源序列。如本公开中所使用的,同源序列是具有一定百分比一致性(例如,至少5%、至少10%、至少15%、至少20%、至少25%、至少30%、至少35%、至少40%、至少45%、至少50%、至少55%、至少60%、至少65%、至少70%、至少71%、至少72%、至少73%、至少74%、至少75%、至少76%、至少77%、至少78%、至少79%、至少80%、至少81%、至少82%、至少83%、至少84%、至少85%、至少86%、至少87%、至少88%、至少89%、至少90%、至少91%、至少92%、至少93%、至少94%、至少95%、至少96%、至少97%、至少98%、至少99%、或100%(包含之间的全部值)百分比一致性)的序列(例如,核酸序列或氨基酸序列)。同源序列包含但不限于旁系同源序列或直系同源序列。旁系同源序列由物种的基因组内的基因的复制产生,而直系同源序列在物种形成事件之后趋异。

[0116] 在一些实施方案中,多肽变体(例如,LeuDh酶变体、KivD酶变体或Adh酶变体)包括与参考多肽(例如,参考LeuDh酶、参考KivD酶或参考Adh酶)共有二级结构(例如,α螺旋、β片层)的域。在一些实施方案中,多肽变体(例如,LeuDh酶变体、KivD酶变体或Adh酶变体)与参

考多肽(例如,参考LeuD酶、参考KivD酶或参考Adh酶)共有三级结构。作为非限制性实例,变体多肽(例如,LeuD酶变体、KivD酶变体或Adh酶变体)与参考多肽相比可以具有低的一级序列一致性(例如,小于80%、小于75%、小于70%、小于65%、小于60%、小于55%、小于50%、小于45%、小于40%、小于35%、小于30%、小于25%、小于20%、小于15%、小于10%、或小于5%的序列一致性),但共有一个或更多个二级结构(例如,包含但不限于环、 α 螺旋或 β 片层),或者具有与参考多肽相同的三级结构。例如,环可以位于 β 片层与 α 螺旋之间、两个 α 螺旋之间、或两个 β 片层之间。同源建模可以用于比较两个或更多个三级结构。

[0117] 任何适合的方法(包含环状变换(Yu and Lutz, Trends Biotechnol. 2011 Jan; 29 (1): 18-25))都可以用于产生这样的变体。在环状变换中,可以环化多肽的线性一级序列(例如,通过连接序列的N末端和C末端),并且可以在不同位置处切断(“断裂”)多肽。因此,如由线性序列比对方法(例如,Clustal Omega或BLAST)所确定的,新多肽的线性一级序列可以具有低的序列一致性(例如,小于80%、小于75%、小于70%、小于65%、小于60%、小于55%、小于50%、小于45%、小于40%、小于35%、小于30%、小于25%、小于20%、小于15%、小于10%、小或小于5%(包含之间的全部值))。然而,两种多肽的拓扑分析可以揭示其三级结构类似。不受特定理论的束缚,通过参考多肽的环状变换创建并且具有与参考多肽的三级结构类似的三级结构的变体多肽可以共有类似的功能特性(例如,酶活性、酶动力学、底物特异性或产物特异性)。在一些情况下,环状变换可以改变二级结构、三级结构或四级结构,并且产生具有不同功能特性(例如,增加或减少的酶活性、不同的底物特异性、或不同的产物特异性)的酶。参见,例如,Yu and Lutz, Trends Biotechnol. 2011 Jan; 29 (1): 18-25。

[0118] 应当理解的是,在已经经历环状变换的蛋白质中,蛋白质的线性氨基酸序列将不同于尚未经历环状变换的参考蛋白质。然而,本领域普通技术人员将能够通过例如比对序列和检测保守基序、和/或通过比较蛋白质的结构或预测结构(例如,通过同源建模)容易地确定已经经历环状变换的蛋白质中的哪些残基对应于尚未经历环状变换的参考蛋白质中的残基。本申请中描述的变体包含本申请中描述的序列的经环状变换的变体。

[0119] 在一些实施方案中,本申请中描述的确定感兴趣的序列与参考序列之间的百分比一致性的算法说明了序列之间的环状变换的存在。可以使用本领域已知的任何方法(包含,例如,RASPODOM(Weiner et al., Bioinformatics. 2005 Apr 1; 21 (7): 932-7))检测环状变换的存在。在一些实施方案中,在计算感兴趣的序列与本申请中描述的序列之间的百分比一致性之前,对环状变换的存在进行校正(例如,重排至少一个序列中的域)。应当理解本申请的权利要求包含在考虑序列的潜在环状变换后计算与参考序列的百分比一致性的序列。

[0120] 本公开也涵盖本申请中公开的重组LeuD酶、KivD酶或Adh酶的功能变体。例如,功能变体可以结合相同底物中的一种或更多种或者产生相同产物中的一种或更多种。可以使用本领域已知的任何方法鉴别功能变体。例如,上文描述的Karlin and Altschul Proc. Natl. Acad. Sci. USA 87:2264-68, 1990的算法可以用于鉴别具有已知功能的同源蛋白质。

[0121] 也可以通过搜索具有功能注释域的多肽鉴别推定的功能变体。数据库(包含Pfam (Sonnhammer et al., Proteins. 1997 Jul; 28 (3): 405-20))可以用于鉴别具有特定域的多肽。

[0122] 同源建模也可以用于鉴别适合突变而不影响功能的氨基酸残基。这样的方法的非

限制性实例可以包含位置特异性评分矩阵 (position-specific scoring matrix) (PSSM) 和能量最小化协议的使用。

[0123] 位置特异性评分矩阵 (PSSM) 使用位置权重矩阵来鉴别共有序列 (例如, 基序)。可以在核酸序列或氨基酸序列上进行PSSM。比对序列, 并且方法考虑在特定位置处观察到的特定残基 (例如, 氨基酸或核苷酸) 的频率和所分析的序列的数量。参见, 例如, Stormo et al., Nucleic Acids Res. 1982 May 11; 10 (9) : 2997-3011。可以计算在给定位置处观察到特定残基的可能性。不受特定理论的束缚, 具有高变异性的序列中的位置可以适合突变 (例如, PSSM评分 ≥ 0) 以产生功能同系物。

[0124] PSSM可以与Rosetta能量函数的计算配对, Rosetta能量函数确定野生型与单点突变体之间的差异。Rosetta能量函数将该差异计算为 ($\Delta \Delta G_{\text{calc}}$)。利用Rosetta函数, 突变的残基与周围原子之间的键合相互作用被用于确定突变是增加还是减小蛋白质稳定性。例如, 然后可以使用Rosetta能量函数分析由PSSM评分 (例如, PSSM评分 ≥ 0) 指定为有利的突变, 以确定突变对蛋白质稳定性的潜在影响。不受特定理论的束缚, 潜在稳定化的突变对于蛋白质工程 (例如, 功能同系物的产生) 是期望的。在一些实施方案中, 潜在稳定化的突变具有小于-0.1 (例如, 小于-0.2、小于-0.3、小于-0.35、小于-0.4、小于-0.45、小于-0.5、小于-0.55、小于-0.6、小于-0.65、小于-0.7、小于-0.75、小于-0.8、小于-0.85、小于-0.9、小于-0.95, 或小于-1.0) Rosetta能量单位 (R.e.u.) 的 $\Delta \Delta G_{\text{calc}}$ 值。参见, 例如, Goldenzweig et al., Mol Cell. 2016 Jul 21; 63 (2) : 337-346. Doi:10.1016/j.molcel.2016.06.012。

[0125] 在一些实施方案中, LeuDh酶、KivD酶或Adh酶编码序列包括在与参考 (例如, LeuDh酶、KivD酶或Adh酶) 编码序列对应的1个、2个、3个、4个、5个、6个、7个、8个、9个、10个、11个、12个、13个、14个、15个、16个、17个、18个、19个、20个、21个、22个、23个、24个、25个、26个、27个、28个、29个、30个、31个、32个、33个、34个、35个、36个、37个、38个、39个、40个、41个、42个、43个、44个、45个、46个、47个、48个、49个、50个、51个、52个、53个、54个、55个、56个、57个、58个、59个、60个、61个、62个、63个、64个、65个、66个、67个、68个、69个、70个、71个、72个、73个、74个、75个、76个、77个、78个、79个、80个、81个、82个、83个、84个、85个、86个、87个、88个、89个、90个、91个、92个、93个、94个、95个、96个、97个、98个、99个、100个或超过100个位置处的突变。在一些实施方案中, 相对于参考 (例如, LeuDh酶、KivD酶或Adh酶) 编码序列, LeuDh酶、KivD酶或Adh酶编码序列在编码序列的1个、2个、3个、4个、5个、6个、7个、8个、9个、10个、11个、12个、13个、14个、15个、16个、17个、18个、19个、20个、21个、22个、23个、24个、25个、26个、27个、28个、29个、30个、31个、32个、33个、34个、35个、36个、37个、38个、39个、40个、41个、42个、43个、44个、45个、46个、47个、48个、49个、50个、51个、52个、53个、54个、55个、56个、57个、58个、59个、60个、61个、62个、63个、64个、65个、66个、67个、68个、69个、70个、71个、72个、73个、74个、75个、76个、77个、78个、79个、80个、81个、82个、83个、84个、85个、86个、87个、88个、89个、90个、91个、92个、93个、94个、95个、96个、97个、98个、99个、100个或更多个密码子中包括突变。如本领域普通技术人员将理解的, 由于遗传码的简并性, 密码子内的突变可以改变由密码子编码的氨基酸或者可以不改变由密码子编码的氨基酸。在一些实施方案中, 相对于参考多肽 (例如, LeuDh酶、KivD酶或Adh酶) 的氨基酸序列, 编码序列中的一个或更多个突变不改变编码序列 (例如, LeuDh酶、KivD酶或Adh酶) 的氨基酸序列。

[0126] 在一些实施方案中,相对于参考多肽(例如,LeuD酶、KivD酶或Adh酶)的氨基酸序列,重组LeuD酶序列、重组KivD酶序列或重组Adh酶序列中的一个或更多个突变改变多肽(例如,LeuD酶、KivD酶或Adh酶)的氨基酸序列。在一些实施方案中,相对于参考多肽(例如,LeuD酶、KivD酶或Adh酶)的氨基酸序列,一个或更多个突变改变重组多肽(例如,LeuD酶、KivD酶或Adh酶)的氨基酸序列,并且相对于参考多肽改变(增强或降低)多肽的活性。

[0127] 可以使用常规方法测量本公开中描述的重组多肽(例如,LeuD酶、KivD酶或Adh酶)中任何一种的活性(例如,比活性)。作为非限制性实例,可以通过测量重组多肽的底物特异性、产生的一种或多种产物、产生的一种或多种产物的浓度、或其任何组合来确定重组多肽的活性。如本公开中所使用的,重组多肽的“比活性”指每单位时间针对给定量(例如,浓度)的重组多肽产生的特定产物的量(例如,浓度)。

[0128] 本领域技术人员还将认识到,重组多肽(例如,LeuD酶、KivD酶或Adh酶)编码序列中的突变可以造成保守氨基酸置换,以提供前述多肽的功能等效变体(例如,保留多肽的活性的变体)。如本公开中所使用的,“保守氨基酸置换”指不改变进行氨基酸置换的蛋白质的相对电荷特性或尺寸特性或功能活性的氨基酸置换。

[0129] 在一些情况下,氨基酸的特征在于其R基团(参见,例如,表1)。例如,氨基酸可以包括非极性脂族R基团、带正电荷的R基团、带负电荷的R基团、非极性芳族R基团、或极性不带电荷的R基团。包括非极性脂族R基团的氨基酸的非限制性实例包含丙氨酸、甘氨酸、缬氨酸、亮氨酸、甲硫氨酸和异亮氨酸。包括带正电荷的R基团的氨基酸的非限制性实例包含赖氨酸、精氨酸和组氨酸。包括带负电荷的R基团的氨基酸的非限制性实例包含天门冬氨酸盐和谷氨酸盐。包括非极性芳族R基团的氨基酸的非限制性实例包含苯丙氨酸、酪氨酸和色氨酸。包括极性不带电荷的R基团的氨基酸的非限制性实例包含丝氨酸、苏氨酸、半胱氨酸、脯氨酸、天门冬酰胺和谷氨酰胺。

[0130] 可以根据本领域普通技术人员已知的用于改变多肽序列的方法(如在编纂这样的方法的参考文献(例如,Molecular Cloning:A Laboratory Manual,J.Sambrook,et al.,eds.,Fourth Edition,Cold Spring Harbor Laboratory Press,Cold Spring Harbor,New York,2012、或Current Protocols in Molecular Biology,F.M.Ausubel,et al.,eds.,John Wiley&Sons,Inc.,New York,2010)中发现的)来制备变体。

[0131] 多肽的功能等效变体的非限制性实例可以包含本申请中公开的蛋白质的氨基酸序列中的保守氨基酸置换。如本公开中所使用的,“保守置换”与“保守氨基酸置换”可互换地使用,并且指表1中提供的氨基酸置换中的任何一种。

[0132] 在一些实施方案中,在制备变体多肽时,可以改变1个、2个、3个、4个、5个、6个、7个、8个、9个、10个、11个、12个、13个、14个、15个、16个、17个、18个、19个、20个或超过20个残基。在一些实施方案中,氨基酸被保守氨基酸置换替代。

[0133] 表1.保守氨基酸置换

[0134]	原始残基	R基团类型	保守氨基酸置换
	Ala	非极性脂族R基团	Cys、Gly、Ser
	Arg	带正电荷的R基团	His、Lys
	Asn	极性不带电荷的R基团	Asp、Gln、Glu
	Asp	带负电荷的R基团	Asn、Gln、Glu

Cys	极性不带电荷的R基团	Ala、Ser
Gln	极性不带电荷的R基团	Asn、Asp、Glu
Glu	带负电荷的R基团	Asn、Asp、Gln
Gly	非极性脂族R基团	Ala、Ser
His	带正电荷的R基团	Arg、Tyr、Trp
Ile	非极性脂族R基团	Leu、Met、Val
Leu	非极性脂族R基团	Ile、Met、Val
Lys	带正电荷的R基团	Arg、His
Met	非极性脂族R基团	Ile、Leu、Phe、Val
Pro	极性不带电荷的R基团	
Phe	非极性芳族R基团	Met、Trp、Tyr
Ser	极性不带电荷的R基团	Ala、Gly、Thr
Thr	极性不带电荷的R基团	Ala、Asn、Ser
Trp	非极性芳族R基团	His、Phe、Tyr、Met
Tyr	非极性芳族R基团	His、Phe、Trp
Val	非极性脂族R基团	Ile、Leu、Met、Thr

[0135] 可以通过改变多肽(例如,LeuDh酶、KivD酶或Adh酶)的编码序列来进行多肽的氨基酸序列中的氨基酸置换以产生具有期望性质和/或活性的重组多肽(例如,LeuDh酶、KivD酶或Adh酶)变体。类似地,通常通过改变重组多肽(例如,LeuDh酶、KivD酶或Adh酶)的编码序列来进行多肽的氨基酸序列中的保守氨基酸置换以产生多肽的功能等效变体。

[0136] 可以通过本领域普通技术人员已知的多种方法在核苷酸序列中进行突变(例如,置换)。例如,可以通过PCR定向突变、根据Kunkel的方法(Kunkel, Proc.Nat.Acad.Sci.U.S.A.82:488-492,1985)的定点突变、通过编码多肽的基因的化学合成、通过基因编辑技术、或者通过插入(如标签(例如,HIS标签或GFP标签)的插入)来进行突变。

[0137] 编码支链氨基酸(BCAA)途径酶的核酸

[0138] 本公开的方面涉及重组酶、其功能化修饰和变体、以及与其相关的应用。例如,本申请中描述的酶和细胞可以用于例如通过将亮氨酸转化成异戊醇来促进亮氨酸消耗。方法可以包括使用包括本申请中公开的一种或更多种酶的宿主细胞、细胞裂解物、分离的酶、或其任何组合。本公开涵盖包括在宿主细胞中重组表达编码本申请中公开的酶的多核苷酸的方法。本公开涵盖包括向有需要的受试者施用包括至少一种BCAA途径酶(例如,LeuDh酶、KivD酶或Adh酶)的宿主细胞的方法。本公开还涵盖包括使反应混合物中的一种或更多种支链氨基酸(BCAA)与本申请中公开的BCAA途径酶反应的体外方法。在一些实施方案中,BCAA途径酶是LeuDh酶、KivD酶、或Adh酶、或其组合。

[0139] 编码重组多肽(例如,LeuDh、KivD、Adh和/或BrnQ)中的任何一种或更多种的核酸被本公开涵盖,并且可以被包括在宿主细胞内。在一些实施方案中,核酸是操纵子的形式。在一些实施方案中,至少一个核糖体结合位点存在于核酸中存在的编码序列中的一个或更多个之间。

[0140] 在一些实施方案中,本公开涵盖的LeuDh核酸序列、KivD核酸序列、Adh核酸序列

和/或BrnQ核酸序列是与本公开中提供的LeuDh核酸序列、KivD核酸序列、Adh核酸序列和/或BrnQ核酸序列在高严格性条件或中等严格性条件下杂交并且具有生物学活性的核酸序列。例如,可以使用在65°C的0.2×SSC至1×SSC、然后在65°C以0.2×SSC洗涤的高严格性条件下与编码LeuDh、KivD、Adh和/或BrnQ的核酸杂交的核酸。可以使用在室温的6×SSC、然后在室温以2×SSC洗涤的低严格性条件下与编码LeuDh、KivD、Adh和/或BrnQ的核酸杂交的核酸。其他杂交条件包含40°C或50°C的3×SSC、然后在20°C、30°C、40°C、50°C、60°C或65°C用1×SSC或2×SSC洗涤。

[0141] 可以在存在甲醛(例如,10%、20%、30%、40%或50%)的情况下进行杂交,甲醛的存在进一步增加了杂交的严格性。例如,S.阿格拉瓦尔(S.Agrawal)(编者)分子生物学方法(Methods in Molecular Biology),第20卷;和泰森(Tijssen)(1993)生物化学和分子生物学实验技术-核酸探针杂交(Laboratory Techniques in biochemistry and molecular biology-hybridization with nucleic acid probes)(例如,第I部分第2章“杂交原理和核酸探针测定的策略的概述(Overview of principles of hybridization and the strategy of nucleic acid probe assays)”,纽约爱思维尔提供核酸杂交的基本指导)中描述了核酸杂交的理论和实践。示例性蛋白质可以与LeuDh蛋白、KivD蛋白、或Adh蛋白或其域(例如,催化域)具有至少约50%、70%、80%、90%(优选至少约95%,甚至更优选至少约98%,并且最优选至少99%)的同源性或一致性。其他示例性蛋白质可以由与LeuDh核酸、KivD核酸或Adh核酸(例如,本申请中描述的那些LeuDh核酸、KivD核酸或Adh核酸)至少约90%(优选至少约95%,甚至更优选至少约98%,并且最优选至少99%)的同源性或一致性的核酸编码。

[0142] 可以通过本领域已知的任何方法将编码本申请中描述的重组多肽(例如,LeuDh、KivD、Adh和/或BrnQ)中任何一种或更多种的核酸并入任何适当的运载体中。例如,运载体可以是表达载体(包含但不限于病毒运载体(例如,慢病毒运载体、逆转录病毒运载体、腺病毒运载体、或腺相关病毒运载体)、适合于瞬时表达的任何运载体、适合于组成型表达的任何运载体、或者适合于诱导型表达的任何运载体(例如,半乳糖诱导型运载体或强力霉素诱导型运载体))。

[0143] 在一些实施方案中,运载体在细胞中自主复制。在一些实施方案中,将运载体整合到细胞内的染色体中。运载体可以含有一个或更多个核酸内切酶限制性位点,核酸内切酶限制位点被限制性核酸内切酶切割以插入和连接含有本申请中描述的基因的核酸,以产生能够在细胞中复制的重组运载体。运载体通常由DNA组成,尽管RNA运载体也是可用的。克隆运载体包含(但不限于):质粒、F黏粒(fosmid)、噬菌粒、病毒基因组和人工染色体。如本申请中所使用的,术语“表达运载体”或“表达构建体”指重组或合成生成的、具有一系列容许特定核酸在宿主细胞(例如,微生物)(如酵母细胞)中转录的指定核酸元件的核酸构建体。在一些实施方案中,将本申请中描述的基因的核酸序列插入克隆运载体中,使得其可操作地连接至调控序列,并且在一些实施方案中表达为RNA转录物。在一些实施方案中,运载体含有一种或更多种标志物(如本申请中描述的可选择的标志物),以鉴别用重组运载体转化或转染的细胞。在一些实施方案中,本申请中描述的基因的核酸序列是经密码子优化的。密码子优化可以将基因产物的产量相对于未经密码子优化的参考序列增加至少10%、至少15%、至少20%、至少25%、至少30%、至少35%、至少40%、至少45%、至少50%、至少55%、

至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、或100% (包含之间的全部值)。

[0144] 在一些实施方案中,本申请中描述的核酸序列在质粒中表达。例如,本申请中描述的核酸序列可以在克隆质粒中表达。本申请中描述的核酸序列可以在质粒中表达以用于瞬时表达。本申请中描述的核酸序列也可以在质粒中表达以将核酸序列并入基因组DNA中。

[0145] 当编码序列和调控序列共价地连接并且编码序列的表达或转录受到调控序列的影响或控制时,编码序列和调控序列被称为“可操作地连接(operably joined)”或“可操作地连接(operably linked)”。如果编码序列被翻译成功能蛋白,则如果5'调控序列中启动子的诱导容许编码序列被转录,并且如果编码序列与调控序列之间的联接的性质不会(1)造成移码突变的引入;(2)干扰启动子区指导编码序列的转录的能力、或(3)干扰相应RNA转录物被翻译成蛋白质的能力,则编码序列和调控序列被称为是可操作地连接。

[0146] 在一些实施方案中,编码本申请中描述的蛋白质中的任何一种或更多种的核酸受调控序列(例如,增强子序列)的控制。在一些实施方案中,核酸在启动子的控制下表达。启动子可以是天然启动子(例如,基因在其内源环境中的启动子,该启动子提供基因表达的正常调控)。可替代地,启动子可以是与基因的天然启动子不同的启动子,例如,启动子与基因在其内源环境中的启动子不同。

[0147] 在一些实施方案中,启动子是真核启动子。真核启动子的非限制性实例包含如本领域普通技术人员已知的TDH3、PGK1、PKC1、PDC1、TEF1、TEF2、RPL18B、SSA1、TDH2、PYK1、TPI1、GAL1、GAL10、GAL7、GAL3、GAL2、MET3、MET25、HXT3、HXT7、ACT1、ADH1、ADH2、CUP1-1、ENO2和SOD1(参见,例如,Addgene网站:blog.addgene.org/plasmids-101-the-promoter-region)。在一些实施方案中,启动子是原核启动子(例如,噬菌体启动子或细菌启动子)。如本领域普通技术人员所已知的,噬菌体启动子的非限制性实例包含P_λs1con、T3、T7、SP6和PL。细菌启动子的非限制性实例包含P_{bad}、P_{mgrB}、P_{trc2}、P_{C1857}、P_{lac/ara}、P_{lac/fnr}、P_{tac}、P_{tet}、P_{cmt}和P_m。

[0148] 在一些实施方案中,启动子是诱导型启动子。如本申请中所使用的,“诱导型启动子”是受到分子的存在或不存在控制的启动子。例如,这可以用于可控地诱导酶的表达。在一些实施方案中,在诱导型启动子与LeuDH、KivD和/或Adh连接的情况下,可以在某些时间处诱导LeuDH、KivD和/或Adh的表达或者可以在某些时间处不诱导LeuDH、KivD和/或Adh的表达。例如,在一些实施方案中,可以在某些时间处不诱导表达,以便限制亮氨酸消耗(例如,在细胞生长期)。诱导型启动子的非限制性实例包含化学调控的启动子和物理调控的启动子。对于化学调控的启动子,转录活性可以由一种或更多种化合物(如醇、四环素、半乳糖、类固醇、金属、或其他化合物)调控。对于物理调控的启动子,转录活性可以受现象(如光或温度)的调控。四环素调控的启动子的非限制性实例包含脱水四环素(aTc)响应性启动子和其他四环素响应性启动子系统(例如,四环素阻遏蛋白(tetR)、四环素操纵子序列(tetO)和四环素反式激活子融合蛋白(tTA))。类固醇调控的启动子的非限制性实例包含基于大鼠糖皮质激素受体、人雌激素受体、蛾蜕皮激素受体的启动子,以及来自类固醇/类维生素A/甲状腺受体超家族的启动子。金属调控的启动子的非限制性实例包含来源于金属硫蛋白(结合并且螯合金属离子的蛋白质)基因的启动子。发病机制调控的启动子的非限制性实例包含由水杨酸、乙烯或苯并噻二唑(BTH)诱导的启动子。温度/热诱导型启动子的非限制性实例包含由水杨酸、乙烯或苯并噻二唑(BTH)诱导的启动子。温度/热诱导型启动子的非限制性

实例包含热激启动子。光调控的启动子的非限制性实例包含来自植物细胞的光响应性启动子。在某些实施方案中,诱导型启动子是半乳糖诱导型启动子。在一些实施方案中,通过一种或更多种生理条件(例如,pH、温度、辐射、渗透压、盐水梯度、细胞表面结合、或者一种或更多种外在诱导剂或内在诱导剂的浓度)来诱导诱导型启动子。外在诱导物或诱导剂的非限制性实例包含氨基酸和氨基酸类似物、糖类和多糖、核酸、蛋白质转录激活子(activator)和阻遏子(repressor)、细胞因子、毒素、石油基化合物、含金属的化合物、盐、离子、酶底物类似物、激素或其任何组合。

[0149] 在一些实施方案中,启动子是组成型启动子。如本申请中所使用的,“组成型启动子”指允许基因的连续转录的未经调控的启动子。组成型启动子的非限制性实例包含TDH3、PGK1、PKC1、PDC1、TEF1、TEF2、RPL18B、SSA1、TDH2、PYK1、TPI1、HXT3、HXT7、ACT1、ADH1、ADH2、ENO2和SOD1。

[0150] 本申请中也预期了本领域普通技术人员已知的其他诱导型启动子或组成型启动子。

[0151] 基因表达所需的调控序列的确切性质可能在物种或细胞类型之间变化,但通常视需要包含分别涉及转录和翻译的起始的5' 非转录序列和5' 非翻译序列(如TATA框、加帽序列、CAAT序列等)。特别地,这样的5' 非转录调控序列将包含启动子区,该启动子区包含用于可操作地连接的基因的转录控制的启动子序列。调控序列还可以包含增强子序列或上游激活子序列。本申请中公开的运载体可以包含5' 前导序列(leader)或信号序列。调控序列还可以包含终止子序列。在一些实施方案中,终止子序列在转录期间标记DNA中基因的末端。适合于诱导异源生物体中的本申请中描述的一个或更多个基因的表达的一种或更多种适当的运载体的选择和设计在本领域普通技术人员的能力和判断范围之内。

[0152] 含有表达必需元件的表达运载体是可商业获得的,并且是本领域普通技术人员已知的(参见,例如,Sambrook et al., Molecular Cloning:A Laboratory Manual,Fourth Edition,Cold Spring Harbor Laboratory Press,2012)。

[0153] 宿主细胞

[0154] 用大肠埃希氏菌细胞(例如,大肠埃希氏菌Nissle 1917)举例说明所公开的方法和组合物以及宿主细胞,但在一些实施方案中适用于其他宿主细胞。

[0155] 适合的宿主细胞包含(但不限于):酵母细胞、细菌细胞、藻类细胞、植物细胞、真菌细胞、昆虫细胞和动物细胞(包含哺乳动物细胞)。在一个说明性实施方案中,适合的宿主细胞包含大肠埃希氏菌(例如,从马萨诸塞州伊普斯威奇的New England BioLabs可获得的ShuffleTM感受态大肠埃希氏菌或者从德国微生物和细胞培养物保藏中心可获得的大肠埃希氏菌Nissle 1917(DSMZ Braunschweig,大肠埃希氏菌DSM 6601))。

[0156] 适合的酵母宿主细胞包含(但不限于):假丝酵母属、汉逊酵母属、酵母属、裂殖酵母属、毕赤酵母属、克鲁维酵母属和耶氏酵母属。在一些实施方案中,酵母细胞是多形汉逊酵母、酿酒酵母、卡尔斯伯酵母(*Saccharomyces carlsbergensis*)、糖化酵母(*Saccharomyces diastaticus*)、诺地酵母(*Saccharomyces norbensis*)、克鲁弗酵母(*Saccharomyces kluyveri*)、粟酒裂殖酵母、毕赤酵母、芬兰毕赤酵母(*Pichia finlandica*)、喜海藻糖毕赤酵母(*Pichia trehalophila*)、*Pichia kodamae*、膜醭毕赤酵母(*Pichia membranaefaciens*)、仙人掌毕赤酵母(*Pichia opuntiae*)、耐热毕赤酵母(*Pichia*

thermotolerans)、柳毕赤酵母 (*Pichia salictaria*)、栎树毕赤酵母 (*Pichia quercuum*)、皮杰普氏毕赤酵母 (*Pichia pijperi*)、树干毕赤酵母 (*Pichia stipitis*)、甲醇毕赤酵母 (*Pichia methanolica*)、安格斯毕赤酵母 (*Pichia angusta*)、乳酸克鲁维酵母 (*Kluyveromyces lactis*)、白假丝酵母 (*Candida albicans*)、或解脂耶氏酵母 (*Yarrowia lipolytica*)。

[0157] 在一些实施方案中,酵母菌株是工业多倍体酵母菌株。真菌细胞的其他非限制性实例包含获自曲霉属、青霉属、镰刀菌属、根霉属、支顶孢属、脉孢菌属、粪壳菌属、稻瘟菌属、异水霉属、黑粉菌属、葡萄孢属和木霉属的细胞。

[0158] 在某些实施方案中,宿主细胞是藻类细胞(如衣藻属(例如,莱茵衣藻)和席藻属(席藻属ATCC29409))。

[0159] 在其他实施方案中,宿主细胞是原核细胞。适合的原核细胞包含革兰氏阳性细菌细胞、革兰氏阴性细菌细胞和革兰氏不定细菌细胞。宿主细胞可以是(但不限于):农杆菌属、脂环酸芽孢杆菌属(*Alicyclobacillus*)、鱼腥藻属、倒囊藻属、不动杆菌属、热酸菌属(*Acidothiermus*)、节杆菌属、固氮菌属、芽孢杆菌属、双歧杆菌属、短杆菌属、丁酸弧菌属、布赫纳氏菌属(*Buchnera*)、油菜属菌属(*Campestris*)、弯曲菌属、梭菌属、棒状杆菌属、着色菌属、粪球菌属、埃希氏菌属、肠球菌属、肠杆菌属、欧文氏菌属、梭杆菌属、粪杆菌属、弗朗西斯氏菌属、黄杆菌属、地芽孢杆菌属、嗜血杆菌属、螺杆菌属、克雷伯菌属、乳杆菌属、乳球菌属、泥杆菌属(*Ilyobacter*)、微球菌属、微杆菌属、中间根瘤菌属(*Mesorhizobium*)、甲基杆菌属、甲基杆菌属、分枝杆菌属、奈瑟氏菌属、泛菌属、假单胞菌属、原绿球藻(*Prochlorococcus*)、红细菌属、红假单胞菌属、红假单胞菌属、罗氏菌属(*Roseburia*)、红螺菌属、红球菌属、栅藻属、链霉菌属、链球菌属、聚球藻属(*Synechococcus*)、糖单孢菌属、糖多孢菌属、葡萄球菌属、沙雷氏菌属、沙门氏菌属、志贺氏菌属、嗜热厌氧杆菌属(*Thermoanaerobacterium*)、*Tropheryma*、*Tularensis*、*Temecula*、嗜热聚球藻属(*Thermosynechococcus*)、热球菌属(*Thermococcus*)、尿素原体(*Ureaplasma*)、黄杆菌属、小菌属(*Xylella*)、耶尔森氏菌属和发酵单胞菌属。

[0160] 在一些实施方案中,细菌宿主菌株是工业菌株。许多细菌工业菌株是已知的,并且适合于本申请中描述的方法和组合物。

[0161] 在一些实施方案中,细菌宿主细胞是农杆菌属(例如,放射形农杆菌(*A.radiobacter*)、发根农杆菌(*A.rhizogenes*)、悬钩子农杆菌(*A.rubi*))、节杆菌属(例如,金黄节杆菌(*A.aureescens*)、柠檬节杆菌(*A.citreus*)、球形节杆菌(*A.globiformis*)、裂烃谷氨酸节杆菌(*A.hydrocarboglutamicus*)、迈索尔节杆菌(*A.mysorens*)、烟草节杆菌(*A.nicotianae*)、石蜡节杆菌(*A.paraffineus*)、原玻璃蝇节杆菌(*A.protophonniae*)、玫瑰色石蜡节杆菌(*A.roseoparaffinus*)、硫磺节杆菌(*A.sulfureus*)、产脲节杆菌(*A.ureafaciens*))、芽孢杆菌属(例如,苏云金芽孢杆菌(*B.thuringiensis*)、炭疽芽孢杆菌(*B.anthracis*)、巨大芽孢杆菌(*B.megaterium*)、枯草芽孢杆菌(*B.subtilis*)、迟缓芽孢杆菌(*B.lentus*)、环状芽孢杆菌(*B.circularis*)、短小芽孢杆菌(*B.pumilus*)、灿烂芽孢杆菌(*B.lautus*)、凝结芽孢杆菌(*B.coagulans*)、短芽孢杆菌(*B.brevis*)、坚强芽孢杆菌(*B.firmus*)、嗜碱芽孢杆菌(*B.alkaophilus*)、地衣芽孢杆菌(*B/licheniformis*)、克劳氏芽孢杆菌(*B.clausii*)、嗜热脂肪芽孢杆菌(*B.stearothermophilus*)、耐盐芽孢杆菌

(*B.halodurans*) 和解淀粉芽孢杆菌 (*B.amyloliquefaciens*))。在特定实施方案中,宿主细胞将是工业芽孢杆菌菌株(包含但不限于枯草芽孢杆菌、短小芽孢杆菌、地衣芽孢杆菌、巨大芽孢杆菌、克劳氏芽孢杆菌、嗜热脂肪芽孢杆菌和解淀粉芽孢杆菌)。在一些实施方案中,宿主细胞将是工业梭菌属(例如,丙酮丁醇梭菌 (*C.acetobutylicum*)、破伤风梭菌E88 (*C.tetani* E88)、象牙海岸梭菌 (*C.lituseburens*e)、糖丁基梭菌 (*C.saccharo**butylicum*)、产气荚膜梭菌 (*C.perfringens*)、拜氏梭菌 (*C.beijerinckii*))。在一些实施方案中,宿主细胞将是工业棒状杆菌属(例如,谷氨酸棒状杆菌 (*C.g**lutamicum*)、嗜醋酸棒状杆菌 (*C.acetoacidophilum*))。在一些实施方案中,宿主细胞将是工业埃希氏菌属(例如,大肠埃希氏菌)。在一些实施方案中,宿主细胞将是工业欧文氏菌属(例如,噬夏孢欧文氏菌 (*E.uredovora*)、软腐欧文氏菌 (*E.carotovora*)、菠萝欧文氏菌 (*E.ananas*)、草生欧文氏菌 (*E.herbicola*)、斑点欧文氏菌 (*E.punctata*)、*E.terreus*)。在一些实施方案中,宿主细胞将是工业泛菌属(例如,柠檬泛菌 (*P.citrea*)、成团泛菌 (*P.agglomerans*))。在一些实施方案中,宿主细胞将是工业假单胞菌属(例如,恶臭假单胞菌 (*P.putida*)、绿脓假单胞菌 (*P.aeruginosa*)、梅瓦隆假单胞菌 (*P.mevalonii*))。在一些实施方案中,宿主细胞将是工业链球菌属(例如,相似型链球菌 (*S.equisimiles*)、酿脓链球菌 (*S.pyogenes*)、乳房链球菌 (*S.uberis*))。在一些实施方案中,宿主细胞将是工业链霉菌属(例如,产二素链霉菌 (*S.ambofaciens*)、不产色链霉菌 (*S.achromogenes*)、阿维链霉菌 (*S.avermitilis*)、天蓝色链霉菌 (*S.coelicolor*)、金霉素链霉菌 (*S.aureofaciens*)、金色链霉菌 (*S.aureus*)、杀真菌链霉菌 (*S.fungicidicus*)、灰色链霉菌 (*S.griseus*)、变铅青链霉菌 (*S.lividans*))。在一些实施方案中,宿主细胞将是工业发酵单胞菌属(例如,运动发酵单胞菌 (*Z.mobilis*)、解脂发酵单胞菌 (*Z.lipolytica*))等。

[0162] 本公开也适合于与多种动物细胞类型(包含哺乳动物细胞,例如,人类细胞系(包含293细胞、海拉细胞、WI38细胞、PER.C6细胞和Bowes黑素瘤细胞)、小鼠细胞系(包含3T3、NS0、NS1、Sp2/0)、仓鼠细胞系(CHO、BHK)、猴细胞系(COS、FRhL、Vero)和杂交瘤细胞系)使用。

[0163] 在各种实施方案中,可以在本公开的实践中使用的菌株包含原核菌株和真核菌株两者,并且公众易于从多个培养物保藏中心(如美国典型培养物保藏中心(American Type Culture Collection) (ATCC)、德国微生物菌种保藏中心(Deutsche Sammlung von Mikroorganismen and Zellkulturen GmbH) (DSM)、荷兰微生物菌种保藏中心(Centraalbureau Voor Schimmelcultures) (CBS) 和美国农业研究服务专利培养物保藏中心北方地区研究中心(Agricultural Research Service Patent Culture Collection, Northern Regional Research Center) (NRRL) 获取。本公开也适合于与多种植物细胞类型使用。

[0164] 如本申请中所使用的,术语“细胞”可以指单个细胞或细胞群体(如属于相同细胞系或菌株的细胞群体)。不应当将单数术语“细胞”的使用解释为明确地指单个细胞而不是细胞群体。相对于野生型对应物,宿主细胞可以包括基因修饰。

[0165] 可以使用本领域已知的任何方法将编码本申请中描述的重组多肽(例如,LeuD酶、KivD酶、Adh酶和/或BrnQ)中任何一种或更多种的运载体引入适合的宿主细胞中。可以在本领域普通技术人员理解的任何适合的条件下培养宿主细胞。例如,可以使用本领域已

知的任何培养基、温度和孵育条件。对于携带诱导型运载体的宿主细胞,可以用适当的诱导剂培养细胞以促进表达。

[0166] 可以在接触和/或核酸的整合之前、在接触和/或核酸的整合期间、和/或在接触和/或核酸的整合之后在任何类型(富集的或基本的)和任何组成的培养基中培养本申请中公开的任何细胞。如本领域普通技术人员所理解的,可以通过常规实验优化培养物或培养过程的条件。在一些实施方案中,选择的培养基补充有各种组分。在一些实施方案中,优化补充组分的浓度和量。在一些实施方案中,通过常规实验优化培养基和生长条件(例如,pH、温度等)的其他方面。在一些实施方案中,优化培养基补充一种或更多种补充组分的频率、以及培养细胞的时间量。

[0167] 可以在本领域已知且使用的培养容器中进行本申请中描述的细胞的培养。在一些实施方案中,充气反应容器(例如,搅拌釜反应器)用于培养细胞。在一些实施方案中,生物反应器或发酵器用于培养细胞。因此,在一些实施方案中,在发酵中使用细胞。如本申请中所使用的,术语“生物反应器”和术语“发酵器”可互换地使用,并且指在其中发生生物反应、生物化学反应和/或化学反应(涉及活生物体或活生物体的一部分)的包围物或部分包围物。“大规模生物反应器”或“工业规模生物反应器”是用于以商业规模或准商业规模生成产物的生物反应器。大型生物反应器通常具有在升、数百升、数千升、或更大范围内的体积。

[0168] 在一些实施方案中,生物反应器包括细胞(例如,细菌细胞)或细胞培养物(例如,细菌细胞培养物)(如本申请中描述的细胞或细胞培养物)。在一些实施方案中,生物反应器包括孢子和/或分离微生物的休眠细胞类型(例如,处于干燥状态的休眠细胞)。

[0169] 生物反应器的非限制性实例包含:搅拌釜发酵器、通过旋转混合装置搅动的生物反应器、恒化器、通过振动装置搅动的生物反应器、气升式发酵器、填充床反应器、固定床反应器、流化床生物反应器、采用波诱导的搅动的生物反应器、离心生物反应器、滚瓶、以及中空纤维生物反应器、滚转器设备(例如,台式种类、推车安装式种类、和/或自动化种类)、竖直堆叠的板、旋转瓶、搅拌瓶或摇动瓶、振动的多孔板、MD瓶、方瓶、洛克斯氏瓶、多表面组织培养繁殖器、改良的发酵器、以及经涂覆的珠(例如,用血清蛋白、硝化纤维素或羧甲基纤维素涂覆的珠以防止细胞附着)。

[0170] 在一些实施方案中,生物反应器包含细胞培养系统,其中细胞(例如,细菌细胞)与运动的液体和/或气泡接触。在一些实施方案中,细胞或细胞培养物悬浮生长。在其他实施方案中,细胞或细胞培养物附着于固相载体。载体系统的非限制性实例包含微载体(例如,可以是多孔或无孔的聚合物球、微珠和微盘)、带有特定化学基团(例如,叔胺基团)的交联珠(例如,右旋糖酐)、2D微载体(包含捕获在无孔聚合物纤维中的细胞)、3D载体(例如,载体纤维、中空纤维、多筒反应器(multicartridge reactor)、以及可以包括多孔纤维的半渗透膜)、具有降低的离子交换能力的微载体、微囊化细胞、毛细管、以及聚集体。在一些实施方案中,由材料(如右旋糖酐、明胶、玻璃、或纤维素)制造载体。

[0171] 在一些实施方案中,以连续模式、半连续模式或非连续模式操作工业规模的过程。操作模式的非限制性实例是分批、补料分批(fed batch)、扩展分批(extended batch)、重复分批(repetitive batch)、抽取/填充、旋转壁、旋转瓶、和/或灌注操作模式。在一些实施方案中,生物反应器允许连续或半连续补充底物原料(例如,碳水化合物来源)和/或从生物反应器连续或半连续分离产物。

[0172] 在一些实施方案中,生物反应器或发酵器包含传感器和/或控制系统以测量和/或调整反应参数。反应参数的非限制性实例包含生物学参数(例如,生长速率、细胞尺寸、细胞数量、细胞密度、细胞类型、或细胞状态等)、化学参数(例如,pH、氧化还原电位、反应底物和/或产物的浓度、溶解的气体的浓度(如氧气浓度和CO₂浓度)、营养物浓度、代谢物浓度、寡肽的浓度、氨基酸的浓度、维生素的浓度、激素的浓度、添加剂的浓度、血清浓度、离子强度、离子的浓度、相对湿度、摩尔浓度、同渗容摩、其他化学物质(例如,缓冲剂、佐剂或反应副产物)的浓度)、物理/机械参数(例如,密度、传导率、搅拌程度、压力、和流速、剪切应力、剪切速率、粘度、颜色、浊度、光吸收、混合速率、转化率、以及热力学参数(如温度、光强度/质量)等)。测量本申请中描述的参数的传感器对于相关机械和电子领域的普通技术人员来说是公知的。控制系统基于来自本申请中描述的传感器的输入来调整生物反应器中的参数是生物反应器工程领域的普通技术人员公知的。

[0173] 在一些实施方案中,方法涉及分批发酵(例如,摇瓶发酵)。分批发酵(例如,摇瓶发酵)的一般考虑因素包含氧气和葡萄糖的水平。例如,分批发酵(例如,摇瓶发酵)可能受限于氧气和葡萄糖,因此在一些实施方案中,菌株在设计良好的补料分批发酵中进行的能力被低估。另外,最终产物可以在溶解性、毒性、细胞累积和分泌方面显示出与底物的一些差异,并且在一些实施方案中可以具有不同的发酵动力学。

[0174] 在一些实施方案中,本公开的细胞适合于在体内消耗亮氨酸。在一些实施方案中,细胞适合于产生一种或更多种用于经由转化为异戊醇消耗亮氨酸的酶(例如,LeuDH、KivD和/或Adh)。在这样的实施方案中,酶可以通过在体外过程或离体过程中的生物转化来催化用于亮氨酸的消耗的反应。

[0175] 本公开的蛋白质或酶中的任何一种可以在宿主细胞中表达。如本申请中所使用的,宿主细胞是可以用于表达至少一种异源多核苷酸(例如,编码本申请中描述的蛋白质或酶)的细胞。就多核苷酸(如包括基因的多核苷酸)而言,术语“异源”与术语“外源的”和术语“重组”可互换地使用,并且指:已经被人工地提供给生物系统的多核苷酸;已经在生物系统内修饰的多核苷酸、或者已经在生物系统内操纵其表达或调控的多核苷酸。被引入宿主细胞中或者在宿主细胞中表达的异源多核苷酸可以是来自与宿主细胞不同的生物体或物种的多核苷酸,或者可以是合成的多核苷酸,或者可以是也在与宿主细胞相同的生物体或物种中内源表达的多核苷酸。例如,当在宿主细胞中内源表达的多核苷酸非天然地位于宿主细胞中;在宿主细胞中稳定或瞬时重组表达;在宿主细胞内被修饰;在宿主细胞内被选择性编辑;在宿主细胞内以不同于天然存在的拷贝数的拷贝数表达;或者在宿主细胞内以非天然方式表达(如通过操纵控制多核苷酸的表达的调控区)时,在宿主细胞中内源表达的多核苷酸可以被认为是异源的。在一些实施方案中,异源多核苷酸是一多核苷酸,所述多核苷酸在宿主细胞中内源表达,但所述多核苷酸的表达由不天然调控多核苷酸表达的启动子驱动。在其他实施方案中,异源多核苷酸是一多核苷酸,所述多核苷酸在宿主细胞中内源表达,并且所述多核苷酸的表达由天然调控多核苷酸表达的启动子驱动,但所述启动子或另外的调控区被修饰。在一些实施方案中,启动子被重组激活或阻抑。例如,基于基因编辑的技术可以用于调控多核苷酸(包含内源多核苷酸)自启动子(包含内源启动子)的表达。参见,例如,Chavez et al., Nat Methods. 2016 Jul; 13(7):563-567。与参考多核苷酸序列相比,异源多核苷酸可以包括野生型序列或突变序列。

[0176] 任何适合的宿主细胞可以被用于产生本申请中公开的重组多肽(例如,LeuDH、KivD和/或Adh)(包含真核细胞或原核细胞)中的任何一种。

[0177] 组合物

[0178] 本公开提供组合物(包含药物组合物),组合物包括本申请中描述的宿主细胞(例如,包括编码选自由LeuDH、KivD和Adh组成的组的至少一种酶的异源多核苷酸的宿主细胞)或本申请中描述的一种或更多种酶(例如,LeuDH、KivD和/或Adh)、以及可选的药学上可接受的赋形剂。

[0179] 在某些实施方案中,本申请中描述的宿主细胞以有效量提供在组合物(如药物组合物)中。在某些实施方案中,本申请中描述的一种或更多种酶以有效量提供在组合物(如药物组合物)中。在某些实施方案中,有效量是治疗有效量。在某些实施方案中,有效量是预防有效量。在一些实施方案中,有效量是足以治疗或改善MSUD的一种或更多种症状的量。

[0180] 在某些实施方案中,受试者是动物。在某些实施方案中,受试者是人类。在其他实施方案中,受试者是非人类动物。在某些实施方案中,受试者是哺乳动物。在某些实施方案中,受试者是非人类哺乳动物。在一些实施方案中,受试者是非哺乳动物。在某些实施方案中,受试者是驯养动物(如狗、猫、牛、猪、马、绵羊、鸡或山羊)。在某些实施方案中,受试者是伴侣动物(如狗或猫)。在某些实施方案中,受试者是家畜动物(如牛、猪、马、绵羊、鸡或山羊)。在某些实施方案中,受试者是动物园动物。在另外的实施方案中,受试者是研究动物(如啮齿动物(例如,小鼠、大鼠)、狗、猪或非人类灵长类动物)。

[0181] 可以通过本领域已知的任何方法制备本申请中描述的组合物(如药物组合物)。通常,这样的制备方法包含将本申请中描述的化合物(例如,“活性成分”)与载体或赋形剂和/或一种或更多种其他辅助成分联合,并且然后(如果必要的话和/或如果期望的话)成形和/或将产品包装成期望的单剂量单位或多剂量单位。

[0182] 方法

[0183] 在一些方面,本公开提供使用宿主细胞的方法。在一些实施方案中,本公开提供包括培养本申请中描述的宿主细胞(例如,包括编码选自由LeuDH、KivD和Adh组成的组的至少一种酶的异源多核苷酸的宿主细胞)的方法。在本申请的其他处描述了用于培养细胞的方法。在一些实施方案中,本公开提供从亮氨酸产生异戊醇的方法,方法包括培养本申请中描述的宿主细胞(例如,包括编码LeuDH、KivD和Adh的异源多核苷酸的宿主细胞)。在一些实施方案中,在体内(例如,在已经施用宿主细胞的人类受试者中)进行产生和培养。在一些实施方案中,离体(例如,在体外细胞培养环境中)进行产生。本申请中描述的组合物、细胞、酶和方法也适用于工业环境(包含其中可能存在支链氨基酸(例如,亮氨酸、异亮氨酸和缬氨酸)累积的任何应用)。

[0184] 通过以下实施例进一步阐明本发明,但绝不应当将以下实施例解释为限制。贯穿本申请所引用的全部参考文献(包含文献参考、授权专利、公布的专利申请、以及待审专利申请)的全部内容特此通过引用被明确并入。如果并入本申请的参考文献含有定义与本公开中定义的相同术语的定义不一致或不相容的术语,则应当以本公开中归于该术语的含义为准。然而,对本申请中引用的任何参考文献、文章、出版物、专利、专利公布和专利申请的提及不被认为是其构成有效现有技术或者形成世界上任何国家中的公知常识的一部分的承认或任何形式的暗示,也不应当被认为其构成有效现有技术或者形成世界上任何国家中

的公知常识的一部分的承认或任何形式的暗示。

实施例

[0185] 为了可以更充分地理解本申请中描述的发明,示出下列实施例。本申请中描述的实施例被提供以阐明本申请中提供的系统和方法,并且不被以任何方式解释为限制其范围。

[0186] 实施例1:酶库设计和合成

[0187] 材料和方法

[0188] 宏基因组酶发现

[0189] 基于机器学习的生物信息学工具用于在公共序列数据库(SwissProt和TrEMBL,一起被称为UniProt)中针对三种期望活性中的每种鉴别候选酶(亮氨酸脱氢酶,1.4.1.9;酮异戊酸脱羧酶,4.1.1.1;以及醇脱氢酶1.1.1.1)。对于LeuDh和Adh,使用先前开发的算法使序列多样性最大化。对于KivD,使用分层取样方法。候选酶的总数为1175种LeuDh序列、1296种KivD序列和1177种Adh序列。

[0190] 合理的酶设计

[0191] 对于LeuDh和Adh,使用Rosetta软件建立酶-过渡态复合物的分子模型,并且设计活性位点残基对20种氨基酸中的每种的系统突变。

[0192] 库合成

[0193] 全部LeuDh酶、KivD酶和Adh酶的DNA序列都被密码子优化以用于在大肠埃希氏菌中表达。在T7启动子的控制下在诱导型大肠埃希氏菌表达运载体中合成编码序列。

[0194] 结果

[0195] 为了改进消耗亮氨酸的支链氨基酸(BCAA)途径,进行实验以鉴别原型菌株(1980,也称为SYN1980)中相对于亲本酶具有更高活性的LeuDh酶、KivD酶和Adh酶,其中亲本菌株包含蜡样芽孢杆菌LeuDh、乳酸乳球菌KivD和酿酒酵母ADH2。原型菌株还包含来自大肠埃希氏菌的BrnQ,BrnQ是可以将支链氨基酸(如亮氨酸)转运到细胞中的支链氨基酸的转运体。亲本LeuDh酶展现出底物杂泛性,从而使缬氨酸和异亮氨酸(除亮氨酸以外)脱氨基化。为了改进通过BCAA途径对亮氨酸的特定消耗,途径设计的附加目标是鉴别相对于缬氨酸(Val)和异亮氨酸(Ile)对亮氨酸(Leu)具有增加的特异性的LeuDh酶。

[0196] 两种互补的方法用于针对每个酶家族(LeuDh、KivD和Adh)设计库:宏基因组来源和合理设计(表2)。对于每种酶,设计>1000种酶的宏基因组库,以对序列数据库中可用的完整宏基因组序列空间进行采样(图1A-图1C)。对于LeuDh库和Adh库,可用的结构数据用于蜡样芽孢杆菌LeuDh酶和酿酒酵母Adh酶的合理设计。全部库的酶序列针对在大肠埃希氏菌中的表达进行优化,并且在诱导型大肠埃希氏菌表达运载体中合成,并且转化到大肠埃希氏菌中以用于高通量筛选。

[0197] 表2.酶库组成

	库	细菌	真菌	动物	植物	合理的	总设计
[0198]	LeuDh	1129	11	23	12	270	1445
	KivD	783	508	1	4	0	1296
	Adh	654	273	128	122	140	1317

[0199] 实施例2:途径酶库的表征

[0200] 材料和方法

[0201] 细胞生长和酶制备

[0202] 对于筛选的酶库中的每个,将具有库质粒的菌株转化到大肠埃希氏菌T7表达宿主细胞中。将5 μ L/孔的解冻甘油原料压入用AeraSeal密封的半高深孔板中的500 μ L/孔的LB+100 μ g/mL羧苄青霉素(LB-Carb100)中。在37°C孵育样品,并且在80%湿度下以1000RPM摇动过夜。将50 μ L/孔的所得到的预培养物压入用AeraSeal密封的半高深孔板中的450 μ L/孔的LB-Carb100+1mM IPTG中。在30°C孵育样品,并且在80%湿度下以1000RPM摇动过夜。将250 μ L/孔的所得到的产生培养物压入含有500 μ L磷酸盐缓冲盐水(PBS)的深孔板中,并且以4000*G离心10分钟。去除上清液,并且将所得到的细胞沉淀物重悬于200 μ L的BugBuster蛋白提取试剂+1 μ L/mL纯化的核酸酶(Benzonase)+1 μ L/6mL纯化的溶菌酶中。将样品在室温孵育10分钟,以生成在体外酶测定中使用的细胞裂解物。

[0203] LeuD活性测定

[0204] 将LeuD库菌株的10 μ L裂解物转移到含有90 μ L/孔测定缓冲液(20mM氨基酸[L-亮氨酸、L-缬氨酸或L-异亮氨酸]、200mM甘氨酸、200mM KC1、0.4mM NAD, pH 10.5)的半面积平底板中。在读板器上进行光学测量,其中在340nm处读取吸光度读数10分钟。所得到的动力学数据用于解析NAD⁺还原的最大速率(LeuD活性的代用指标)。

[0205] KivD活性测定

[0206] 将KivD库菌株的10 μ L裂解物转移到含有90 μ L/孔测定缓冲液(100mM PIPES-KOH、100mM谷氨酸钾、1mM二硫苏糖醇、0.4mM NAD、1.5mM硫胺素焦磷酸、10mM谷氨酸镁、20mM酮异己酸(KIC), pH 7.5)的半面积平底板中。偶联酶用于间接地测量对KIC的KivD活性。在10分钟内进行光学吸光度测量。所得到的动力学数据用于确定KivD活性。

[0207] Adh活性测定

[0208] 将Adh库菌株的10 μ L裂解物转移到含有90 μ L/孔测定缓冲液(50mM MOPS缓冲液、0.4mM NADH和30mM异戊醛,pH 7.0)的半面积平底板中。在读板器上在340nm处进行光学吸光度测量达10分钟。所得到的动力学数据用于解析NADH氧化的最大速率(ADH活性的代用指标)。

[0209] LeuD选择性测定

[0210] 为了测量LeuD选择性(在存在L-Ile和L-Val的情况下L-Leu的特异性脱氨基化),将裂解物在裂解缓冲液中稀释四倍,并且将10 μ L/孔新稀释的裂解液压入90 μ L/孔的来自上文的改良测定缓冲液(特征在于0.5mM每种氨基酸(L-亮氨酸、L-异亮氨酸、L-缬氨酸)、200mM甘氨酸、200mM氯化钾和4mM NAD)中。在不同的时间点淬灭反应,并且进行亮氨酸、异亮氨酸和缬氨酸的LC-MS定量。

[0211] 结果

[0212] 为了筛选3×~1300个成员的酶库,开发高通量(HTP)方法以筛选大肠埃希氏菌细胞裂解物中的LeuD酶活性、KivD酶活性和Adh酶活性。简言之,在96深孔板中培养菌株以诱导蛋白质产生,其中每个板中包含阳性对照菌株和阴性对照菌株。裂解细胞,并且使用本文描述的酶特异性分光光度测定测量细胞裂解物中的酶活性。在全自动机械工作单元上进行酶测定。对于每个酶家族,在生物学复制品中测量完整的库(各自~1300个成员),并且选择

每个酶家族中具有最高活性的50-200种酶作为该家族的主要“命中”。用附加的复制(4个生物学重复)在二级筛选中重新筛选主要命中以验证酶排序。

[0213] 亮氨酸脱氢酶(LeuDH)

[0214] 首先针对使Leu脱氨基化的能力筛选共计1378种LeuDH酶。最初的一轮筛选鉴别活性类似于或优于来自枯草芽孢杆菌的亲本LeuDH酶的活性的220种酶(表4)。在二级筛选中进一步分析这些主要命中(图2)。在二级筛选中,验证具有对Leu的LeuDH活性多达1.8倍增加的LeuDH酶。

[0215] 活性计算为:酶活性除以背景酶活性减1。对照设置为0,并且具有值>0的菌株被认为是潜在命中。值表示相较于对照的分数改进。作为非限制性实例,具有50%改进的菌株将在表4中以0.5的值表示。

[0216] 为了确定主要LeuDH命中中的任何一种是否展现出相较于Ile和Val针对Leu的增加的特异性,还针对对Val和Ile的活性筛选全部220种主要命中。特异性测量为对Leu的活性与对Ile或Val的活性的比率。如图3中所示,从初级筛选中命中的酶展现出相较于Val对Leu高达2.7倍的偏好、以及相较于Ile对Leu的高5倍的偏好。当在该测定中测量时,阳性对照蜡样芽孢杆菌LeuDH显示出对Leu、Val和Ile的同等偏好。

[0217] 在该库中观察到Leu特异性对Leu活性的折衷,其中最具特异性的LeuDH酶不是最具活性的LeuDH酶。通过比较对Leu/Ile的特异性与对Leu/Val的特异性,鉴别了相对于Leu和Val两者对Leu具有增加的特异性的命中(图4)。对照蜡样芽孢杆菌LeuDH展现出对Leu、Val和Ile大约同等的偏好。

[0218] 酮异戊酸脱羧酶(KivD)

[0219] 针对对酮异己酸的脱羧酶活性筛选总计1248种KivD酶。最初的一轮筛选鉴别了活性高于来自金黄色葡萄球菌的亲本KivD酶的活性的55种酶(表5),来自金黄色葡萄球菌的亲本KivD酶在该测定中未展现出大于背景裂解物脱羧酶活性的活性,并且等同于非零可测量的背景活性。在二级筛选中进一步分析这些主要KivD命中(图5)(表5)。在二级筛选中,鉴别了在该测定中相对于背景裂解物活性具有KivD活性的至少6倍至8倍增加的>40种KivD酶。KivD活性计算为:酶活性除以背景酶活性减1。

[0220] 醇脱氢酶(Adh)

[0221] 针对将异戊醛还原为异戊醇的能力筛选共计1215种Adh酶。最初的一轮筛选鉴别了活性高于来自酿酒酵母的亲本ADH2酶的55种酶(表6),来自酿酒酵母的亲本ADH2酶在该测定中未展现出大于背景裂解物醇脱氢酶活性的活性,并且等同于非零可测量的背景活性。由于ADH2酶对酿酒酵母的活性与裂解物的背景活性无法区分,因此使用活性高于背景活性的马(Equus caballus)Adh作为筛选的阳性对照。在二级筛选中进一步分析这些主要命中(图6)(表6)。在二级筛选中,鉴别了相对于背景裂解物活性具有Adh活性的至少20倍增加的5种Adh酶。酿酒酵母的ADH2酶用作二级筛选的对照。Adh活性计算为:酶活性除以背景酶活性减1。

[0222] 实施例3:最高LeuDH候选酶的选择性

[0223] 材料和方法

[0224] LeuDH选择性测定

[0225] 为了测量LeuDH选择性(在存在L-Ile和L-Val的情况下L-Leu的特异性脱氨基化),

将裂解物在裂解缓冲液中稀释四倍，并且将10 μ L/孔新稀释的裂解液压入90 μ L/孔的来自上文的改良测定缓冲液(特征在于0.5mM每种氨基酸(L-亮氨酸、L-异亮氨酸、L-缬氨酸)、200mM甘氨酸、200mM氯化钾和4mM NAD)中。在不同的时间点淬灭反应，并且进行亮氨酸、异亮氨酸和缬氨酸的LC-MS定量。

[0226] 结果

[0227] LeuDH催化Leu、Val和Ile的脱氨基化，并且因此全部底物都具有在底物池混合的体内环境中充当竞争者的可能性。为了更好地预测关于混合底物池的最高LeuDH命中的性能，测量LeuDH酶对Leu的选择性(即，当Leu、Val和Ile全部存在于反应混合物中时，LeuDH对Leu的偏好)。在与HTP筛选类似的细胞裂解物测定中筛选总计21种LeuDH酶，除了反应混合物含有摩尔比为1:1:1的Leu、Val和Ile。在反应混合物中监测Leu、Val和Ile消失的速率。图7示出每种LeuDH酶在反应混合物内Leu、Ile和Val的消耗。当与亲本枯草芽孢杆菌LeuDH相比时，至少10种LeuDH酶显示出相较于Val和Ile对Leu提高的偏好。对于几乎全部的LeuDH酶，显示出对缬氨酸的最小偏好。

[0228] 实施例4：途径酶命中选择和操纵子组装

[0229] 为了改进BCAA途径的整体Leu消耗，针对每个步骤选择相对于亲本酶显示出更优性能的多种酶。对于LeuDH，基于两个标准选择6种命中：相对于Val和Ile，对Leu的酶活性和对Leu的特异性。由于LeuDH选择性分析与操纵子组装并行运行，因此选择性数据集不包括在LeuDH选择中。对于KivD和ADH，基于体外酶活性对每个酶家族选择3种命中。总计12种酶进入最终操纵子设计(表3)。操纵子由依照下列顺序的酶的四种编码序列组成：LeuDH-KivD-Adh-BrnQ。选择用于Leu消耗的优选操纵子，并且如下所述进一步测试。

[0230] 表3. 被选择用于推进操纵子设计的酶

酶	标识符	来源	SEQ ID NO (核酸)	SEQ ID NO (氨基酸)
LeuDH	t160946	<i>Cetobacterium ceti</i>	1	2
LeuDH	t160389	大清膜壳杆菌 (<i>Hymenobacter daecheongensis</i>)	3	4
LeuDH	t160283	薄层菌属 CRA2	5	6
LeuDH	t160434	单胞菌属 SCN 70-307	7	8
[0231]	LeuDH	<i>Candidatus kapabacteria</i> sp. 59-99	9	10
	LeuDH	消化球菌科细菌 CEB3	11	12
	KivD	耳念珠菌	13	14
	KivD	芽孢杆菌属 FJ AT-1801	15	16
	KivD	<i>Erwinia injecta</i>	17	18
	Adh	<i>Tortispora caseinolytica</i> NRRL Y-17797	19	20
	Adh	根瘤菌目细菌 NRL2	21	22
	Adh	柴油食烷菌	23	24

[0232] 实施例5:操纵子测试

[0233] 材料和方法

[0234] 细胞制备

[0235] 将支链氨基酸 (BCAA) 途径操纵子质粒转化到购自德国微生物和细胞培养物保藏中心 (DSMZ Braunschweig, 大肠埃希氏菌 DSM 6601) 的大肠埃希氏菌 Nissle 菌株 1917 中。转化的细胞在冰上解冻, 并且通过 600nm 处的光吸收 (OD_{600}) 测量细胞密度。在该方法中, 假定 1.0 的 OD_{600} 等于 10^9 个细胞/mL。计算体积以便以 1mL 的 2×10^9 个细胞/mL 细胞重悬液为目标, 并且将细胞转移到 96 深孔板中, 并且用冷 PBS 洗涤一次。在离心 (4000 rpm, 4°C, 10 min) 之后, 弃去 PBS, 并且然后将细胞沉淀物重悬于 1mL 的 $1 \times M9 + 50\text{mM MOPS} + 0.5\% \text{葡萄糖 (MMG)}$ 缓冲液中。将八百 (800) μL 的每个样品转移到新的 96 深孔板中, 并且添加 800 μL 的含有 16mM 亮氨酸的 MMG, 通过移液充分混合。此时, 收集指定为时间零的样品 (200 μL)。然后用透气膜覆盖板, 并且移动至厌氧室中以在 37°C 孵育。在厌氧室中孵育期间, 在 2 小时和 4 小时也收集样品。将样品在收集之后立即在 4°C 以 4000 rpm 离心 10 分钟。将 100 μL 的上清液转移到新的 96 孔板中, 并且储存在 -80°C 以用于将来分析。

[0236] 亮氨酸活性测定

[0237] 使用 Ultimate 3000 UHPLC-TSQ Quantum 或 Vanquish UHPLC-TSQ Altis 系统通过液相色谱联用串联质谱 (LC-MS/MS) 在细菌上清液中对亮氨酸进行定量。用 9 份 2:1 乙腈:水

(含有1 μ g/mL亮氨酸-d3作为内标物)提取样品,涡旋,并且离心。用9份0.1%甲酸稀释上清液,并且同时用0.8 μ g/mL至1000 μ g/mL如上处理的标准物进行分析。在Phenomenex Synergi 4um Hydro-RP80A, 75 \times 2mm上使用0.1%甲酸(A)、0.1%甲酸/乙腈(B)以0.3mL/min和50摄氏度分离样品。在2 μ L注射和从0分钟至0.5分钟保持的初始5%B之后,在0.5分钟至1.5分钟内从5%至90%B梯度洗脱分析物,然后是高有机洗步骤和水平衡步骤。在电喷雾正离子模式下使用化合物特异性碰撞诱导的片段的选择反应监测(SRM)来检测分析物(亮氨酸:132>86,异亮氨酸:亮氨酸-d3:135>89)。对SRM色谱图进行积分,并且未知/内标峰面积比用于标准曲线计算浓度。

[0238] 结果

[0239] 将通过HTP筛选鉴别的消耗Leu最高的操纵子转化到大肠埃希氏菌Nissle 1917中(并且标记为菌株5941、菌株5942和菌株5943)中,并且与原型菌株1980进行比较。菌株5941含有Cetobacterium ceti的LeuDH酶、Erwinia iniecta的KivD酶和柴油食烷菌的Adh酶。菌株5942具有Cetobacterium ceti的LeuDH酶、Erwinia iniecta的KivD酶和根瘤菌目细菌NRL2的Adh酶。菌株5943具有Cetobacterium ceti的LeuDH酶、Erwinia iniecta的KivD酶和根瘤菌目细菌NRL2的Adh酶。操纵子还含有大肠埃希氏菌的BrnQ。原型菌株含有蜡样芽孢杆菌LeuDH、乳酸乳球菌KivD、酿酒酵母ADH2、以及大肠埃希氏菌BrnQ。

[0240] 针对Leu消耗对来自消耗Leu最高的操纵子和原型菌株的样品进行分析(图8)。发现含有消耗Leu最高的操纵子的菌株(5941、5942和5943)与原型菌株(1980)相比以显著更快的速率消耗Leu。

[0241] 实施例6:LeuDH酶的工程化和活性LeuDH酶的生物信息学分析

[0242] 如表4中所示,生成来自蜡样芽孢杆菌的UniProt P0A392(SEQ ID NO:27)的突变体,并且测试以确定突变体是否相对于UniProt P0A392(SEQ ID NO:27)显示出提高的活性或酶表达。使用实施例2中描述的LeuDH活性测定。观察到下列独特位置处的点突变提高活性或酶表达:42、43、44、67、71、76、78、113、115、116、136、293、296、297和300。

[0243] 观察到UniProt P0A392(SEQ ID NO:27)中的下列点突变提高活性或蛋白质表达:A115N、A115Q、A115S、A115T、A115V、A297C、A297D、A297E、A297F、A297H、A297K、A297L、A297M、A297N、A297Q、A297R、A297T、A297W、A297Y、E116A、E116L、E116M、E116N、E116R、E116S、E116V、E116W、G43E、G43F、G43T、G43W、G43Y、G44H、G44I、G44K、G44Y、I113F、I113M、I113Q、I113V、I113W、I113Y、L300A、L300C、L300D、L300F、L300H、L300K、L300M、L300N、L300Q、L300R、L300S、L300T、L300W、L300Y、L42A、L42Q、L42T、L76E、L76F、L76H、L76I、L76K、L76M、L76R、L76S、L76T、L76W、L76Y、L78C、L78F、L78H、L78K、L78Q、L78V、L78Y、M67A、M67E、M67K、M67Q、M67S、M67T、N71C、N71D、N71H、N71K、N71M、N71T、T136E、T136F、T136L、T136R、T136S、T136Y、V293A、V293C、V293Q、V293S、V293T、V296A、V296C、V296E、V296I、V296K、V296L、V296N、V296S和V296T。

[0244] 对SEQ ID NO:27的突变体和来自宏基因组库的命中的序列进行生物信息学分析。在下文表7中提供在命中中发现的独特残基的列表。示出SEQ ID NO:27中的相应位置。命中是相对于SEQ ID NO:27具有增加的活性(大于0)的LeuDH。对于多序列比对中的每个位置,将个体残基一致性归入命中和非命中,并且计算集合差。这些是经由系统点突变库或宏基因组序列对命中集而言独有的残基。

[0245] 实施例7:活性KivD酶的生物信息学分析

[0246] 对相对于SEQ ID NO:29显示出增加的活性的命中KivD酶进行生物信息学分析。表8中提供命中中发现的独特残基的列表。对于多序列比对中的每个位置,将个体残基一致性归入命中和非命中,并且计算集合差。这些是对命中集而言独有的残基。表8中指示SEQ ID NO:29中的相应位置。

[0247] 来自乳酸乳球菌的UniProt Q684J7是在酪乳和奶酪的生产中广泛使用的微生物。虽然不是天然酶的命名反应,但KivD催化4-甲基-2-氧代戊酸酯的脱羧基化,以形成异戊醇。发现来自KivD酶库的命中具有超过其天然底物(其天然底物是 α -酮异戊酸)的扩大的底物特异性。

[0248] 实施例8:活性ADH酶的生物信息学分析

[0249] 对相对于SEQ ID NO:31显示出增加的活性的命中ADH酶进行生物信息学分析。表9中提供命中中发现的独特残基的列表。对于多序列比对中的每个位置,将个体残基一致性归入命中和非命中,并且计算集合差。这些是对命中集而言独有的残基。表9中指示SEQ ID NO:31中的相应位置。

[0250] 实施例9:异戊醇途径的摩尔平衡闭合(Molar Balance Closure)

[0251] 在AMBR®15生物反应器中评估菌株5941中异戊醇途径的性能和摩尔平衡闭合。菌株5941包括SEQ ID NO:2的LeuD酶、SEQ ID NO:18的KivD酶和SEQ ID NO:24的Adh酶。用具有0.5%葡萄糖、10mM Leu、10mM Val和5mM Ile的M9培养基将反应器填充至17mL。将条件控制为0%溶解氧和7.0的pH。接种激活的生物质以达到1的OD600,并且随时间采集上清液样品以监测代谢物浓度。

[0252] 图10中示出途径中间产物的细胞外浓度分布。在180分钟的过程中,消耗4.1±0.3mM的亮氨酸,并且在培养基中积聚4.4±0.5mM的异戊醇。在上清液中未观察到酮酸(2-氧代异己酸)和醛(异戊醛)。因此,通过途径的通量被平衡,并且被考虑。途径中间产物的总摩尔数的守恒(与图10中的“加和”对应的数据)也证明了这一点。

[0253] 方法-发酵

[0254] 在来自Sartorius的AMBR15f微生物反应器系统中进行测定。容器填充有17mL补充有2.0mM MgSO₄、0.1mM CaCl₂、5%葡萄糖、10mM L-亮氨酸、5mM L-异亮氨酸和10mM缬氨酸的1×M9培养基盐。在接种前18小时填充容器,以使pH光极和DO光极两者水合。将反应器中的温度保持在37°C,使用2N NaOH将pH值保持在7,并且使用0.14vvm N₂流速将溶解氧保持在0。搅动设置为500RPM,以实现在整个实验中的良好混合。由Synlogic提供的激活的生物质接种生物反应器以达到1的OD600。在接种后0分钟、30分钟、90分钟、150分钟和180分钟对生物反应器取样。立即将样品在微量离心机中以15000×g离心30秒,并且去除上清液以用于分析。将上清液储存在-20°C直至准备好进行分析。

[0255] 方法-分析

[0256] 开发分析以用于两种方法。一种方法涉及用于亮氨酸(Leu)、酮异己酸(Leu酸)和异戊醛(Leu醛)的定量的液相色谱质谱(LCMS)。该方法也经验证,并且用于缬氨酸和异亮氨酸(及其各自的酸产物和醛产物)的定量。第二种方法涉及用于异戊醇(Leu醇)的定量的气相色谱质谱(GCMS)。总之,这些分析方法允许对菌株5941的全部途径中间产物的定量。GCMS方法也经验证,并且用于缬氨酸醇产物和异亮氨酸醇产物的定量。

[0257] 在具有Thermo Q-Exactive四极杆-orbitrap质谱检测器和Thermo Accucore PFP柱(2.1×100mm, 2.6μm填料)的Thermo Ultimate 3000 UPLC系统上使用下列洗脱溶剂进行LCMS分析:A=0.1%甲酸和水中0.1%TFA;B=乙腈中0.1%甲酸。梯度为0.5mL/min的A中1%B达60秒,然后是在270秒内从A中1%B至A中40%B的线性直线上升。然后,用A中95%B冲洗柱达60秒,并且用A中1%重新平衡达180秒。MS采集为从0.8分钟至5.3分钟。

[0258] 经由具有+3800V的正模式电离、400℃的汽化器温度、375℃的离子迁移管温度的标准Thermo ESI源将柱流出物引入质谱仪中。Thermo以任意单位报告气体流速(在STP下可能接近L/min)。设定点为:鞘流气,60;辅助气,30;吹扫气,1。为了增加数据采集率,将静电场轨道阱分辨率设置为17500。四极杆分辨率为1m/z。

[0259] 该方法还衍生醛和酮酸两者,从而提高那些分析物的稳定性。探索多种衍生剂,并且发现甲醇中的2-(二甲基氨基)乙基阱在正模式下引起最佳灵敏度。甲醇中0.5M乙酸和0.5M乙酸钠的缓冲液用于LEU酸和LEU醛的定量,同时也测量未经衍生化的LEU。

[0260] 使用J&W DB-WAX GC柱(15m)和作为提取溶剂的氯仿在具有Gerstel自动进样器的Agilent GCMS/MSD上进行GC-MS分析。将前部注射器设置在250℃和1mL/min的流速。烘箱温度在40℃保持1分钟,然后直线上升至130℃(15℃/min),并且然后直线上升至200℃(65℃/min)。MS采集扫描窗口在40-150mz,其中MS源和MS四极杆分别在250C和200C。

[0261] 为了便利高通量和自动化,Gerstel自动进样器用于将提取的底部氯仿层注入在顶部具有充当覆盖物以防止产物蒸发的水性ambr15培养基的96孔板形式中。为了考虑任何其他潜在的醇产物蒸发,向氯仿中添加2-庚醇作为内部物。

[0262] 表3中的酶的序列

[0263] LeuD (标识符:t160946; 登录号:AOA1T4PGG9)

ATGAACATCTCAAGAAAATGGAGGAATTAAATTATGAACAACTGGCTACTTCTACGACAGCGAACGGAACTC
AAAGGTATTACCTGTATACACAACACAACCTTAGGGCCGGCATGGCGTACCCGCTTGGAACATAACTCT
GAGGAAGATGCCGTTGAAGACGTAATCCGCTGGCTGGCATGACTTACAAAGCGGCTTGCGCGGTCTGAAT
CTGGCGGCGGTAACCGTGCTGATCGGTGATGCTAAAAGATTAATCAGAGTCCTACTTCCGTGGACTGGGG
CGCTACGTTAGTCGCTAACGGCAGATATATCACCGCGGAAGACGTAATAACTTCTACGAAGGATATGGCATA
GTTGCTATGAAACTGACTATGTGGTAGGCCTGGGAGGTAAATCCGCAACCCTAGTCCAGTTACTGCTTACGGT
GCATTTATGGGTATCAAAGCGCGCTGATGAAAAAATTGAGGATAGCTTACTTGAAGGCCAACCTTCGCACTG
CAGGGTGTGGCAGACGGTTACTATCTTATCGATTACCTCTAGGAAACAACAAGTTCAAAGAAAAGGCTAAA
AAAATTACTTCACCAGAAATTACGAGAGCTATATCGAGCGTATGAACAAAGAACATCCGGAAGTTGAATTATT
TCCCCGGACAAATCTACTCGCTGGAGTAGACGTCTCGCCCTGCGCCCTGGCAAATCGTTAATGACAAA
ACTATCGATGAATTAAAGTGTCCGATCATCGCAGGTACTGCAACACAGTACTGGAAAGGGAGCGCACGCAAC
ATGCTTAAAGAACGTGGCATTCTTACGCCGGACTATGTGATCAATGCTGGCTGATCAACGTTTACAC
GAGCTGAACGGTTACAATAAGAGAACGCTATTCTGGAAGTGGAAATTATGATCGCTACTGGAAATATT
AACATCGCTGATTCTGAACATCAGCACCAATATCGCTGCCAACGAGTTCGCGAAAAACGTATCAAGCAAATT
AAGTCCTTGAACAACTCATTAAACGC (**SEQ ID NO: 1**)

MNIFKKMEEFNYEQLVYFYDSETELKGITCIHNTTLGPALGGTRILWNYNSEDAVEDVIRLARGMTYKAACAGLN
LGGGKTVLIGDAKKIKSESYSFRGLGRYVQSLNGRYITAEDVNTSTKDMAYVAMETDYVVLGGKSGNPSPVTAYG
AFMGIKAALMKKFEDSSIEGRTFAVQGAGQTGYYLIDYLLGNNKFKEAKKIYFTEINESYIERMNKEHPEVEFI
SPDKIYSLEVDFVPCALGKIVNDKTIDEFKCPIIAGTANNVLEREAHGNMLKERGILYAPDVINAGGLINVYH
ELNGYNKENAILEVELIYDRLLIEFNIADSLNISTNIAANEFAEKRIKQIKSLKNNFIKR (**SEQ ID NO: 2**)

[0265] LeuD (标识符:t160389; 登录号:AOA1M6BE59)

[0266]

ATGGTAGAGATCAAGGCTTGACGGACACTCCGTGTTGGCAAATTGAGAACACCAGCATGAACAGGTGCGTT
 TTCTGCCACGATCACGAAACCGGCCCTCCGTGCGATCATCGTATTATAACACAGTTCTGGCCCCGCCTAGGT
 GGAACTCGCATGTGGCACTATGCTCTGACGAGGCGCTGAATGATGTTCTGCGCTGTCGCGCGGTATGACC
 TACAAAGCTGCTATAAGTGGCTGAACCTGGTGGCGTAAAGCAGTGTACATTGGGACGCCAAAACCTGAAA
 ACCGAAGCGCTGCTGCGAAGTTCGGCAGATTCTGAAAAACCTGAATGGTAAATACATCACTGCTGAAGATGTC
 AACATGACTACAAAAGACATGGAGTACATCAGGATGGAAACCAAGCACGTTGCTGGCTTACCTGAATCAATGGGT
 GGAAGCGGTGATCCGTCCCCGGTGAUTGCATTGGTACGTATGGGATGAAAGCGGCCAAAAGCGTTC
 GGCTCTGACTCTCTGGCTGGCAAACGTATCGCTGTTCAAGGGTGTAGGTATGTCGCACTTACCTGTTGGAGTAT
 TTGAGAAGGAAGGTGCTAAGCTGGTACTGACTGACTATGAAGATCGTGCCTGGAGGCAGCAACGCGTTT
 GCGCAAAATGGTTGGCCTGGACGAAATTACGATCAAGACGTTGATATCTACAGTCCATGTGCTTGGAGCT
 ACCATTAACGATGACACTATCGGTCGCTGAATGCCAGGTTATCGCTGGTGCACAAACCAGCTGCAAAAC
 GAAAATGTGATGGCCGGCCCTCGTGGAGCGGGATTGTCAGCCTCCGGATTCCCTGATCAACGCCGGC
 CTGATCAACGTTTACTCGGAAGTAGTGGTAGCTCCGTCAAGGGTCTTGAACCAGACGAAAAAATTTCGAC
 ATCACCACTCAGGTTCTAACAAAGCGGAACAAGAGGGTCTCACCCGCAAGCGCAGCTACTAAGCAGGTGAA
 GAGCGTATTGCAAGCCTGGCAAAGTTAAGAGCACCTAC (SEQ ID NO: 3)

MVEIKALTDTSVFGQIAEHQHEQVFCHDHETGLRAIIGIHNTVLGPALGGTRMWHYASDAEALNDVRLSRGMT
 YKAAISGLNLGGGKAVIIGDAKTLKTEALLRKFRFVNLNQKYITAEDVNMTTKDMEYIRMETKHVAGLPESMG
 GSGDPSPTAFGTYMGKAAAKKAFGDSLAKRIAVQGVGHVGTYLLEYLQKEGAKLVLTDYYEDRAEAATRF
 GAKMVLDEIYDQDVDIYSPCALGATINDDTIGRLKCQVIAGCANNQLQNENVHGPALVERGIVYAPDFLINAGG
 LINVYSEVVGSSRQGALNQTEKIFDITTVQVLNAEQEGSHPQAAATKQAEERIASLGKVKSTY (SEQ ID NO:
 4)

[0267]

LeuDH(标识符:t160283;登录号:A0A1S9B636)

ATGGTAGAGATCCAGGCTTGCCGAAACTCCATTGGCAAATCGCAGACCACCGCATGAACAGGTGGTC
 TTCTGCCACGATCACGAAACCGGCCCTCCGTGCGATAATCGTATTCTACGTTCTGGCCCCGCCTAGGT
 GGAACTCGCATGTGGCACTATGCTACCGAGGCGAGAACGCTGAATGACGTTCTGCGCTGTCGCGGTATGACC
 TACAAAGCTGCTATCTGGCCCTGAACCTGGTGGCGTAAAGCAGTAAATCATTGGGATGCCAAAACATCAAA
 ACCGAAGCGCTGCTGCGAAATTCCGGCAGATTCTGTCAGAACCTGAATGGTAAATACATCACTGCTGAAGACGTT
 AACATGACTACAAAGGATATGGAGTACATTAGGATGGAAACCAAACACGTCGCTGGCTTACCTGAAAGTATGGGT
 GGAAGCGGTGACCCGTACCGGTAACTGCATATGGTACGTACATGGGATGAAAGCGGCCAAAAGCGTT
 GGCTCTGATTCCTGGCTGGCAAACGTATCGCTGTTCAAGGTGTGGTACATGTTGGCACTTATCTGCTTGGACAT
 TTGACCAAAGAAGGTGCTCAGATTGTGCTGACTACTATAAGGAACGTCGCCAGGAAGCAGGCCGGCTTT
 GCGCACAGGTTGTTGGCTGGACGATATCTACGATCAAGAGGTCGACATTACTCTCATGTGCTCTCGGTGCT
 ACCATCAACGATGACACTATCGATCGCCTGCGCTGCGCTGTTGAGCCGGTGCACAAACACCAGCTGAAAGAA
 GAAAACGTCCACGGTCCGGCTGGTGGAGCGCGGGATAGTATACGCCCGACTCTGATCAATGCAGGTGGC
 CTGATTAACGTGTATAGCGAAGTTACAGGGTCTACCCGTCAAGGGGCTTTAACTCAGACGAAAAATCTATGAC
 TACACACTCCAAGTTCTGAAAAAGCCGCGGCTGAAGGTCTGCACCCCGCAGCAGGCTGCGATCCGTAGCGGAA
 CAACGCATCGCTGCAATTGGTAAGGTGAAAGCACCTAC (SEQ ID NO: 5)

MVEIQALPETSIFGQIADHQHEQVFCHDHETGLRAIIGIHNTVLGPALGGTRMWHYATEAEALNDVRLSRGMT
 YKAAISGLNLGGGKAVIIGDAKTIKTEALLRKFRFVNLNQKYITAEDVNMTTKDMEYIRMETKHVAGLPESMG
 GSGDPSPTAYGTYMGKAAAKKAFGDSLAKRIAVQGVGHVGTYLLEHLTKEGAQIVLTDYYKERAEEAGARF
 GAQVVGLDIYDQEVDIYSPCALGATINDDTIDRLRCAVVAAGCANNQLKEENVHGPALVERGIVYAPDFLINAGG
 LINVYSEVTGSTRQGALTQTEKIYDYLQVLEAAAEGLHPQQAAIRQAEQRIAIGKVKSTY (SEQ ID NO:
 6)

[0269]

LeuDH(标识符:t160434;登录号:A0A1D2RXB2)

[0270]

ATGATCTTCGAGACAATTCTACGTCGAATCACGAAGAAGTTGTATTGCCATAACAAGGACGCCGGCTTGAAA
 GCAATCATCGCATTACAACACTGTAUTCGGCCGGCTCTGGGTGGCACTCGCATGTGGCCCTACGCTAGCGAA
 GAGGAAGCACTGAAAGATGTCCTCGTTATCCCCTGGGATGACCTACAAAGCTGCCGTTTCAAGGCTAAACCTG
 GGCAGCCTGGGAAAGCTGTGATCTGGGTGATCCGAAATAAGACAAGCTGAAGCGCTGTTAGACGCTTGGACGG
 TTTGTAAACAGCCTGGGCGGACGCTACATTACCGCGAGGACGTTGGATTGATGTTAACGACATGGAATATGTG
 CTGCGTCAAACACTGATTACGTCACCGGTGATCACAGGTTCACGGTGGAGTGGTGTACCTTCCATTACCGCA
 TATGGCACTCTGCAAGGCCGTGATGGCGCTCTGCAAGTGAATTGCGTAACGAAGACGTAGGCAATTACAGCTAC
 GCTGTTCAAGGTTGTCAGGTCAGCTGGATGAAATTGTTAAACTGCTGCGTGAGCGCGGTGCAAAGGTTTCGTC
 ACTGACATCAACAAAGATGCGGTCCAGCGTGTGGACGAAATTGGTTGAGGCACTGGATGAAATC
 TATGACGTTGATTGCGACGTGTAUTCCCCGACCCTGCGGACCGTGAACGATAAAACTTACCGCTCTG
 AAATGTAAGGTAATCTGCGGTGCGGCAAACAAACCAAGCTTAGCTAATGATGAGATAGCGTGGAACTGGAAAAAAA
 GGCATCCTCTATGCTCCGACTACCGCGTCAACGCGGGTGGCTGATGAACTGCTGGAAATCGATGGATAC
 AACCGCGAACGTCGATCGTATGCGTACCTTATTACAATTGGTGTGCAATTTCGAAATCTCTAACGCG
 GACGGCATCCCTACATCCGAGCCGCCGATCGTATGGCTGAAAGAACGCAATAACGCCATCGTAAACTGCGTTA
 CGCATTGGCGCTGGCACCGCGCTTCCAGGGCCGACGTGGCAAC (SEQ ID NO: 7)

[0271]

LeuDH (标识符:t160048)

ATGCAGATCTCGACACTTGCAATCAATGGGCCATGAGCAGGTGGCCTATGTAGCGATAAGACCACGGCTG
 CGCGCATTATCGCTATACAGATACTCCCTAGGGCGGCGCTGGTACCCGTATGTGGCAGTATGCAACT
 GACGACGATGCTATTACTGACGCACTCCGTCTGTCTCGGGCATGACCTACAAAGCTGCCGTTCTGGCGTAAAT
 CTGGCGGTGGTAAAGCCGTTATCATCGGAAACCCCTCACAGTGTAAAGCGAAGCGTGTGGCCTACGGC
 AGAATGGTGAATCCAGCGTGGCGTTACATACCGCCGAAGACGCTGGTACTAGCGTACGTGATATGGAGTGG
 ATTGCGATGAAACCAAATATGTAACGGCGTGGCAACGGAGGCTGGTACCCCTCTCCAGTTACCGCT
 CTGGGTGTTACTCGGCATGAAGGCATGCGCTAAATCAGTCTATGGTACTGATGCGCTGAGCGTAAAGGATC
 GTGGTTAGGGCGGGTAACGTTGCATCCATCTGGTACAGTCTGGTAAAGAAGGCGTGTGGTTTCGTC
 ACTGACATCTACGAAGAAAAGCCAAAGCATTAGCGGCTGAAACGGCGCTACCGTGTGATTCGACCGAGGTT
 TTACTACACAATGCGATATCTCTCCGAAAGCTCTGGGGCGCTCTGAACGATGAAACTATTCCGCAGCTC
 ACATGCGCTATCGTAGCTGGTGCACAAACATCAGCTAAACATGAAACACGTACGCCACGGCTCTGCAAGAG
 AAAGGCATTCTGTATGCGCCGGATTACGTAATCAACGCCGGGCTCATGAATGTGGCAGCGAAGTTGACGGC
 TACAACCGTAAAAGGTATGCGCAGGCTGAAGGTATTACGATATTACTATGACATCTAAACCGCGT
 GAGCGTAAACATCCTGACCATCGAACGATCCAACCGCATTGCTGAAGAGCGGATCAACAAAGTCGCCATGTCAC
 GGGAACTTCATCGGTTCCCCGTCTATCGCGGAGTA (SEQ ID NO: 9)

[0272]

MQIFDTLQSMGHEQVVLCSDKTTGLRAlIAIHDTSLGPALGGTRMWQYATDDDAITDALRLSRGMTYKAAVSGVN
 LGGGKAVIIGNPHSDKSEALFRAYGRMVESQRGRYITAEDVGTSVRDMEWIRMETKYVTGVGGNGGSDPSPVTA
 LGVYSGMKACAKSVYGTDALSGKRIIVQGAGNVASHLVHSLVKEAVVFVTDIYEKAKALAETGATVIRTDEV
 FTTQCDIFSPNALGAVLNDETI PQLTCAIVAGGANNQLKIEQRHATALQEKGILYAPDYVINAGGLMVASEVDG
 YNREKVMRQAEGIYDITMNI LNTARERNILTIEASNAIAEERINKVRVHGNFIGSPSIRGV (SEQ ID NO:
 10)

[0273]

LeuDH (标识符:t160141; 登录号:AOA0J1FEE3)

[0274] ATGACAACGTTGAGTATATGGAAAAGTACGACTACGAACAACTGGCCTTGTCAGGATAACACTTCTGGCCTC
AAAGCAGTAATTGCATCCATGACACCACTCTGGGCCAGCTTGGTGGCACCCGTATGTGAAATTACGCCAGT
GAAGAAGATGCTATCCTGGATGCCTACGCCTGGCGAGGTATGACTTATAAAACGCTGCCAGGTCTGAAC
CTGGCGGCGGTAAAGCTGTATTATGGCGACAGCCGTACCCAGAAATCAGAGGAACGTGTTCGCGCTTCGGT
CGTACGTGCAGGCCTAACGCCGTATATCACCGCTGAGGACGTTGACTAACGTACAAGATATGGACTGG
ATACACATGGAAACAAAGTTGTGACCGGGATCTCCTCTTCGTACGGTCTAGCGAGATCCGCCCCCTCGACC
GCACTGGCGTTACCGCGTATGAAAGCGCCCAAAGAAGCCTGGCAGCGACTCTTAGAGGGTAAAG
GTTGCTATTCACTGGGTCTGGCCACGTCGGCTATTACCTGGCAAACACCTCACTGATGAAGGCCTAAACTGATC
GTGACGGATATCAATTCTGAAGCCGTTAAGAGGGTAGCGCGTAGGTTGCTACCGCAGTCCGTACCGAAGAA
ATTTTGGCGTTAAATGCGACATCTTGCGCCGTGCTCTGGGTGCAAGTTATCAACGATGAAACCATTCCGCAG
CTGAAGTCCAGGTAGTTGCCGTGCTGCGAACATGTGTTGAAAGAGGATGCCATGGTACGAACATACGAA
AAAGGAATCCTGTACGCTCCGGACTATGTAATTAAACGCGCCGGCTTATCAACGTTGCGACGAACGGAAAGGT
TACAACGCTGAACGTGCTCTGAAAAAAGTTGAGATGGTATATGATAATGTGGCACCGCTCATCGTATTGCAAG
CGTACCATATCCGACTTATAAAGCAGCGAACGATGGTGGAGGAACGTATTGCGAAATTGCGAAAGTTCC
AACACTTCTCGC (SEQ ID NO: 11)

MTTFEYMEKYDYEQLVLCQDNTSGLKAVICIHDTLGPALGGTRMWNYASEEDAILDALRLARGMTYKNAAAGLN
LGGGKAVIMGDSRTQKSEELFRAFGRYVQALNNGRYITAEDVGTNVQDMDWIHMETKFVTGISSSYGASGDPSPLT
ALGVYRGMKAAAKEAFGSDSLEGKTVAIQGLGHVGYYLAKHLDEGAKLIVTDINSEAVKRVAREFVATAVRTEE
IFGVKCDIFAPCALGAVINDETIPQLKCQVVAGAANNVLKEDRHGDELYEKGILYAPDYVINAGGVINVADELEG
YNAERALKKVEMVYDNVARVIAIAKRDHIPTYKAADRMAERIARIKIGKVSNTFLR (SEQ ID NO: 12)

[0275] KivD(标识符:t163988;登录号:A0AOLOP8D8)

ATGTCGGAGATCACATTGGTAGATACCTTTGCAACGCTAAACCAACTGCAAGTGCAGACTATTTTGGCTG
CCCGCGACTTCAATCTGCCCTGCGATAAGATCTATGAAAGTTGATGGCATGCCATGGTGGCAGGTAACGCTAAC
GAACCTCAACGCCGTTACGCCGGCTGACGGTTATAGCCGTGCAAAGGCCCTCGCATGTCTGGTACCGACTTTGGT
GTAGGCGAGCTAAGTGCCTGAATGGTGTGGTGGCGCTTACGCAGAACAGCTGCTGCACCACCCGGTAACGGAGATT
GTCACCATCTCTAGCCAGCGAACAGCTGCTGCACCACCCGGTTATCAGCGACATTAATTCTGCTCTGGTAAATCGAT
AGGTGCATCCGTGAGGCTGGTACATCAGCGCCGGTTACGTCGGCTGCCGGAACCTAGTTGACCTGACT
GTGCCGGCGTCTCTGTTAGACACTCCGATCGATCTGCTCTGAAAAAAACGACCCGGATGCCAGGAAGAGTT
ATTGAAACCGCTCTGATCTGGTAGACAAGTCTAAACCCCTATAATCTTAGTTGACGCATGCCAGCTAGCCGTAC
TCATGCCCGCATGAAGTACGCCGGTTGGTGGACTCCACCAGCTTCCCGGTTCTGTTACTCCAATGGTAAATCT
GCTGAAATGAGAGTACCCCGCTTGGCGGTGTTACGTCGGCAGCCTCAGCGAGCCAAACGTAAGAACGCC
GTGAAACCGCTGACCTGGTGTGCTGCCATAGGCCCTGTTGAGCGACTTCACACTGGATCGTCTCTTATTCC
TACAAAACTAAGAACATGTTGAATTCTACTCTGATTATACCAAAATCCGTCAGCAACGTTCCCGGGTTTCAG
ATGAAAAGAAGCACTGAATGTCCTGTTGGAAAAAAATCCGAGCCATGTCGTAACCTACAAACCTGCGGGTTCCG
CAGCGTCGCCATTCCGAGGCCAGGGGATAAGGCTGCGATCTCAGGAGTGGCTGTGGTCCGCTGCTAG
TGGTCCGCGAGGGCGACATCGTCATTACAGAAACCGGTACCGAGTGGCTTGGAAATTGTCAGCTCCTATTCC
GATAACTGCATCGGCATCAGTCAGGTGCTGCTGGGTTGATCGGCTTCACCGTAGGTGCAACGCTGGCGGGTG
ATGGCTGCACAAAGAACATCCGAAAAACACTCCCTACCTGTTGCTGAAACAGTGGTACACTATC
CAGGAATTTCTACCATGGTTAAGTGGAAACCACTCCCTACCTGTTGCTGAAACAGTGGTACACTATC
GAACGCCCTATCCATGGCGAGACTGCTACGTATAACGATATTAGCCGGTGGGATAATCTGGTCTGTTGCCGTG
TTCAAGCTCGTACTGAACCAACCGAGTGGACTTAGGCGAAATTGAAGCGTATTCAACAAATTCA
TTCAATGAGAATACAAAGATCCGTATGGTGGAGGTGATGTCGCCGCGATGGTACACCACAGAACCTGGTTAA
CAGGCTGAATTTCCTCCAAGACCAACAGCGAAAC (SEQ ID NO: 13)

MSEITLGRYLFERLNQLQVQTIFGLPGDFNLSLLDKIYEVDGMRWAGNNELNAAAYAADGYSRVKGLACLVTTFG
VGELSALNGVGGAYAEHVGLLHVVGVPISISSQAKQLLHHTLGNDFTVFHRMSNNISQTTAFISDINSAPGEID
RCIREAWVHQRPVYVGLPANLVDLTVPPASLLDTPIDLSSLKNDPDAQEEVIETVLVDLVDKSKNPIILVDACASRH
SCRDEVRLVDSTSFPVFTPMGKSAVNESHPRFGGVYVGSLSPEPNVKEAVENADLVLSIGALLSDFNTGSFSYS
YTKTNIVEFHSDYTKIRQATFPGVQMKEALNVILLEKIPSHVANYKPLPVQRRVIPSPGDKAISQEWLWSRLSS
WFREGDIVITEGTSAGFIVQSYFPDNCIGISQVLWGSIGFTVGATLGAVMAAQEIIDPKKRVILFVGDSLQLTV
QEISTMVWKETTPYLFVLNNNDGYTIERLIHGETATYNDIQPWDNLGLPLFKARDYETNRVATVGEIEALFNNSA
FNENTKIRMVEVMLPRMDAPQNLVKQAEFSSKTNSEN (SEQ ID NO: 14)

[0277] KivD(标识符:t164076;登录号:A0AO0M5JJZ2)

ATGACAAGCATGGACAATTCTAGTCAGCAAATCCCCATGGGTAGAAAACCGTCGGGAGTACTTGTTGATTGC
 CTCAGCAGGAAGGCATAACGGAAATCTTGGTGTGCCGGCATTATAACTTCACCTTACTGGACGCCCTGCAA
 GAATACAACGGTATTCTGTTCTATAACGGCCGAACGAGCTGAATGCTGGCTACGAGCTGACGGTTACGCGCGT
 ATAAAGGAATCTCGCGCTAATCACTACTTTGGTGTGGTAACGTACGAACTAACGCTATTGCCGGCG
 AACAGCGAACACGTACCTATCATCCATTGTTGGTCCCACCGGAAAAGCTCAGAAGGAGCGCAAACGTGATG
 CACCATAACCTGATGGATGGCAACTTCGACGTATTCCGTAAGTTACGAACCGCTTACCGTTATACTACCATC
 GTCACGGCAGATAACGCCGGATGGAGATCCGGCTGCTATCCGTATTGCCAAAGAACGAAGAAAGCCAGTGTAC
 CTGGTTGTTGCGGATGACGTAGTGGCTAAACCGATTACTGGTCGTGAAGTCCCGCATCTCCTCTGCCGGTAGC
 AATCAGGACAAACTGCTTGTGCGGTTGAGCACGTTAGGCGTCTCTGGAACCTGCACGCCAGCCGGTAATATTG
 GTTGATGTGAAAGCCATGCGCTTGGATTACAGACGCCGTACAGGAACACTATGAATGTTCCAGTG
 GCTACAATGATGATGGAAAGGCACCTTCGACGAAACCATCCAAACTACATCGCGTATATGCCGGTAGTTC
 GGGTCTGAGTTCAATCTATCGTAGAAAACCTGGACTGTGTTATCGCGTTGGTTGAGGAGCAGAAAGCTGAAACTCGCT
 AACCCGCAAACCTTACTCGAAATTAAACCCGACAATACCATTGAGGTTAGCCGACAAAGCTGAGTGGATTATCGTAGCCAGTCT
 GAGTCCCAGTACCCGATGTCGTGCCGAGACATCCTGCAAGAAATGAGCAGCTGGATTATCGTAGCCAGTCT
 AAACCGGAAAAATCTCATTCCGTACGAGAGATAACCGGGTCCAGTGATGAACCGCTCCGCCAGAAAACCTAC
 TTCCCTCGTTTCAGCGCATGCTGAAGGAAAAGCATATTGTTATCGCTGAGACCCGACGTTCTACTACGGTATG
 AGTCAAGTTAAACTGCCCGAACACTACGTAACATCATGAGGGCGGTGGCAGAGCATTGGTTATGCCACCCCG
 GCGGCATACGGCGCTATCGCTCCGGACCGTCGCTACTGTTACTGGTATGGCTCCATGCAGCTG
 ACCGCACAGGAAATCTTCTATGCTTATTACGGTTGCAAGCCGATTATCTTGTACTGAACATGACGGTAC
 ACCATTGAGCGGTATCTGAATGAGAAATCTCCCTGACGAAACAAAACATACGATATTCCGAACTGGCTTAT
 ACTAAACTGGCTGAGGCGTTGGTGAACGTGTTACTAAACAGTGCCTACCAATGAAGAATTGGATGAAGCG
 ATCACACAGGCTGAGCAAGAGTACGCCGAAAACGTGCTGATCGAGATGATTGCTGATCCAATGGACGCA
 CGGAAATACATGCACCGTATCCGTAACCATAACGAGGAACAGAAAAG (**SEQ ID NO: 15**)

MTSMNNSQQIPMGQKTVGEYLFDCLKQEGITEIFGVPGDYNFTLLDALQEYNGIRFYNGRNELNAGYAADGYAR
 IKGISALITTFGVGELESATNAIAGANSEHVPITHIVSPPEKAQKERKLMHHTMDGNFDVFRKVYEPLTAYTTI
 [0279] VTADNARMEI PAAIRIAKERRKPVYLVVADDVVAKPITGREVPASPLPASNQDKLIAAVEHVRRLLEPARQPVIL
 VDVKAMRFGLQTAVRELANTMNPVATMMYKGKFDETHPNYIGVYAGTFSSEVQSIVENSDCVIAVGLWSDT
 NTANFTAKLNPHNTIEVQPTKVKIAESQYPDVRAADILQEMQKLDRYSQSKPEKISFPYEEITGSSDEPLAENY
 FPRFQRMLKENDIVIAETGTFYYGMSQVKLPANTTYIMQGGWQSIGYATPAAYGASIAAPDRVLLFTGDGSQL
 TAEQEISSMLYYGCKPIIFVLNNNDGYTERYLNEISPDEQNYNDIPNWSYTKLAEAFGGELFTKTVRTNEELDEA
 ITQAEQEYAEKLCLIEMIAADPMADAPEYMHRIRNHKQEQQK (**SEQ ID NO: 16**)

[0280] KivD(标识符:t163842;登录号:A0A0L7TB96)

ATGTCGACGACAACCGTGGTACTACTTGCTGTATCGCTAAACGAAATCGCATTGAGCACCTTCGGAGTG
 CCAGGTGATTACAATCTGCAATTCTGGATCATGTAATCGACCACCCCTCAGCTGACTTGGGTCGGCTGCACTAAC
 GAACTTAACGCTGCCACGCTGATGGTTATGCGCCTGCTGAGCTGGCTGCCGACTGCTGACCACCTTCGGG
 GTTGGCGAAGTGGCTATTAAATGGCATCGCAGGTTCTACGCGGAGTATCTGCCGTAATACATATCGTGGT
 GCACCGAGTCTATCAGCCAGCAGCAGGGCAGCTGATTCCACACTCTCTGGCGAAGGTGATTTCAGCTTC
 CTGAGGATGTCCTACCGGGTGTCTGCGCAGGCTGCTGACTCTGATAACGCACTGCAAGGAAATGACCGC
 GTACTGGCGGAAGTCCTCATTAGCGCTCGTCCGGTACCTGCTGCTGCTACCGACGTTCTGGCTGCTGCCGGCG
 GCTCTGCCACAAAGCACTCTTCTTGGCAGCCGGGATCATCGCCAGTTCTGGCTGCTTCAGCGACGCT
 GCTGAGCAGATGCTGGCTCAGGCCAAAGCGCTCTACTGGCGGACTTCTGGCTGATCGTTTCGGTGTACT
 CGAGCAACCGGGTCTGTTGGTACCTACGCTGGCGCATCTGATAACGCACTGCGCAGTTCTCGCAGATCGACGCC
 GACAGCAACCCGCTGATCAACTGTGCTTCGTTAGCGAGCTGCTGAGCAGTGGCTATCACCCTCTAGCCGCTGCA
 CTGCGCCGGCTGCTGGTAGCGAGCTGCTGAGCAGTGGCTATCAGCAGCAGTTCTCGCAGCAGATCGACGCC
 CGCCGTTTATAGACATTCAACCTCTTCTCTGTTAGCGAGCTGCTGAGCAGTGGCTATCACCCTCTAGCCGCTGCA
 GCCGTCGAGCCCTGATCAACTGTGCTTCGTTAGCGAGCTGCTGAGCAGTGGCTATCACCCTCTAGCCGCTGCA
 CTGCGCCGGCTGCTGGTAGCGAGCTGCTGAGCAGTGGCTATCAGCAGCAGTTCTCGCAGCAGATCGACGCC
 GACCTGTTGGGCCGACCAAGGTACTGCGCGCTGGCGAGCGCGCTGCGCTTACCGCAGAATTGCCAGCTG
 CTTGTGCAAGCCGCTGTTAGCGAGCTGCTGAGCAGTGGCTATCAGCAGCAGTGGCTATCACCCTCTAGCCGCTGCA
 CGTCGTGAACTCTATCATCTTCTCTGAAACAACACGGTTACCGCTGAGCAGTGGCTATCAGCAGCAGAATTGCCAGCTG
 GGCTGAAACCTATCATCTTCTCTGAAACAACACGGTTACCGCTGAGCAGTGGCTATCAGCAGCAGAATTGCCAGCTG
 CGTTATAACGATATCGCTGCTGGAGACCGCAACTGCCAGGCGCTGAGTGTCTTGGCTGAGCTGAGTGGT
 TGGCGAGTCGTTGAAACGGTGCAGCTGACCGAGCTAATGAAAGTCATCGCTGCTTCGGCTGAGCTGAGTGGT
 GAAGTTGTTCTGCCCTGCAATGGATGCCCACCGCTGCTGCAAGCAGTGAAGTGGCTCTGAACCAGCGCAAACCTCC
 TCT (**SEQ ID NO: 17**)

MSTTTVGDYLLYRLNEIGIEHLFGVPGDYNLQFLDHVIDHPQLTWVGCTNELNAAYAADGYARCRPAALLTTFG
 VGELSAINGIAGSYAEYLPVIHVGPAPLSAQQQGDLIHHSLGEGDFSSLRMSQPVSVAQAAALTPDNACKEIDR
 VLAEVLIQRRPGYLLLSTDVAAAPALPQSTLSPAPDHRAVLAASFDAAEQMLAQAKSVSLLADFLADRGV
 RALAALWLQQVPLPHATLLMGKGVLSSEQQPGFVGTYAGAASIDSTRGAIEEAGVIIGVGVRFSDTITAGFSQQIDA
 RRFIDIQPFPSRIGDRQFDHLPMQAAVALHQCLRYQQQWSITAPSPALPPAAGSELSQNAFWQAMQNFRPG
 DLLVADQGTAAGFAAALRLPQNQNCQLLVQPLWGSIGYSLPATFGAQTADERRVILIIGDGSQALTIQELSSMMRD
 GLKPII FLLNNNGYVERAIHGAERQYNDIAAWNWTQLPQALSVCAPAQSWRVVTQVLTDMKVIASPRILSLV
 EVLPAMDVPPLLQAVSAALNQRNSS (**SEQ ID NO: 18**)

[0282] Adh (标识符:t159319; 登录号:A0A1E4TMA4)

ATGCAGACGGCGTTCTGTATAAGCAGGTCAAGAAAACTTAGTGCGCTCGGAGATCCGATAACCTAAAGCTGGG
 CGTGGCGAAGTCGTTCTGGAAATTAAAGCCGCTGGCATGTGCCATTCCGATCTGCACGTTCTGACGGTGAACATC
 CCCCTGCCGGGTCATAATTGTAATGGCCATGAAATCGTTGGTACTATTACAGAGATCGGCCAGGACGTGACCGGT
 TTCAAACAGGGCGATCTGACGACTCACGGCCGAATCGTGTGGTATTGACCCCTGTGCAGAGAAGGATT
 GATAACGACTGCACTACAGTGGCAAACCGGTCAATGGTCTGGACTGGGCTTACGGGCTACCGAGAAGT
 ATCCGTATCCGAACGTAAGGTCTATCGTTAAAGTTCAGAGGTGTTACGCTGAGGCAGCTGCAGCTGTACT
 GATGCACTGACCCCCGTACCGTGCACAAAAACAGGCTGGGCCAGCAACTCTACTCGGGTACTGATTCTGGG
 CTGGGTGGCTTAGGTCTGAATGCCCTAAACTGGCTAACAGCCTTACGTTACGTTACGCTACGCTGACCTGAAA
 CCTCTGCGCGTGAAGCTGCTAACGGCGCTGGGGGGATGAAAGTGTGGAGTCCCCTGCCGAAGACCCGCTGGG
 GTTGATATCGTGTAGACGTCGGCGTGCAGAGCACCTCAACCTCGCTAAAAAACACGTTGGCCGAAATTGCG
 ATCATTGTACCTGTAGGGCTGGCATCCCCACAGCTTAACTAACGATCTGGCGCTCCGCAGAAGACCCGCTGGG
 GTTCAGGGCAGTTTGGGGCAGCAATGAGCTGGCTGAATGTCTGCCTGCCAGCTGGGCTGATCAAC
 CCGAAATATACTGTGGTGCCTTGAAGAAGGCCGAATATGAAAGCAATGGCTCATGGGAAAGTAGAAGGT
 CGTATCGTTTCCACCCG (SEQ ID NO: 19)

[0284] MQTAFLYKPGHENLVRSEIPIPKAGRGEVVLIEKAAGMCHSDLHVLDDGIPLPGQFVMGHEIVGTIHEIGQDV
 FKQGDLYAVHGPNPCGICTLCREGFNDCTTVAKTGQWFLGLDGGYQKYIRIPNVRSLVKVPEGVSAEAAASCT
 DAVLTPYRALQKAGASNSTRVLILGLGGLNLALKTFGSYVYASDLPSAREAAKAAGADEVLESPLPEDPLG
 VDIVLDVVGVQSTFNLAQKHVGPRGIIVPVGLASQLSFNLTDALREIRVQGTFWGTSNELAECLRLCQLGLIN
 PKYTVVPLEEAPKymeAMAHGKVEGRIVFHP (SEQ ID NO: 20)

[0285] Adh (标识符:t159028; 登录号:A0A192IDS9)

ATGCGCAGCATGCAGTTGATGAGTACGGTGCACCCCTGAAAGCGTTCTCATATGAAGACCCGACCCCGCAAGGG
 AAGGAAGTAGTCGTTAGGATCGAACGCTGTGGTGTGCCACTCTGATATTACATCTTCACGAGGGCTACTTCGAC
 ATGGCGGTGGCAATAAGCTGATTTACTCGTGCCTCGGAACCTCCCTTTACATTGGGTCTGAAATCGTGGC
 GAAGTGGTAGCAACTGGACCAGGTGTCACCGGGCTAACCGGGCGACAAACGTTGTTACCTGGGTGTTCACGTT
 TCGGGCAGCTGCCGAATGCAACAGTGGAGGATCAGTCTGTGCGCGTCCACGTAACCTGGGTATTCTACTGAG
 GACGGTGGCTATTGACCGCACGTAAGATACGGACGAAAATTCTGTTGCCTACGATGGTATTCTACTGAG
 TTACGGGAAACCTATGCTTGCACGGCATCACCGTTATGGTCACTGATGAAAGCAAGGAAGCGGCTGAAAGA
 TCTGGCTACATCGGTCTGATTGGCGTGGTGGCGTGGCATGGCTGATGCTGGCCAAGCAGCGATCGGG
 GCTAAAATGTAAGTCTTGATATCGACGACGAAAATGGAAGCTGCGACCCGTGCCGGGGCGATTACGTT
 AACCTCCGGTCAAAAGAACACGCAAGGAAGTTATGAAACTAACGAATGGTGGCTGCTGGTCTGTTGATTTC
 GTTGGCAGCGATAAAAGCGCTCTGTTGGAATCAACGCTGGGTCAAGAACGGCTGCTGGTCTGTTGATAATTGACTG
 TTCGGTGGCGCTATGACTGTTCCGGTACCCCTGTTCCCGTGAAGGGATCACCCTGATGCTGGCTCATACGTT
 TCCCTGCAAGAGATGAGTGATATGATGGAGTTAGTTCGCGTGGGAAAGTTCTCCGATGCCGTTAAAACCTCGG
 CCACTGGACGCTGCCCTGGAAACCCCTTGAGGATCTACGCCATGGTAAATCGTGGCGTGTGTTCTGACCCCA
 (SEQ ID NO: 21)

MRSMQFDEYGAPLKAFSYEDPTPQGKEVVRIEACGVCHSDIHLHEGYFDMDGGGNKADVTRARELPFTLGHEIVG
 EVVATGPGVTGAKPGDKRIVYPWIGCGDCPKCNSGEDQSCARPRNLGVHVDGGYSTHVKIPDEKFLFAYDGIPE
 LAGTYACSGITAYGALMAKEAAERSGYIGLIGAGGVMAGMLAKAAIGAKTVVFIDDDAKLEAATRAGADYVF
 NSGAKETRKEVMKLTNGGLSGAVDFVGSOKSALFGINALQNGVLVIIGLFGGAMTPVPLFLKGITVRGSYVG
 SLQEMSDDMELVRAGKVPPMPVKTRPLDAAWETLEDLRHGKIVGRVVLTP (SEQ ID NO: 22)

[0287] Adh (标识符:t158538; 登录号:A0AOP1J1W4)

ATGACAGCGGAGCAGAAAATGGGTATCCGACTCACGCCCTTCGAATTTCAGGAATTGGTGGCCCTATGCC
 CCACAGACCTATCAGCTCCCCGACCGGCTAGCGATGAAGTTGTTAAAGGTGAACACTGCGGTGCTGTCAC
 AGTGATGTTCATCTTCACGACGGCTACTCGAGCTGGGTGGCGATAACGCTGATCGCTATGCCGCTGCC
 CTGACGCTGGTACGAACTGGCACCGTGTGGCTGCGCGACAGGTTACTGGTGTAAAACCGGGGGAC
 CAGCGACTGATCTGGCGTGGAGGATGGCGGAAATGGCGCGTGTGCAAAAGGAGAAGAAAACCTGTGCGTT
 ACTCTGCACATCTGGCGTGAACAGCGGGCGGTTACCGCTGATCACGCTGTTACCCATTCTCGCTACCTT
 CTGGACATTCGGGTCTGACCGGGTGTGCGCTACCCCTCGCGTGTGCTGCCGCTGACCACTTCAGCGGATC
 AACAAAGTGTGCGCTTGACGATGACCGAGTGGATTGTTGTTATCGGTTGTTGGTGGCCCTGGCCAGATGGCGTG
 CGTATCTGCAAGTATGGAAATGGAATGTTACGCTGTTGACGCTGAAAGGAAACTGGCTCAT
 GAAAGCGGTGACGTCACTCTCGATCCAAACACTCCGAAGCTGACCCGCTGGTCGCCGAAACCTGCCGGGT
 ACGGTACAGGCCGCTTAGACTTGTGGGCAATGAGCAAACGCTCAGCTGGACTGCTGCTGTTGGAAAAGGT
 GGCAAAATATGTTCTGTCGGGCTGACGGCGGAGCTGCTTACCCATTGCGATCATCGAACAAGCTGTA
 AGTATCATCGTTTCTACGTTGGTACCCCTGAAAGAACTGGAGACTTAGTGTCTTCGCAAGGAAAAAAATCTG
 CCGCCAATTGATTAATGAAACACGCCCGCTGGAATCGGCGCTCAGGCCGTAGAGGACCTGGAAAAGACAGGTT
 GCTGGCGTGTATCTGGATGCAAGTAC (SEQ ID NO: 23)

MTAEQQNGVSDSRRFEFQEFGPIAPQTYQLPAPASDEVLLKVNYCGVCHSDVHLHDGYFELGGDKRLNFAMPLP
 LTGHEVIGTVVAVGDQVTGVKPGDQRILYWPWIGCGKCGACQKGEENLCVTPAHLGVNPKPGYADHIVVPHSRYL
 LDISGLNPDAATLACSGLITFSAINKVLPLADDQWIVVIGCGGLQMLRILQAMGIGNVIGIDLSEEKRKLAH
 ESCARHSFDPTPKLNRRVAETCPGTVQAALDFVGNEQTAQLALSLLGKGGYVPGVLHGTELRYPLPIITNKAV
 SIISGYVGLTKELEDLVAFAKEKNLPPHIERHPLAESAAQAVEDLEKGQVAGRVIDAGN (SEQ ID NO: 24)

[0289] GFP (阴性对照)

ATGACCGCCTTACGGAAGGGCAAAACTGTTGAGAAAGAGATAACCGTATATAACCGAATCTGAAGGGCGACGTAA
GAAGGGATGAAATTATAATTAAAGGCAGGGGACCGGGACGCCACGGGGACCATTAAAGCGAAATACATA
TGCACCTACGGCGACCTGCCGTACCGTGGCAACCTGGTACGGCTACGGGTCCAGTGTTGCC
AAGTACCCGAGCCACATAAAGGATTCTTTAAGAGCGCCATGCCGAAGGGTATAACCCAGAGCGTACCATAGC
TTCGAAGGCAGGGCGTACAAGACGCGTGTATGGTCACCTACGAACCGGGCTATATACATCGTGTACG
CTGACTGGGAGAACCTTAAGAAGACGGCACATTCTGCGTAAGAACGTCGCATTCCAATGCCGCAAAGCATT
CTGTATATTCTGCCTGACACCGTCAACAATGGCATACGCGTCAACCAGGGTACGATATTGAAGGGGTG
ACCGAAAAACTGGTCACCAATGCAGCCAAATGAATCGTCGTTGCCGGCAGTGGGCAGTGCATATACCGCGT
TATCATCACATTACCTACCAACACAAACTGAGCAAAGACCGCGACGAGGCCGTGATCACATGTGTCTGGTGAG
GTAGTGAAAGCGGTCGATCTGGACACGTATCAGTGA (**SEQ ID NO: 25**)

MTALTEGAKLFEKEIPYITELEGDVEGMKFIIKGEGTGDATTGTIKAKYICTTGDLPVWATLVSTLSYGVQCFA
KYPISHIKDFFKSAMPEGYTQERTISFEGDGVYKTRAMVTYERGSIYNRVTLTGENFKKDGHILRKNAFQCPSI
LYILPDVTNNNGIRVEFNQAYDIEGVTEKLVTKCSQMNRLAGSAAVHIPRYHHITYHTKLSKDRERRDHMCLVE
VVKAVDLDTYQ (**SEQ ID NO: 26**)

[0291] 酶筛选数据

[0292] 表4. LeuDH酶和相对于对照的活性

登录号	蛋白质突变	菌株	相对于对照的倍数提高	核苷酸 SEQ ID NO	蛋白质 SEQ ID NO
P0A392	野生型	对照	0	37	257
A0A1T4PGG9	野生型	t160946	2.846	38	258
A4CBM3	野生型	t161014	2.188	39	259
A0A0C1US13	野生型	t160854	2.178	40	260
A0A1M6BE59	野生型	t160389	2.166	41	261
K2M7H0	野生型	t160943	2.027	42	262
A0A1Q6ZIF7	野生型	t160092	2.005	43	263
A0A075JPW8	野生型	t160267	2.002	44	264
A0A0B5AS65	野生型	t160288	1.910	45	265
A0A0V8JFL2	野生型	t160337	1.826	46	266
A0A1S2LUY1	野生型	t160524	1.804	47	267
A0A0A8UN70	野生型	t161111	1.792	48	268
P0A392	G43T	t159984	1.775	49	269
A0A1E7PTP0	野生型	t161162	1.751	50	270
A0A1S9B636	野生型	t160283	1.741	51	271
P0A392	E116V	t160562	1.553	52	272
A0A1D2RXB2	野生型	t160434	1.550	53	273
K4KRS4	野生型	t160706	1.548	54	274
P0A392	L76F	t160502	1.538	55	275
P0A392	T136R	t160559	1.521	56	276
P0A392	A297C	t160202	1.509	57	277
A0A1I1NGX1	野生型	t160947	1.501	58	278
A0A142ITE6	野生型	t161198	1.401	59	279

[0294]

I1DTY5	野生型	t160169	1.364	60	280
P0A392	A297Y	t160199	1.364	61	281
A0A0A0EMP0	野生型	t160499	1.359	62	282
W4PY11	野生型	t160682	1.359	63	283
R8B531	野生型	t161210	1.359	64	284
A0A1Q2KY34	野生型	t160573	1.340	65	285
L1QQC1	野生型	t161091	1.333	66	286
D6XVM2	野生型	t160162	1.301	67	287
P0A392	L78V	t160587	1.281	68	288
A0A1G8KLY7	野生型	t160351	1.267	69	289
A0A0J6CNT2	野生型	t160438	1.254	70	290
P0A392	L300K	t160181	1.196	71	291
U3HCY1	野生型	t161117	1.191	72	292
A0A1K1TVW4	野生型	t160461	1.188	73	293
A0A1Y6CWJ6	野生型	t160154	1.186	74	294
A0A154W9T2	野生型	t160973	1.171	75	295
I1D544	野生型	t161185	1.149	76	296
A0A165NUD8	野生型	t161204	1.149	77	297
A0A0A8JN83	野生型	t160338	1.144	78	298
P0A392	N71T	t160401	1.144	79	299
F7RX04	野生型	t160786	1.110	80	300
A0A1U9K9A9	野生型	t160671	1.108	81	301
A0A0K6GVS2	野生型	t160957	1.105	82	302
A0A136MKS4	野生型	t160417	1.095	83	303
A0A0A5GIG6	野生型	t160609	1.076	84	304
A0A143BJV1	野生型	t160627	1.051	85	305
K6YKY7	野生型	t161088	1.046	86	306
A0A0T5PG63	野生型	t160158	1.032	87	307
A0A1M6L5E8	野生型	t160479	1.032	88	308
P0A392	L42Q	t160013	1.029	89	309
A0A0A2TA47	野生型	t160286	1.017	90	310
P0A392	A297H	t160636	1.012	91	311
A0A0Q5UT14	野生型	t160279	1.002	92	312
I4D8U4	野生型	t160598	1.000	93	313
P0A392	I113V	t160129	0.993	94	314
A0A1G3WLY4	野生型	t159999	0.976	95	315
P0A392	A297N	t160134	0.968	96	316
P0A392	A297M	t160503	0.954	97	317

[0295]

A0A1X4MV49	野生型	t160926	0.949	98	318
P0A392	A297L	t160497	0.912	99	319
A0A0J1FEE3	野生型	t160141	0.897	100	320
P0A392	E116A	t160512	0.892	101	321
P0A392	M67T	t160125	0.883	102	322
A0A0F7HKR2	野生型	t160291	0.873	103	323
K0AAV5	野生型	t160552	0.870	104	324
A0A1Q4XJW1	野生型	t160891	0.868	105	325
P0A392	L300N	t160557	0.866	106	326
A0A0K9GVT6	野生型	t160443	0.863	107	327
W7D8C3	野生型	t160771	0.858	108	328
F7NG13	野生型	t160215	0.851	109	329
A0A1H8Q403	野生型	t160870	0.836	110	330
P0A392	L42T	t160357	0.829	111	331
E1WZZ8	野生型	t160664	0.797	112	332
A0A0K9GC14	野生型	t160444	0.790	113	333
P0A392	V296N	t160184	0.787	114	334
A0A1F3SFY8	野生型	t160002	0.785	115	335
P0A392	L78K	t160487	0.782	116	336
P0A392	T136S	t160176	0.768	117	337
A0A1Y5EK08	野生型	t160841	0.768	118	338
P0A392	T136F	t160489	0.763	119	339
N0AUJ4	野生型	t160823	0.751	120	340
P0A392	M67Q	t159980	0.748	121	341
C4L3E4	野生型	t160256	0.748	122	342
A0A1I6TTT1	野生型	t160115	0.733	123	343
P0A392	A297R	t160509	0.733	124	344
A0A1H7JK8	野生型	t160952	0.733	125	345
A0A1U7M8J0	野生型	t160255	0.724	126	346
P0A392	L300Q	t160226	0.721	127	347
A1S7B6	野生型	t160188	0.719	128	348
P0A392	V293S	t160602	0.711	129	349
C1A7X5	野生型	t160733	0.709	130	350
A0A0W0TJD2	野生型	t161212	0.697	131	351
P0A392	I113F	t160504	0.689	132	352
P0A392	M67E	t160064	0.685	133	353
A0A1U7JH14	野生型	t160966	0.685	134	354
P0A392	L300A	t160612	0.680	135	355

[0296]

P0A392	E116S	t160543	0.675	136	356
P0A392	G43F	t160059	0.672	137	357
P0A392	A297F	t160588	0.670	138	358
M8DS05	野生型	t160310	0.663	139	359
P0A392	L300C	t160633	0.658	140	360
P0A392	L300F	t160128	0.655	141	361
M7N8L2	野生型	t160152	0.655	142	362
P0A392	L78F	t160584	0.653	143	363
G8R2S3	野生型	t160212	0.650	144	364
A0A0P8B102	野生型	t161073	0.650	145	365
S2YPJ0	野生型	t160830	0.643	146	366
A0A1M5CX03	野生型	t159964	0.636	147	367
P0A392	L76E	t160245	0.626	148	368
A0A1M5IEB6	野生型	t160988	0.626	149	369
A0A0F6SHW7	野生型	t160860	0.619	150	370
A0A0U3AUS4	野生型	t160964	0.619	151	371
A0A081G3H3	野生型	t160968	0.604	152	372
A0A1Q4UNH5	野生型	t161006	0.599	153	373
P0A392	A297D	t160548	0.597	154	374
P0A392	V293Q	t160249	0.594	155	375
P0A392	T136E	t160648	0.594	156	376
P0A392	L300D	t160248	0.587	157	377
P0A392	L300T	t160270	0.587	158	378
P0A392	L76H	t160546	0.587	159	379
P0A392	L76W	t160139	0.579	160	380
P0A392	L76M	t160274	0.575	161	381
P0A392	L300M	t160541	0.548	162	382
T0CG61	野生型	t160808	0.538	163	383
A0A166W971	野生型	t160538	0.535	164	384
P0A392	V296C	t160206	0.533	165	385
P0A392	A297E	t160567	0.533	166	386
K2JU58	野生型	t160877	0.523	167	387
P0A392	G44I	t160011	0.516	168	388
A0A0M4FMC6	野生型	t160371	0.516	169	389
P0A392	M67S	t160060	0.509	170	390
A0A0K1JA83	野生型	t160995	0.509	171	391
P0A392	A115T	t159988	0.504	172	392
A0A1N6U8W9	野生型	t160814	0.504	173	393

[0297]

A0A075LQK1	野生型	t160493	0.499	174	394
P0A392	G44Y	t160080	0.494	175	395
P0A392	L300H	t160197	0.494	176	396
A0A0K8QRE8	野生型	t160626	0.489	177	397
A0A1M6M3I5	野生型	t160912	0.487	178	398
A0A0F7JZ22	野生型	t161016	0.477	179	399
P0A392	L78H	t160634	0.469	180	400
A0A1Y6BX33	野生型	t160700	0.460	181	401
P0A392	V296L	t160146	0.447	182	402
A0A1L8CTI5	野生型	t161020	0.445	183	403
P0A392	L300Y	t160145	0.443	184	404
P0A392	E116N	t160539	0.428	185	405
A0A171DN74	野生型	t160716	0.423	186	406
P0A392	A297K	t160491	0.416	187	407
P0A392	L78Y	t160594	0.416	188	408
E6TXR8	野生型	t160618	0.416	189	409
P0A392	N71H	t160120	0.411	190	410
A0A1G3X1T7	野生型	t160910	0.411	191	411
P0A392	E116W	t160246	0.408	192	412
U4KND6	野生型	t160852	0.408	193	413
P0A392	E116R	t160131	0.399	194	414
P0A392	N71C	t160385	0.399	195	415
A0A1G0BBA9	野生型	t160899	0.396	196	416
A0A1Y2L717	野生型	t160990	0.396	197	417
P0A392	A297T	t160227	0.389	198	418
A0A0M4UKZ2	野生型	t160340	0.379	199	419
P0A392	A297W	t160596	0.357	200	420
P0A392	L78C	t160406	0.350	201	421
E2SC01	野生型	t161059	0.350	202	422
A0A1K1PP57	野生型	t160629	0.347	203	423
P0A392	G44K	t159990	0.345	204	424
P0A392	A115S	t160495	0.342	205	425
P0A392	L300S	t160275	0.337	206	426
P0A392	L300W	t160639	0.337	207	427
A0A1G0A9I7	野生型	t160875	0.337	208	428
A0A0W7WYJ8	野生型	t161047	0.337	209	429
P0A392	V296E	t160520	0.325	210	430
P0A392	T136Y	t160638	0.325	211	431

[0298]

P0A392	A115V	t160123	0.320	212	432
A0A1V0ADI4	野生型	t160970	0.318	213	433
W7ZGF1	野生型	t160812	0.315	214	434
P0A392	A115Q	t159982	0.311	215	435
A0A1H6CJX7	野生型	t161141	0.308	216	436
P0A392	M67K	t160356	0.296	217	437
P0A392	L78Q	t160581	0.296	218	438
P0A392	T136L	t160589	0.293	219	439
P0A392	E116L	t160604	0.293	220	440
P0A392	I113M	t160628	0.291	221	441
P0A392	L76Y	t160516	0.289	222	442
P0A392	V293A	t160655	0.274	223	443
P0A392	V296K	t160243	0.267	224	444
P0A392	L76R	t160153	0.264	225	445
P54531	野生型	t160721	0.262	226	446
P0A392	V296I	t160271	0.259	227	447
P0A392	L300R	t160560	0.254	228	448
K9ARW8	野生型	t160789	0.252	229	449
P0A392	L76S	t160133	0.249	230	450
P0A392	I113W	t160094	0.244	231	451
P0A392	A115N	t160194	0.240	232	452
P0A392	V296S	t160644	0.240	233	453
P0A392	E116M	t160643	0.235	234	454
P0A392	L42A	t160402	0.232	235	455
P0A392	V293C	t160500	0.225	236	456
P0A392	N71M	t160324	0.220	237	457
P0A392	V296A	t160143	0.213	238	458
P0A392	G43W	t160099	0.210	239	459
P0A392	A297Q	t160140	0.196	240	460
P0A392	V293T	t160221	0.191	241	461
P0A392	I113Y	t160098	0.188	242	462
P0A392	L76I	t160601	0.188	243	463
P0A392	G44H	t160029	0.176	244	464
P0A392	L76K	t160585	0.171	245	465
P0A392	G43Y	t159996	0.169	246	466
P0A392	N71D	t160415	0.142	247	467
P0A392	I113Q	t160632	0.139	248	468
P0A392	M67A	t160055	0.127	249	469

[0299]

P0A392	V296T	t160630	0.122	250	470
P0A392	L76T	t160603	0.115	251	471
A0A1Q4VRJ4	野生型	t161033	0.112	252	472
B2A513	野生型	t160167	0.108	253	473
P0A392	G43E	t160096	0.083	254	474
P0A392	N71K	t160101	0.044	255	475

[0300]

表5.KivD酶和相对于对照的活性

[0301]

登录号	标记	与对照相比的倍数提高	核苷酸 SEQ ID NO:	蛋白质 SEQ ID NO
Q684J7	对照	0	477	533
A0A085UD38	t163850	1.958	478	534
A0A090DYV6	t163542	3.986	479	535
A0A0A6W4H3	t163732	4.354	480	536
A0A0B1U4F6	t163805	2.972	481	537
A0A0D0SDJ9	t163730	3.292	482	538
A0A0D2CSK3	t163274	3.965	483	539
A0A0D2GWW0	t163016	4.354	484	540
A0A0H4KFT8	t163716	3.958	485	541
A0A0J8UR79	t163869	2.250	486	542
A0A0K2Y209	t163916	3.944	487	543
A0A0L0P8D8	t163988	5.097	488	544
A0A0L7TB96	t163842	4.833	489	545
A0A0M5JJZ2	t164076	4.944	490	546
A0A0M5MY84	t163914	4.139	491	547
A0A0Q4N500	t164007	4.493	492	548
A0A0T9T7Y7	t163705	3.694	493	549
A0A0T9UPI9	t163338	3.493	494	550
A0A0U1CW59	t163964	3.201	495	551
A0A0U2NS09	t163656	2.222	496	552
A0A198FEB4	t163871	4.382	497	553
A0A1B1NY37	t163888	3.646	498	554
A0A1B7ILY5	t163742	3.792	499	555
A0A1B9AUW4	t162995	3.889	500	556
A0A1D4X3F2	t163818	4.708	501	557
A0A1F2KK66	t163546	3.403	502	558
A0A1G7WAJ7	t163085	5.076	503	559
A0A1M7EHD4	t163474	1.813	504	560

[0302]

A0A1Q4T3V5	t163704	4.535	505	561
A0A1T1GFV6	t163784	3.500	506	562
A0A1U4TJK1	t163702	4.722	507	563
A0A1V2L8B3	t164100	3.229	508	564
A0A1V2YXQ3	t163766	2.319	509	565
A0A1V4SV36	t163852	4.396	510	566
A0A1V6TQU7	t162902	3.118	511	567
A0A1W6B724	t163806	3.639	512	568
A0A1X0AE10	t163798	3.104	513	569
A0A1X1XPA7	t163472	3.826	514	570
A0A1X2FKJ1	t163432	3.035	515	571
A0A1Y6E4E9	t163406	3.486	516	572
A0A205J7X5	t163837	3.910	517	573
A0A2B1L7A1	t163722	4.215	518	574
B9DJU8	t163844	4.597	519	575
D4B725	t163868	4.111	520	576
D4C3A5	t163661	2.139	521	577
D4F0I3	t163478	0.896	522	578
D7UWC4	t163880	3.785	523	579
F5SQV4	t163740	2.535	524	580
G9YCD8	t163678	0.090	525	581
I1CGS4	t163934	3.785	526	582
J2LV57	t163902	2.667	527	583
Q6C9L5	t163155	4.222	528	584
R5SST3	t163337	3.014	529	585
R8AV71	t163285	4.535	530	586
S3IST7	t163983	2.979	531	587
W0L941	t163973	3.396	532	588

[0303] 表6. Adh酶和相对于对照的活性

[0304]

登录号	标记	相对于对照的倍数 提高	核苷酸序列 SEQ ID NO	蛋白质序列 SEQ ID NO
P00331	对照	0	589	645
A0A011RFM0	t159061	-0.581	590	646
A0A068NM64	t159163	4.815	591	647
A0A081B9F7	t159282	5.992	592	648
A0A0F7S860	t159174	-0.411	593	649
A0A0F8XA97	t159080	-0.250	594	650

[0305]

A0A0L8BIH2	t158526	0.629	595	651
A0A0M2SIC1	t158995	0.323	596	652
A0A0M8TKC3	t158267	2.427	597	653
A0A0N1F703	t159004	9.032	598	654
A0A0P1J1W4	t158538	10.516	599	655
A0A0Q6FH05	t159022	-0.113	600	656
A0A0Q9AMT3	t158946	1.476	601	657
A0A163KUH6	t159154	0.710	602	658
A0A192IDS9	t159028	10.581	603	659
A0A1A0K0C6	t159162	0.645	604	660
A0A1E4TMA4	t159319	11.113	605	661
A0A1E7X363	t159283	4.234	606	662
A0A1Q7HM90	t159036	-0.492	607	663
A0A1V1TTZ9	t158998	-0.613	608	664
A0A1V2EYM1	t159040	3.750	609	665
A0A1V6E459	t159120	1.008	610	666
A0A1Y0G594	t159236	1.645	611	667
A2V8B3	t159176	4.758	612	668
A9MKQ8	t158774	-0.548	613	669
C0SPA5	t158820	1.113	614	670
D8MZP3	t159280	6.234	615	671
F0IX07	t159318	0.371	616	672
H1ZV38	t158442	1.694	617	673
J1KN15	t158976	4.008	618	674
J5T2P7	t159183	-0.161	619	675
K4IPR3	t158247	5.444	620	676
M1LUC5	t158246	0.073	621	677
M2N9N4	t159152	0.669	622	678
M2QHN1	t159090	0.282	623	679
M2YNQ9	t159054	3.629	624	680
M5FVU5	t158291	0.565	625	681
O74822	t158955	-0.500	626	682
P08843	t158458	0.460	627	683
P0DMQ6	t158893	-0.444	628	684
P13603	t158263	0.645	629	685
P14219	t158869	-0.048	630	686
P14673	t158726	0.952	631	687
P14675	t158728	6.056	632	688

[0306]	P20368	t158816	0.798	633	689
	P25141	t158333	3.677	634	690
	P28032	t158454	2.887	635	691
	P39451	t158390	5.500	636	692
	P39849	t158243	0.516	637	693
	P40394	t158613	3.460	638	694
	P42328	t158520	2.065	639	695
	Q2FJ31	t158326	1.024	640	696
	Q38707	t158580	-0.105	641	697
	Q99W07	t158330	1.597	642	698
	S0EJ18	t159328	11.185	643	699
	W5YKG3	t159122	0.782	644	700

[0307] 表7. 相对于SEQ ID NO:27具有增加的LeuDH活性的酶中的保守氨基酸

SEQ ID NO: 27 中的对应位置	氨基酸
13	V
16	W
42	Q
43	T、Y、F、E、W
44	I、H、K、Y
67	T、E、A、S、K
71	K
73	S
76	R、H、Y、S、K、W
92	Y
93	H
95	G
100	G
105	C
111	G
113	M
115	N、V
116	R、N、W
120	A
122	D
136	E
140	D
141	M

[0309]

160	S
185	F
196	N
228	Y
248	M
256	C
293	Q、C
296	K、N
297	R、Q、K
300	C、D
302	T、S
305	C
319	F
330	M

[0310]

表8. 相对于SEQ ID NO:29具有增加的KivD活性的酶中的保守氨基酸

[0311]

SEQ ID NO:29中的位置中的对应位置	氨基酸
33	Y
44	Q
117	M
129	I
185	W
190	I
225	I
227	Y
311	L
312	G
313	T
328	P
341	W
345	H
347	C
420	R
494	D
508	C
550	F

[0312]

表9. 相对于SEQ ID NO:31具有增加的ADH活性的酶中的保守氨基酸

[0313]

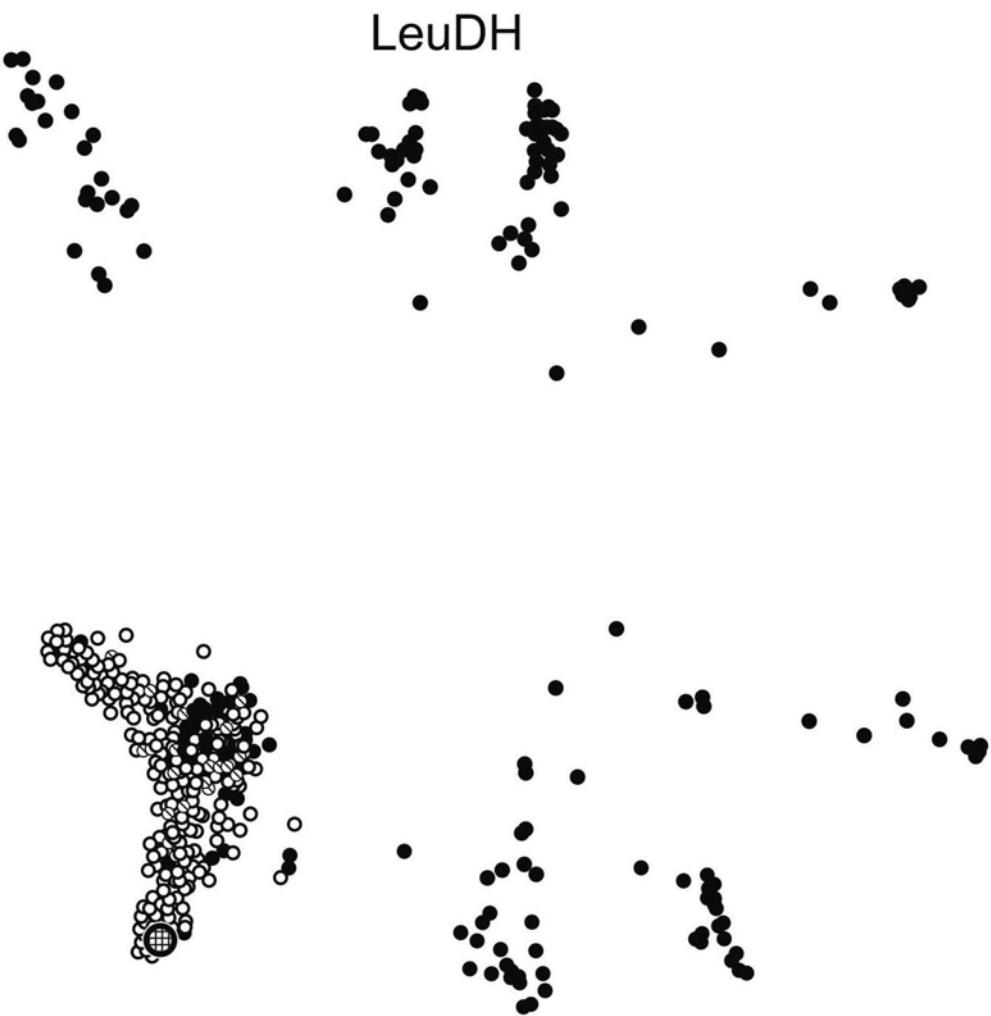
SEQ ID NO:31中的对应位置	氨基酸

9	P
16	G
23	Q
28	R
30	A
93	K
98	L
99	R
114	P
115	K
119	Y
194	Y
242	P
249	K
255	E
260	D
269	H
281	Q
325	L
333	M
334	P
348	Q

[0314] 等同物

[0315] 本领域技术人员仅使用常规实验就将认识到或能够确知本公开中描述的本发明的具体实施方案的许多等同物。这样的等同物旨在由以下权利要求书涵盖。

[0316] 本申请中公开的全部参考文献(包含专利文件)通过引用被整体(特别是本公开中引用的公开内容)并入本文。



簇中最常见的注释

- 亮氨酸脱氢酶细菌
- ▨ Glu/Leu/Phe/Val脱氢酶二聚区thermarum TA2.A1
- 缬氨酸脱氢酶
- ▨ 种子序列

图1A

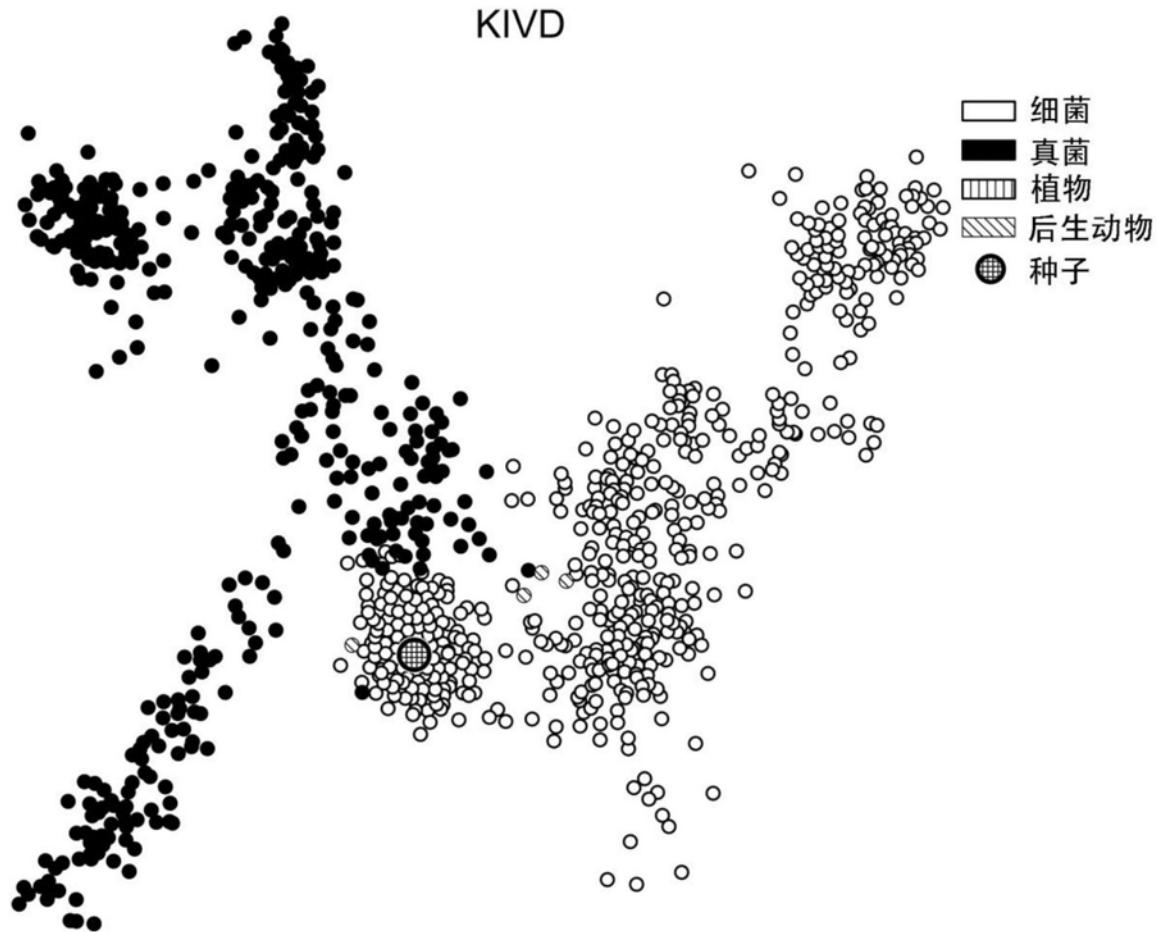


图1B

ADH

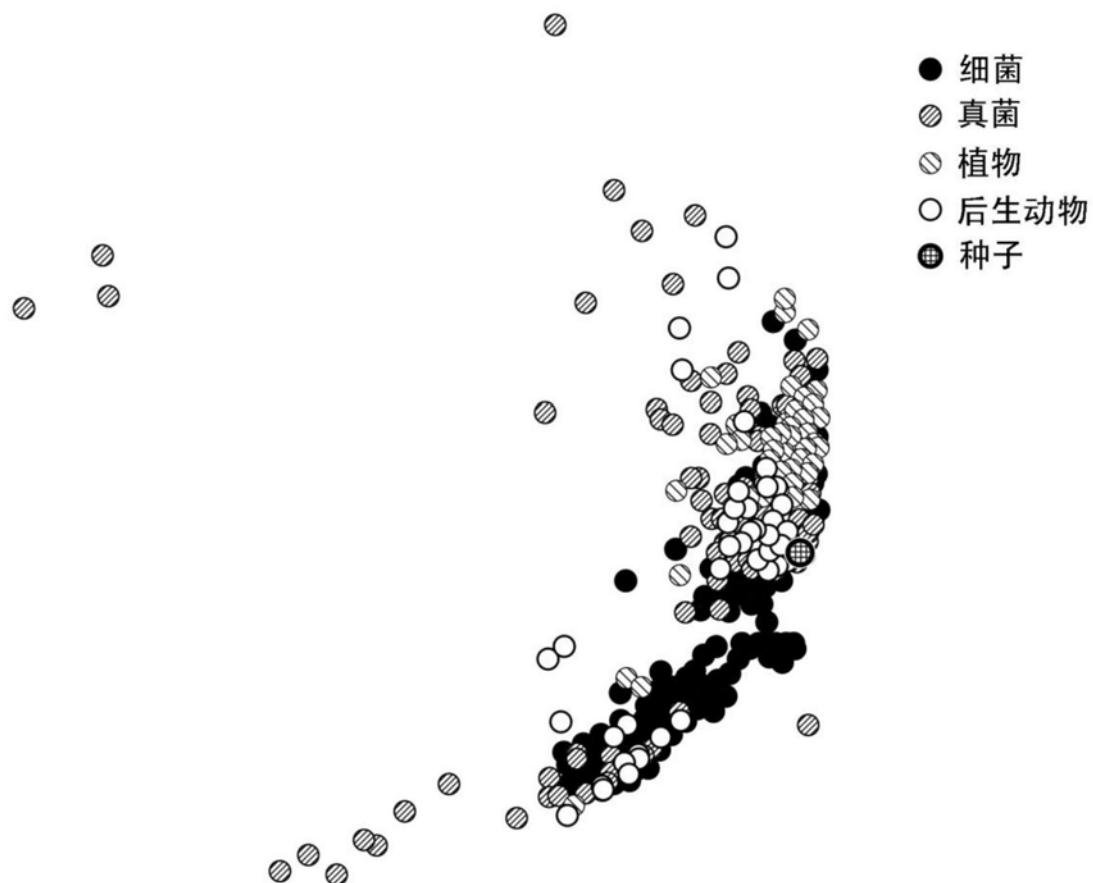


图1C

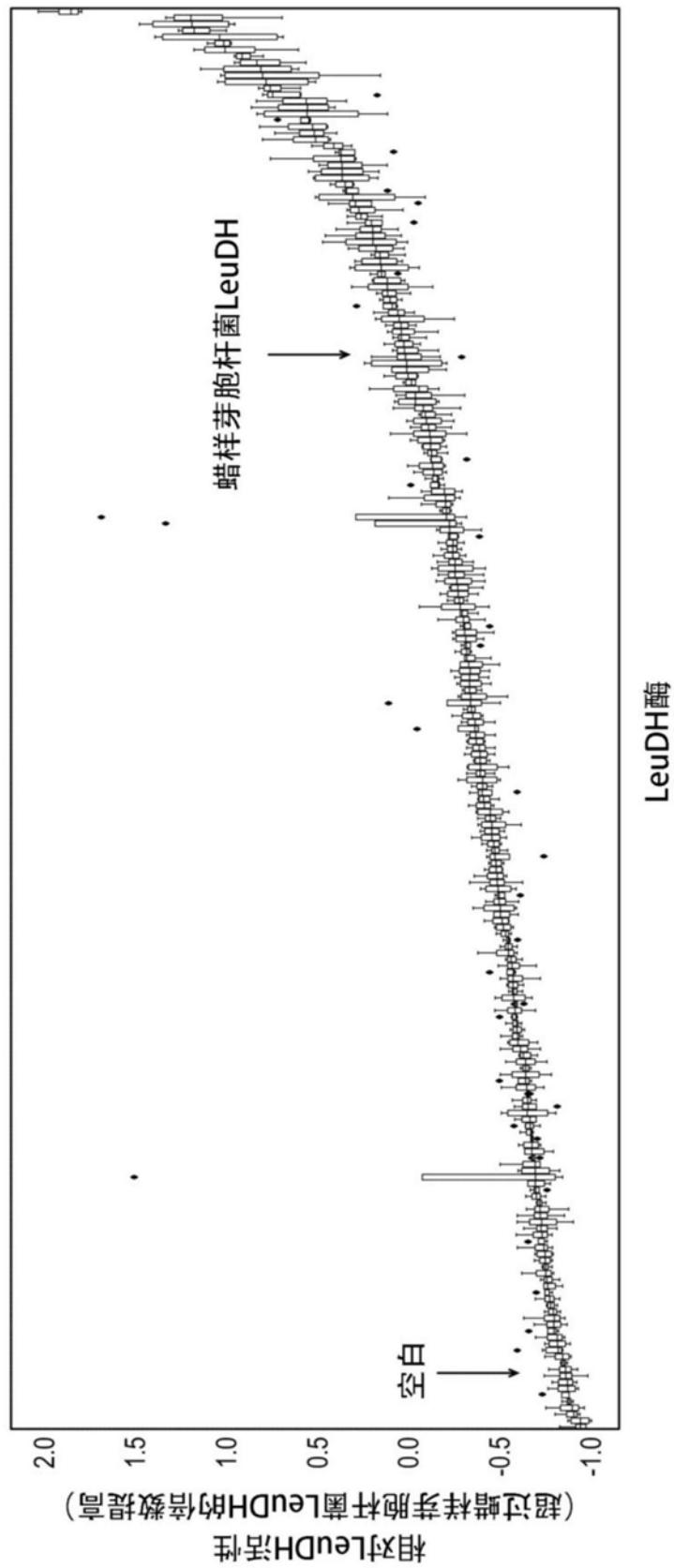


图2

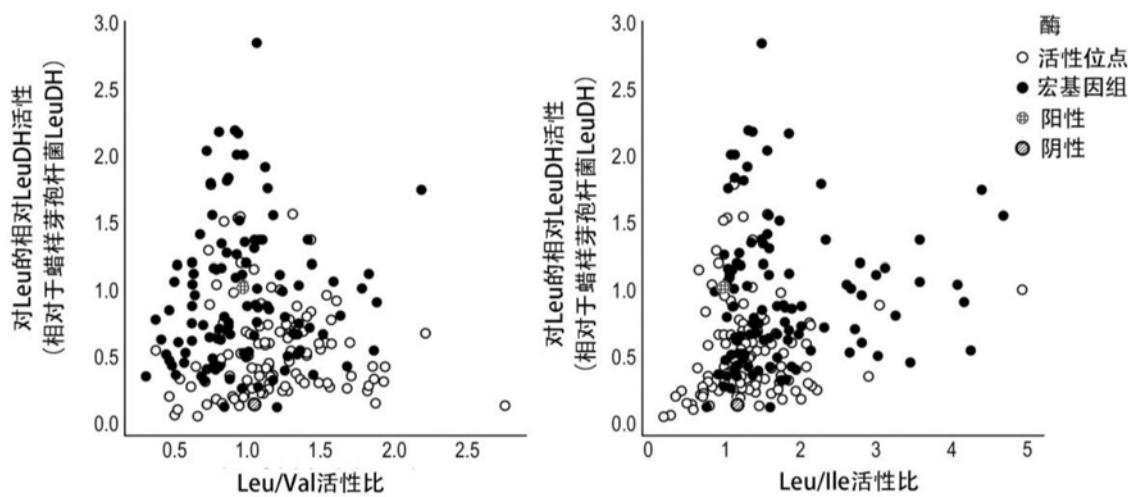


图3

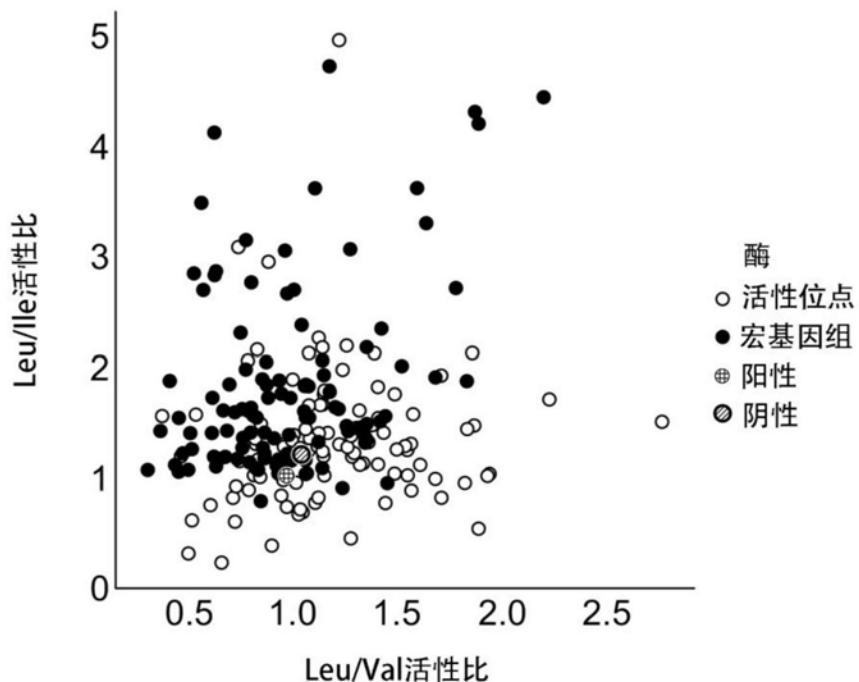


图4

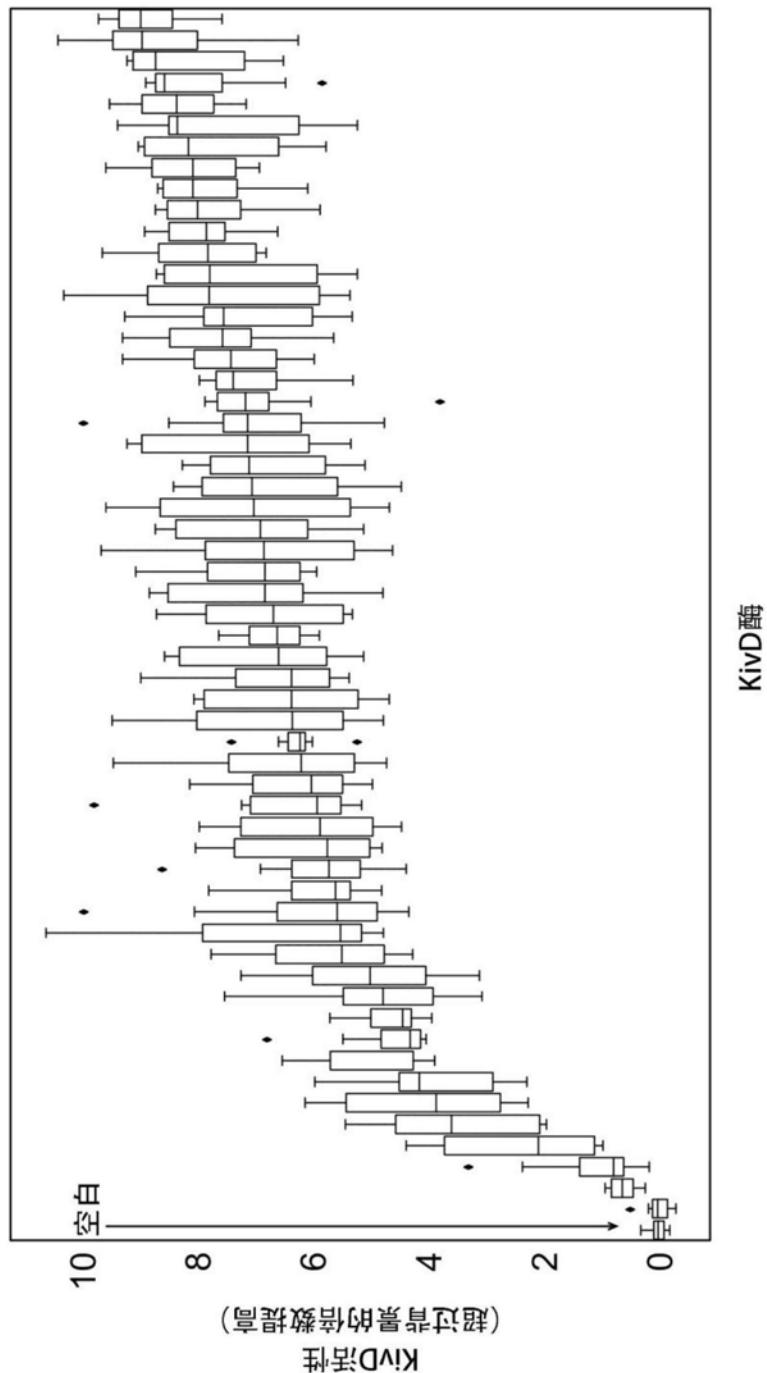


图5

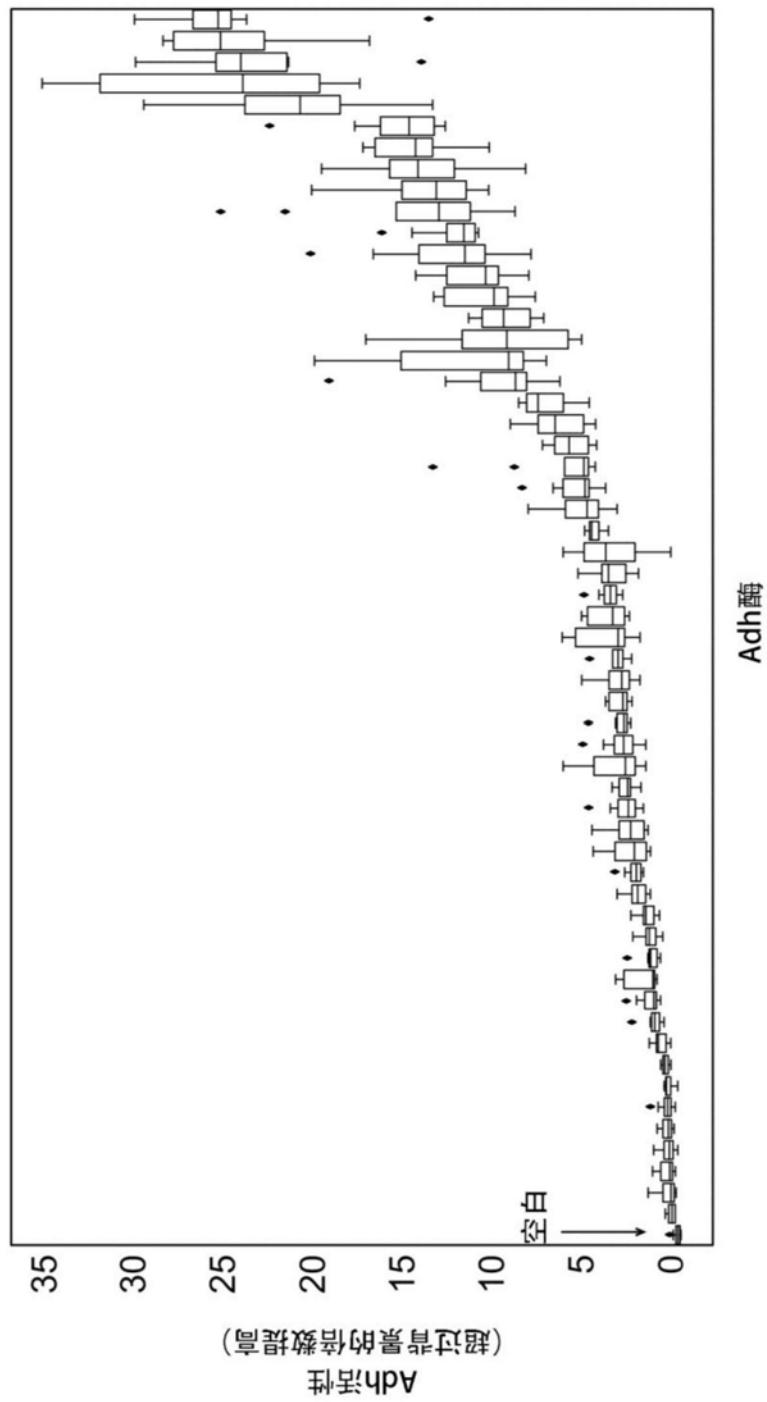
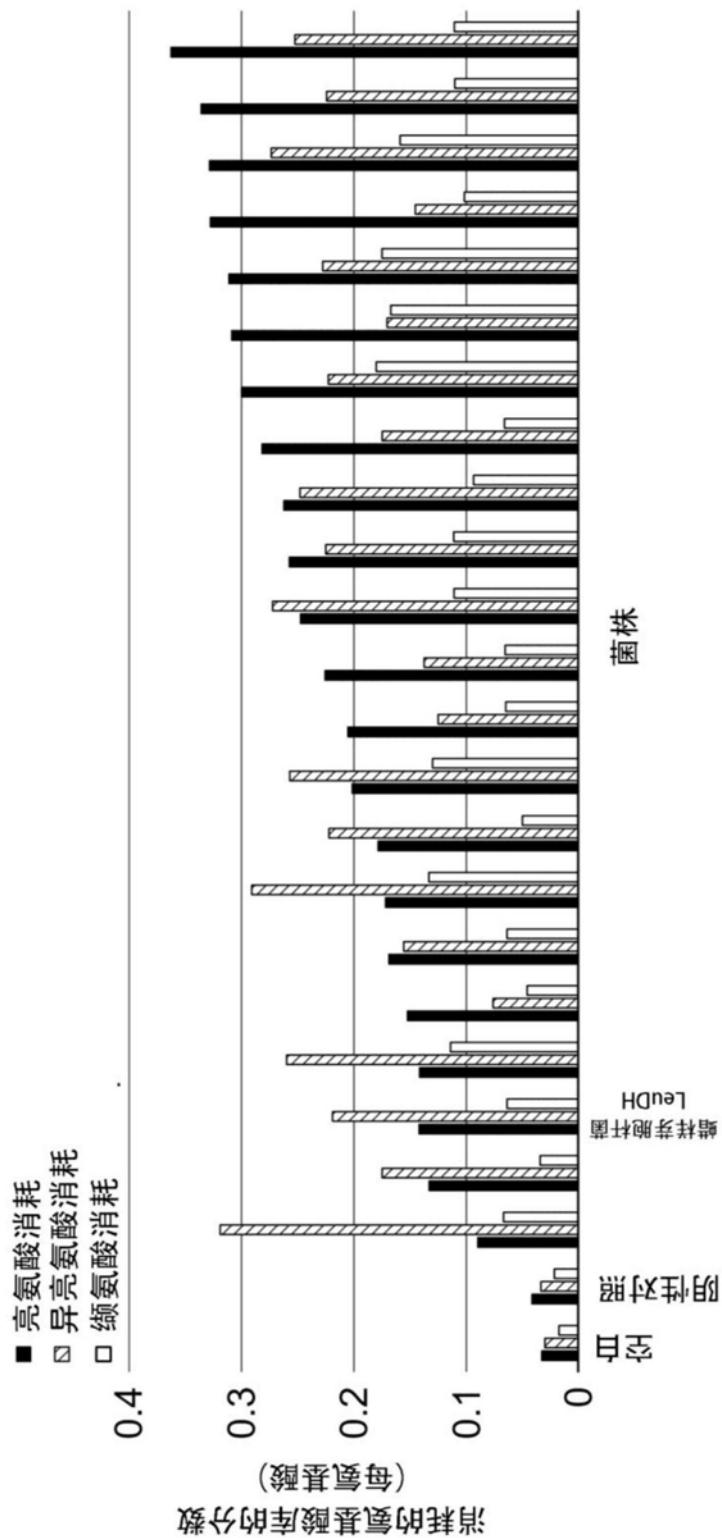


图6



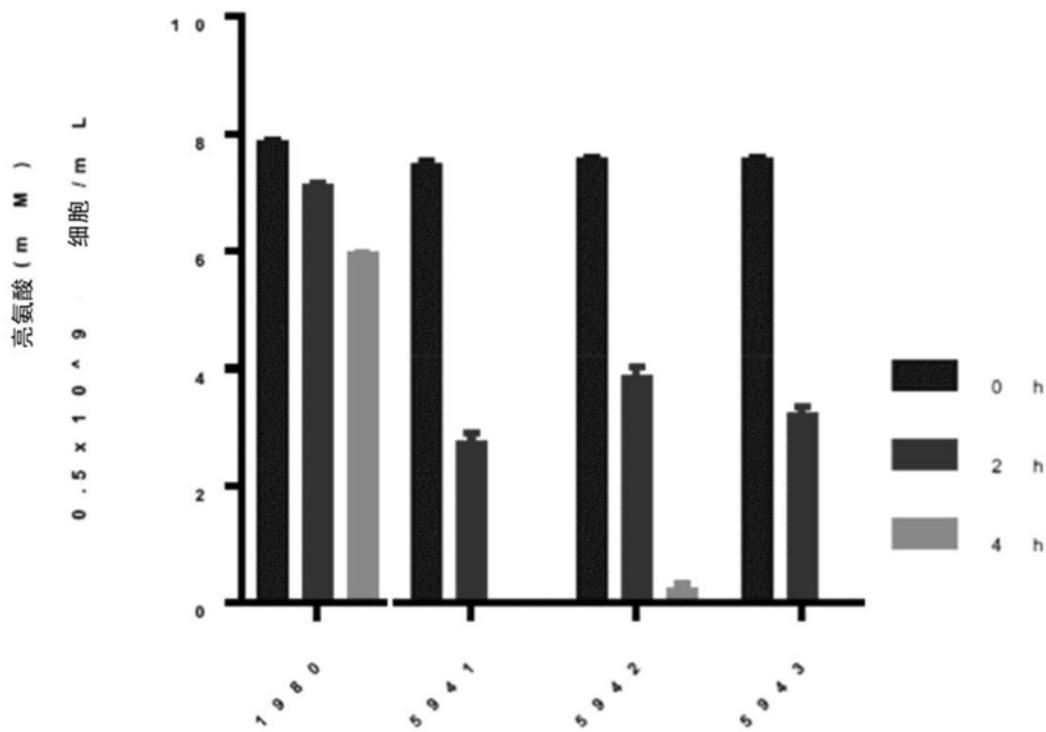


图8

MSUD途径

图9

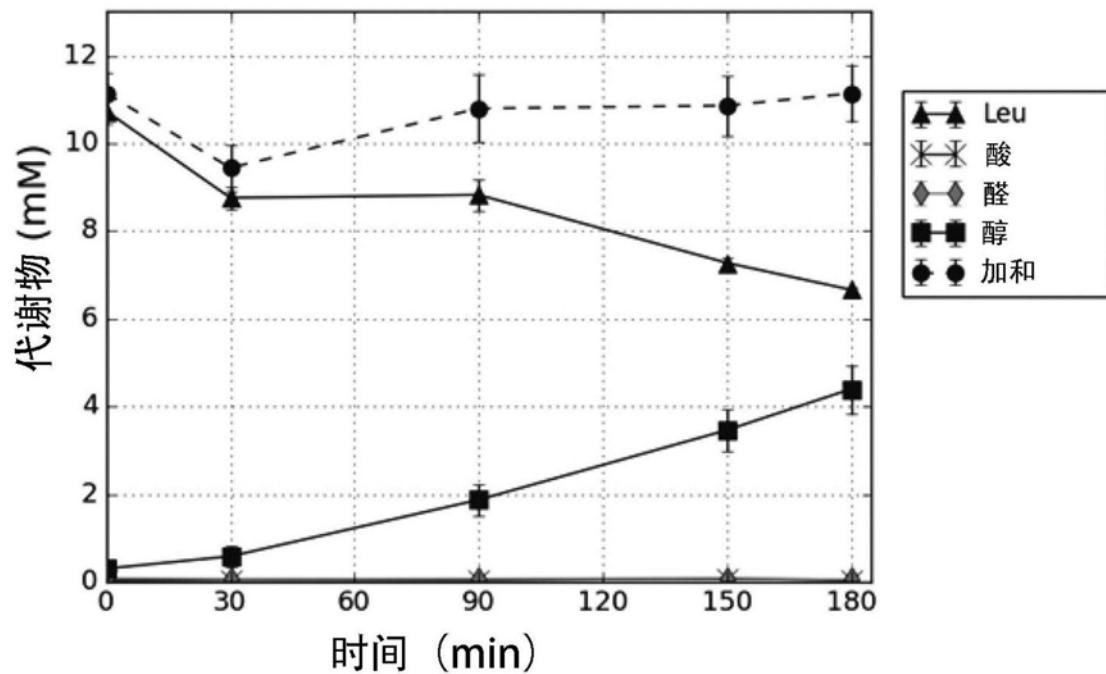


图10