(19)中华人民共和国国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 107058258 A
(43)申请公布日 2017.08.18

(21)申请号 201611027264.1

(22)申请日 2009.04.09

(30)优先权数据
0806562.5 2008.04.10 GB

(62)分案原申请数据
200980121472.X 2009.04.09

(71)申请人 赛默飞世尔科技波罗的海UVB公司
地址 立陶宛维尔纽斯

(72)发明人 阿尔维达斯·亚努拉提斯
雷米吉犹斯·斯卡尔加拉
唐吉拉·斯克斯内内

(74)专利代理机构 北京三友知识产权代理有限公司 11127
代理人 丁香兰 庞东成

(51)Int.Cl.
*C12N 9/12*(2006.01)
*C12N 15/54*(2006.01)

权利要求书1页 说明书92页 附图16页

(54)发明名称

一种逆转录酶和编码其的多核苷酸

(57)摘要

本发明涉及一种逆转录酶，所述逆转录酶包含莫洛尼氏鼠白血病病毒逆转录酶氨基酸序列，并且包含在以下氨基酸位置中的一处或多处的突变。

| | | | | | |
|---|---|---|---|---|---|
| L603, | D200, | L139, | T330, | E607, | T287, D524 |
| Q221, | I49, | N479, | H594, | A502, | |
| D653, | K658, | P130, | Q237, | | |
| A307, | Y344, | Q430, | D449, | A644, | N649, |
| L671, | E673, | M39, | Q91, | M66, | W388, |
| I179 | L333 | R390 | Q374 和 | E5 | |

CN 107058258 A

1.一种逆转录酶，所述逆转录酶包含莫洛尼氏鼠白血病病毒逆转录酶氨基酸序列，并且包含在以下氨基酸位置中的一处或多处的突变：

| | | | | | |
|---|---|---|---|---|---|
| L603, | D200, | L139, | T330, | E607, | T287, D524 |
| Q221, | I49, | N479, | H594, | A502, | |
| D653, | K658, | P130, | Q237, | | |
| A307, | Y344, | Q430, | D449, | A644, | N649, |
| L671, | E673, | M39, | Q91, | M66, | W388, |
| I179 | L333 | R390 | Q374 和 | E5 | |

其中当所述突变位于D653位时所述突变不是D653N，其中当所述突变位于L603位时所述突变不是L603A，其中当所述突变位于H594位时所述突变不是H594A，并且其中所述逆转录酶在大于42℃的温度时具有最佳活性。

2.如权利要求1所述的逆转录酶，所述逆转录酶在至少50℃的温度时具有最佳活性。

3.如权利要求1或2所述的逆转录酶，所述逆转录酶在50℃时的活性比相应野生型酶活性更高。

4.如任一项前述权利要求所述的逆转录酶，所述逆转录酶在37℃时的比活性为相应野生型酶的至少110％。

5.如任一项前述权利要求所述的逆转录酶，所述逆转录酶在50℃处理5分钟后，对在37℃的残留活性进行测定，所述逆转录酶的热稳定性为相应野生型酶的至少1.5倍。

6.如任一项前述权利要求所述的逆转录酶，所述逆转录酶具有一个或多个以下突变：

| | | | | | |
|---|---|---|---|---|---|
| L603W 或 M, | D200N、A 或 G, | L139P, | T330P, | E607K、G 或 A, | T287A, D524A |
| Q221R, | I49V 或 T, | N479D, | H594K、R 或 Q, | | |
| A502V, | D653G, A, H 或 V | K658R 或 Q | P130S, | Q237R, | |
| A307V, | Y344H, | Q430R, | D449G 或 A, | A644V 或 T, | N649S, |
| L671P, | E673G 或 K, | M39V 或 L, | Q91R 或 L, | M66L, | W388R, |
| I179T 或 V | L333Q | | R390W | Q374R 和 | E5K。 |

7.如任一项前述权利要求所述的逆转录酶，所述逆转录酶包含至少两个突变。

8.如权利要求7所述的逆转录酶，其中所述至少两个突变位于D200和L603。

9.如权利要求8所述的逆转录酶，其中所述至少两个突变是D200N和L603W。

10.如权利要求7所述的逆转录酶，其中所述至少两个突变位于N479和H594。

11.如权利要求10所述的逆转录酶，其中所述至少两个突变是N479D和H594K。

12.一种多核苷酸，所述多核苷酸编码如权利要求1～11中任一项所述的逆转录酶。

# 一种逆转录酶和编码其的多核苷酸

[0001] 本申请是分案申请，其原申请的国际申请号为PCT/EP2009/054329，中国国家申请号为200980121472.X，申请日为2009年4月9日，发明名称为“核酸的制备”。

## 技术领域

[0002] 本发明涉及一种逆转录酶和编码其的多核苷酸，具体而言，编码靶蛋白的核酸的制备方法和由此能够获得的靶蛋白，所述靶蛋白包括诸如核酸加工酶等酶，特别是逆转录酶。

## 背景技术

[0003] 蛋白进化是用于对来自大型文库的蛋白进行选择和定向进化的已知技术。选择的基本原则是确保特定表型(蛋白)和编码其的基因型之间存在联系形式。可以以三种不同的方式实现该表型-基因型联系形式：

[0004] • 共价连接诸如mRNA展示，和在一定程度上噬菌体展示、细菌展示、酵母展示等，

[0005] • 利用亲和性相互作用的非共价连接。实例是核糖体展示、CIS展示、质粒展示等，

[0006] • 区室化(compartmentalization)例如体外区室化(IVC)、区室化自我复制(CSR)、简单细菌筛选、高通量筛选等。

[0007] 如上述段落所指出，共价型表型-基因型连接的一个实例可以使用mRNA展示而实现。根据Roberts和Szostak(1997)所述，mRNA和其编码的肽或蛋白之间的共价融合可以通过合成的mRNA的体外翻译来产生，所述mRNA在其3’末端处携带嘌呤霉素，一种肽基受体抗生素。

[0008] 表型和基因型之间的非共价连接可以通过核糖体展示而实现。在核糖体展示中，使包括一种或多种编码靶蛋白的RNA在内的RNA的阵列通过体外翻译系统形成核糖体、mRNA和与tRNA偶联的蛋白的非共价三元复合物。三元复合物的这种阵列必须是稳定的。因此，体外翻译过程中形成的各三元复合物使用末端缺乏终止密码子的mRNA。所述三元复合物在低温(4℃)和高浓度的镁中(50mM)进一步得到稳定。在稳定的复合物中表型和基因型之间的连接得以保持。随后进行选择步骤，由此基于仍附着到三元复合物的蛋白的性质选择靶蛋白。然后可以使所选择的三元复合物解离，通过RT-PCR扩增与靶蛋白关连的mRNA。

[0009] 通常，核糖体展示可成功地应用于肽(Mattheakis等，1994；Matsuura和Pluckthun，2003)和蛋白(Hanes和Pluckthun，1997；He和Taussig，1997；Irving等，2001)的选择，所述肽和蛋白与不同的靶结合。在某些情况下，可以使用核糖体展示来选择酶促活性、使用自杀性抑制剂(Amstutz等，2002)或活性位点配体(Takahashi等，2002)进行蛋白的亲和选择。

[0010] 在体外区室化(IVC)中，通过对油包水乳液中的基因进行体外区室化而实现表型和基因型联系形式。计算用于制备乳液的基因的数量，使得大多数水区室含有不超过一个基因。使区室化的基因转录和翻译。然后评估所合成蛋白的活性。随后基于蛋白活性进行的选择，可以扩增编码具有所需性质的活性蛋白的DNA。大多数IVC应用中所用的水滴的尺寸

为2μm～3μm，反应体积为～5飞升，并且相对于每1ml乳液，有～$10^{10}$个油包水区室(50μl水相)。IVC选择系统的首个成功实例基于靶特异性DNA甲基化活性(Tawfik和Griffiths，1998)。将HaeIII甲基转移酶的基因区室化、转录和翻译。在辅因子存在下，体外合成的甲基转移酶能够将其自身的DNA甲基化。由于甲基化DNA(反应产物)抗HaeIII限制性内切核酸酶的消化，从$10^{7}$倍过量的其它DNA分子中选出了编码甲基转移酶的基因。

[0011] 迄今为止已经设计和实现了更多改进的IVC。进行IVC选择的最简单方式是使用DNA修饰酶，特别是DNA甲基转移酶(Lee等，2002；Cohen等，2004)。类似的实验策略用于选择限制性内切核酸酶FokI的活性变体(Doi等，2004)。对区室化的编码活性限制性内切核酸酶的DNA进行消化，并通过随后导入生物素-dUTP和与链亲和素结合来选择。

[0012] 对DNA聚合酶的进化应用不同的IVC选择策略(Ghadessy等，2001；Ghadessy等，2004；Ong等，2006)。这种新的选择方法基于编码活性DNA聚合酶的基因的‘区室化自我复制’(CSR)。与其中目的蛋白原位表达的常用IVC相反，CSR通过表达嗜热性DNA聚合酶的细菌细胞的区室化进行。对再悬浮于补充有引物和dNTP的PCR缓冲液中的细胞进行乳化，产生尺寸为～15μm的区室。每个微小水滴充当分离的PCR区室。起始PCR变性步骤期间，使细菌细胞破碎，将所表达的嗜热性DNA聚合酶和其编码基因释放到反应混合物中，使自我复制得以进行，同时其它细菌蛋白由于高温而变性。

[0013] IVC的改进是使用水包油包水双乳液。可以对被油层包围的微小水滴用FACS以>$10^{4}$个变体每秒的速率进行分析和分选(Bernath等，2004；Mastrobattista等，2005)。

[0014] 还可以通过IVC对用于结合的蛋白进行选择。油包水区室中表达的目的蛋白与编码其的基因共价偶联(Bertschinger和Neri，2004)或非共价偶联(Doi和Yanagawa，1999；Yonezawa等，2003；Sepp和Choo，2005)。

[0015] 已知的IVC的其它应用有包埋在区室中的微珠用作蛋白和基因偶联的中间体。附着在微珠上的单基因在微小水滴中进行转录和翻译。新合成的蛋白被捕获到反应区室内的相同微珠上。将乳液破碎后，可以将分离的珠进一步用于亲和选择(Sepp等，2002)。通常使用有机溶剂(己烷、醚、氯仿)破碎乳液，但这也会使得展示在所述珠上的某些酶的活性降低，限制了该技术的应用。为了选择催化活性，可以对微珠简单地进行洗涤，使其再悬浮于不同的反应缓冲液中，并通过第二乳化步骤再次进行区室化(Griffiths和Tawfik，2003)。然而有时刚性酶-珠-基因复合物(由于空间位阻和酶移动性限制)不能满足基本的反应要求，并且所附着的酶的活性可能低于游离酶的活性。另外，由于必须使用许多其它成分(即，亲和标签、抗体和珠)，所述方法在技术上是复杂的。

[0016] 对IVC中所用乳液的组成进行设计，从而确保水区室的稳定性和mRNA体外转录以及随后靶蛋白翻译的有效性。体外进化具有许多待改进的目标。一些目的蛋白和酶是稳健的勤奋工作者，并且在不同的缓冲液中，特别在IVC所用的反应混合物中具有足够的活性。然而，存在许多复杂的酶，这些酶仅在经优化的条件下或特定条件下才起作用。除此之外，体外进化的第一定律称-“进化所要选择的(you will evolve what you are selecting for)”。这就是说，转录/翻译反应混合物中进化和优化的酶会在该特定混合物中良好地起作用，并且很有可能在其自身缓冲液中表现差的多。在某些情况下，酶工作条件与区室中用于蛋白表达的体外转录和翻译混合物不相容。部分解决方案是用于将不同溶质运送到乳液区室中的纳米微滴递送系统(Bernath等，2005)。使用微流控装置可以完成使用油包水区室

的甚至更为精密复杂的操作。可以以高达10000个水性微滴每秒钟的速率制备高度单分散的单乳液或双乳液(Thorsen等,2001;Okushima等,2004)。可以将所产生的水区室运送到微流控通道(microfluidic channel)中,使其融合、再分和分选(Song等,2003;Link等,2006)。但是区室中的全缓冲液交换仍然是个问题。

[0017] 逆转录酶是用于从mRNA靶合成cDNA的非常重要的商业化酶。为了改善逆转录酶的性质,已经进行了大量研究。然而迄今为止适合于逆转录酶体外进化的合适的工作选择系统还是未知的。几乎所有的改进和逆转录酶的突变体的选择都是使用高通量筛选和合理设计而进行的。

[0018] 核糖体展示(RD)和体外区室化(IVC)都不能用于全活性逆转录酶的选择。用于核糖体展示的三元复合物在进行逆转录选择所需的较高温度下通常是不稳定的,并且因而会失去表型和基因型之间的连接。虽然使用合成的体外翻译提取物WakoPURE可以产生用于核糖体展示的相对稳定的三元复合物(Matsuura等,2007),但是固定在核糖体和mRNA上的体外翻译的逆转录酶在全长cDNA的合成过程中会遭遇到明显的空间位阻。还存在一种可能性,即固定的酶会反式以及顺式方式起作用,这会同样与蛋白进化策略不相容,原因在于不能保留表型-基因型连接。

[0019] IVC通常使用DNA作为遗传材料。体外的mRNA转录和靶蛋白翻译在DNA的存在下于空间上分离的经乳化水区室中进行。在体外进化逆转录酶的情况下,编码DNA序列的存在消除了选择逆转录酶活性的主要先决条件-必须从头合成cDNA。换言之,更好的逆转录酶变体的选择是基于酶从mRNA合成其自身编码cDNA的能力。新合成的cDNA必须通过PCR扩增,因此cDNA应该是反应中DNA的唯一来源。用于IVC选择的DNA会与cDNA一起扩增,省略了基础选择方案。

[0020] 体外区室化的更复杂的变体方式,例如使用包埋在水区室中的微珠(Sepp等,2002;Griffiths和Tawfik,2003),使得反应缓冲液完全交换。在该方法中,利用微珠实现表型-基因型联系,产生刚性选择单元mRNA-微珠-蛋白,在逆转录酶选择的情况下,该刚性选择单元mRNA-微珠-蛋白同样可以引起空间位阻,结果使得cDNA的合成效率差。

[0021] 使用噬菌体展示技术的改进方法选择能够合成～300个核苷酸长的cDNA的Taq DNA聚合酶(Vichier-Guerre等,2006)。虽然该方法起作用,但具有以下缺点:1)不是所有蛋白都可以在噬菌体上展示;2)绝对需要使用生物素标记的核苷酸以进行选择;3)所展示的酶能够以反式和顺式方式起作用;4)由于空间位阻和酶移动性限制,噬菌体-酶-DNA/RNA复合物会干扰cDNA的有效合成。

[0022] WO0222869中也提到了对逆转录酶进行选择的可能性,该申请涉及区室化自我复制(CSR)法。CSR技术用于选择嗜热性DNA聚合酶,特别是Taq DNA聚合酶(Ghadessy等,2001;Ghadessy等,2004;Ong等,2006)。将表达嗜热性DNA聚合酶的突变体文库的细菌细胞悬浮于PCR混合物中并进行乳化,产生分离的用于体外选择更具活性的聚合酶的PCR区室。

[0023] 细菌RNA酶的存在会阻止逆转录酶活性的实时选择,所述细菌RNA酶在中等温度下仍然有活性并会降解靶mRNA。还存在来自未经选择的质粒DNA(释放自细菌细胞)的DNA污染,并且存在所有大肠杆菌酶、结构蛋白、核糖体、NTP、RNA酶、DNA酶和小分子量分子。

## 发明内容

[0024] 本发明旨在提供蛋白进化的改进方法，该方法不具有现有技术方法的缺点。

[0025] 因此，在第一方面，本发明提供了编码靶蛋白的核酸的产生方法，所述方法包括：

[0026] (a)提供包括一种或多种编码所述靶蛋白的RNA或DNA分子在内的RNA或DNA分子的阵列；

[0027] (b)从所述阵列产生靶蛋白，从而形成RNA-蛋白复合物或DNA-蛋白复合物，其中所述RNA或DNA分子与所述复合物非共价或共价结合；

[0028] (c)将所述复合物分隔到区室中，其中大多数或全部区室含有不超过一个复合物；

[0029] (d)使所述复合物处于允许靶蛋白具有活性的反应条件下；和

[0030] (e)基于与靶蛋白相关的活性选择编码所述靶蛋白的核酸，

[0031] 其中当所述复合物是其中DNA非共价结合的DNA-蛋白复合物时，在不存在用于各复合物的分离的区室的情况下进行步骤b)。

[0032] 本发明人已经设计了一种方法，该方法使用两种不同类型的表型-基因型联系形式，并且获得了用于实验室内蛋白进化方法中使用的新型选择系统。下文进一步描述的该方法的新特征使得其可以应用到与现有技术方法相比范围更宽的靶蛋白上，使用更加简单，适应性增加。特别是，本发明首次提出了进化和改进逆转录酶性质的可能性，所述逆转录酶是分子生物学家工具箱中最重要的酶类之一。

[0033] 在该方面的一个实施方式中，本发明提供了编码靶蛋白的核酸的产生方法，所述方法包括：

[0034] (a)提供包括一种或多种编码所述靶蛋白的RNA或DNA分子在内的RNA或DNA分子的阵列；

[0035] (b)从所述阵列产生靶蛋白，从而形成RNA-蛋白复合物或DNA-蛋白复合物，

[0036] (c)将所述复合物分隔到区室中，其中大多数或全部区室含有不超过一个复合物；

[0037] (d)使所述复合物处于允许靶蛋白具有活性的反应条件下；和

[0038] (e)基于与靶蛋白相关的活性选择编码所述靶蛋白的核酸，

[0039] 其中在所述RNA-蛋白复合物中所述RNA与其非共价结合或共价结合，在所述DNA-蛋白复合物中所述DNA与其共价结合。

[0040] 在该方面的第二实施方式中，本发明提供了编码靶蛋白的核酸的产生方法，所述方法包括：

[0041] (a)提供包括一种或多种编码所述靶蛋白的RNA或DNA分子在内的RNA或DNA分子的阵列；

[0042] (b)从所述阵列产生靶蛋白，从而形成RNA-蛋白复合物或DNA-蛋白复合物，其中所述RNA或DNA分子与所述复合物非共价或共价结合；

[0043] (c)使所述复合物分隔到区室中，其中大多数或全部区室含有不超过一个复合物；

[0044] (d)使所述复合物处于允许靶蛋白具有活性的反应条件下；和

[0045] (e)基于与靶蛋白相关的活性选择编码所述靶蛋白的核酸，

[0046] 其中在不存在用于各复合物的分离的区室的情况下进行步骤b)。

[0047] 可以通过本领域已知的任何技术，例如mRNA展示、噬菌体展示、细菌展示或酵母展示，产生DNA或RNA与靶蛋白之间的共价连接。具体而言，使用mRNA展示技术可以产生共价RNA-蛋白连接，而使用共价抗体展示(CAD)技术可以产生共价DNA-蛋白连接(Reiersen等，

2005)，通过共价DNA展示(Bertschinger和Neri，2004)或使用类似的共价展示技术如Stein等(2005)所述的技术进行区室内的翻译。

[0048] 也可以通过本领域已知的任何技术，例如核糖体展示、CIS展示或质粒展示产生DNA或RNA与靶蛋白之间的非共价连接。具体而言，使用CIS展示在不存在区室时可以产生非共价DNA-蛋白连接(Odergrip等，2004)，而使用核糖体展示技术可以产生非共价RNA-蛋白连接。

[0049] 如上指出，当所述复合物是其中DNA为非共价结合的DNA-蛋白复合物时，在不存在用于各复合物的分离的区室的情况下进行本发明方法中的步骤b)。换言之，步骤b)是非区室化的。具体而言，当产生的复合物是其中DNA为非共价结合的DNA-蛋白复合物时，无需将阵列的各成员彼此分隔即可进行产生步骤。特别是，无需通过体外区室化(IVC)将阵列的各成员分隔即可进行产生步骤。在特别优选的实施方式中无需在油包水乳液中将阵列的各成员分隔即可进行产生步骤。

[0050] 还可以通过本领域已知的任何方法进行区室化，这些方法能够使得所述复合物产生或分离从而使得所有或基本上所有的区室含有不超过一个复合物。特别是，优选至少70％、至少80％或至少90％的区室含有不超过一个复合物。例如，可以通过将阵列的成员或各复合物分隔到位于微量滴定板或纳米滴定板(nanotiter plate)上的不同的孔中，或通过体外区室化(IVC)进行区室化。具体而言，通过IVC进行的分隔可以包括分隔到油包水乳液或水包油包水乳液中的水性微滴中。

[0051] 本发明方法组合了选自以下所列中的至少两种不同类型的基因型-表型联系形式：共价连接、非共价连接和区室化。在本发明的优选方面，所述方法利用这些联系形式中的至多两种。因此在特别优选的实施方式中所述方法利用共价连接或非共价连接作为步骤b)中的唯一表型-基因型联系形式。换言之，在该实施方式中步骤b)中不存在区室化。

[0052] 可以以许多不同的方式，例如通过核糖体展示、mRNA展示(Roberts和Szostak，1997)、CIS展示(Odergrip等，2004)或共价抗体展示(CAD)(Reiersen等，2005)在不存在区室时建立DNA或RNA与蛋白之间的共价/非共价连接。具体而言，通过使用CAD技术可以在不存在区室时实现共价DNA-蛋白连接，而通过mRNA展示和核糖体展示分别可以建立共价RNA-蛋白连接和非共价RNA-蛋白连接。

[0053] 可以通过许多不同联系方式的组合实现本发明。例如，可以通过核糖体展示和体外区室化的组合，或通过mRNA展示、CIS展示或CAD展示与IVC的组合实现本发明。

[0054] 在优选方面，通过核糖体展示和体外区室化的组合实现编码靶蛋白的核酸的产生方法。具体而言所述方法包括：

[0055] (a)提供包括一种或多种编码所述靶蛋白的mRNA在内的mRNA的阵列；

[0056] (b)在用于核糖体翻译的条件下对所述mRNA的阵列进行温育，从而产生三元复合物的阵列，每个三元复合物包含mRNA、核糖体和由所述mRNA翻译的蛋白；

[0057] (c)将三元复合物的阵列导入油包水或水包油包水乳液的水相微滴中，其中大多数或所有的水相微滴含有不超过一个三元复合物；

[0058] (d)使所述水相微滴处于允许靶蛋白具有活性的反应条件下；和

[0059] (e)基于与靶蛋白相关的酶活性选择编码所述靶蛋白的核酸。

[0060] 本发明的这种方法可以称为“区室化核糖体展示”(CRD)。CRD可应用于各种靶蛋

白，包括酶。CRD的优点在于所述酶和mRNA之间的联系方式是非共价的。因此如果步骤(d)中所用的反应条件涉及增高的温度，步骤(b)中产生的三元复合物会解离，并且会释放所述酶。这避免了所述酶固定在珠上的现有技术中上述与酶移动性相关的问题。

[0061] 在步骤(e)中可以以许多方式对内部具有核糖体展示复合物的乳液微滴进行分选或选择。优选的是通过荧光激活细胞分选法(FACS)或使用微流控技术对它们进行分选。两种技术主要利用了基于荧光的微滴分选。然而，取决于步骤(d)中所用的反应条件和要选择的蛋白活性，还可以通过尺寸、光衍射或光吸收对微滴进行分离。

[0062] 当所述靶蛋白是酶时优选使用基于荧光的分选方法。在该实施方式中步骤(d)中所用的反应条件包括能够转化成荧光产物的非荧光底物。所述酶的活性产生荧光产物，使得可以使用FACs来区分含有活性酶的荧光微滴和不含有活性酶或含有较弱活性酶的非荧光微滴或弱荧光微滴。

[0063] 尤其是，CRD可应用于诸如逆转录酶等核酸加工酶，并使得可以进行快速有效的体外进化。

[0064] 在另一方面，本发明提供了编码靶蛋白的核酸的产生方法，所述方法包括：

[0065] (a)提供包括一种或多种编码所述靶蛋白的mRNA在内的mRNA的阵列，其中所述mRNA包括酶的底物，所述酶包括靶蛋白或其辅酶；

[0066] (b)在用于核糖体翻译的条件下对所述mRNA的阵列进行温育，从而产生三元复合物的阵列，每个三元复合物包含mRNA、核糖体和由所述mRNA翻译的蛋白；

[0067] (c)将三元复合物的阵列和可选的辅酶导入油包水或水包油包水乳液的水相微滴中，其中大多数或所有的水相微滴含有不超过一个三元复合物；

[0068] (d)使所述水相微滴处于允许靶蛋白具有活性的反应条件下；和

[0069] (e)基于与靶蛋白相关的酶活性选择编码所述靶蛋白的核酸。

[0070] 在本发明该方面的一种实施方式中，当所述核酸加工酶是DNA依赖性DNA聚合酶时，步骤(a)中的mRNA可以连接到双链DNA衔接体(adaptor)分子，从而提供底物。

[0071] CRD多样性为～$10^9$至$10^{10}$个变体，并且受到IVC步骤限制。该新方法与能够用于筛选～$10^5$至$10^6$个逆转录酶的突变体的高通量筛选(HTS)相比，更加有效的多，耗时更少和更加廉价。CRD多样性比HTS高约4个数量级，因此通过区室化核糖体展示选择可以容易地寻找出被HTS错过的许多更有益的突变体。

[0072] 根据步骤(a)，提供通常是合成mRNA的mRNA的阵列，所述阵列包括编码所述靶蛋白的一个或多个阵列成员。当靶蛋白是逆转录酶时，所述mRNA包含酶的底物，所述酶包含靶蛋白或其辅酶。在随后的选择步骤(e)中，基于与靶蛋白相关的酶活性选择编码所述靶蛋白的核酸。这样，涉及了本发明的两个实施方式：一个实施方式中所述酶活性由靶蛋白提供，一个实施方式中在靶蛋白的存在下所述酶活性由所述靶蛋白的辅酶提供。在需要辅酶的实施方式中，与三元复合物的阵列一样，将辅酶导入步骤(c)的油包水乳液的水相微滴中。当所述酶包括所述靶蛋白时，不必将另外的辅酶导入水相微滴中。

[0073] 在所述方法的步骤(b)中，用核糖体处理mRNA的阵列以产生三元复合物的阵列，每个三元复合物包含mRNA、核糖体和由mRNA翻译的蛋白。可以在任何适合于通常的mRNA体外翻译的条件下，例如于核糖体展示技术中所用的条件下进行该步骤。在这点上，可以对所述三元复合物进行纯化，虽然这不是必须的。在这点上可以对所述三元复合物补充任何随后

酶活性所需的辅助底物(co-substrate)，然后通常将反应混合物乳化以获得约$10^{10}$个油包水区室，各区室的平均直径通常为约2μm～3μm。即使是小体积(25μl)的体外翻译反应也会产生约$10^{11}$～$10^{12}$个贮存核糖体复合物分子。通常的核糖体展示方法使用缺乏终止密码子的mRNA，虽然也可以存在终止密码子(Matsuura等，2007)。为了获得其中大多数或所有水相微滴含有不超过一个三元复合物的水相微滴，三元复合物的浓度与通常的核糖体展示技术所用的相应浓度相比必须减少约两个数量级。在所述方法的该步骤中仅使用非常低浓度的三元复合物。

[0074] 所述酶可以包括核酸加工酶，该核酸加工酶可以是RNA加工酶。所述核酸加工酶可以包括靶蛋白，并且可以选自核酸聚合酶、核酸连接酶和末端脱氧核苷酸转移酶。如本文进一步的详述，所述核酸聚合酶可以包括逆转录酶。在该实施方式中，编码逆转录酶的mRNA自身是该逆转录酶的底物。选择编码靶蛋白的核酸的步骤(e)包括选择通过逆转录酶的作用产生的cDNA，所述cDNA编码逆转录酶。

[0075] 当靶蛋白是核酸连接酶时，可以对能够连接RNA和RNA或连接DNA和DNA的RNA(DNA)连接酶进行选择。优选的是，允许酶具有活性的反应条件包含含有核酸连接子或衔接体的辅助底物，所述辅助底物还包括用于附着到配体结合伴侣的亲和配体或用于RT-PCR中经加工的mRNA的特异性扩增的序列标签。在第一种情况下，编码靶蛋白的mRNA与在导入有编码核酸连接酶的mRNA的水相微滴中的辅助底物连接。优选的是，所述亲和配体包含生物素，并且所述配体结合伴侣包含链亲和素。对编码所述靶蛋白的核酸进行选择的步骤包括通过附着到包含配体结合伴侣的固相而选择引入所述辅助底物的mRNA。在通常的方法中，将连接酶的突变体文库进行体外翻译，将纯化的三元复合物稀释并且在具有生物素标记的DNA/RNA连接子和/或衔接体的反应缓冲液中进行乳化。使乳液升温到37℃后，核糖体三元复合物解体(disassemble)。连接酶将被释放，并且mRNA的3'末端将可接触生物素标记的衔接体并用于随后的连接反应。仅编码连接酶的具有活性(或活性更高)的变体的生物素标记的mRNA在链亲和素珠上得到纯化并可以通过RT-PCR进行扩增。

[0076] 在第二种情况下，选择编码靶蛋白的核酸的步骤包括选择具有附着的序列特异性标签的mRNA，所述序列特异性标签可用作用于逆转录和随后的PCR的引物的选择性退火位点。

[0077] 在通常的方法中，将连接酶的突变体文库进行体外翻译，并且将纯化的三元复合物稀释并且在具有DNA/RNA连接子和/或衔接体的反应缓冲液中进行乳化。将乳液升温到37℃后，核糖体三元复合物解体。连接酶会被释放，并且mRNA的3'末端将可接触所述衔接体并用于随后的连接反应。仅编码连接酶的具有活性(或活性更高)的变体的RNA会具有在逆转录中所用引物的特异性退火所需的特异性连接子序列，并且可以通过RT-PCR进行有效地进行扩增。

[0078] 还可以对末端脱氧核苷酸转移酶(TdT)进行选择。该酶作用于RNA并且可引入脱氧核糖核苷酸、核糖核苷酸和核苷酸类似物等。在该实施方式中，允许酶具有活性的反应条件包括含有dNTP的辅助底物，所述辅助底物还包括用于附着到配体结合伴侣的亲和配体。与核酸连接酶一样，所述亲和配体可以是生物素，所述配体结合伴侣可以是链亲和素。对编码靶蛋白的核酸的选择可以包括通过附着到包含配体结合伴侣的固相来选择引入辅助底物的mRNA。可以将TdT的突变体文库进行体外翻译，并且需要将纯化的核糖体三元复合物进行

稀释并且在具有生物素标记的核苷酸(例如生物素-dUTP)的反应缓冲液中进行乳化。对于野生型酶最适的工作温度是37℃。在该温度下所述核糖体三元复合物会解体,并且mRNA的3'末端将可用于模板非依赖性聚合反应。引入生物素标记的核苷酸的TdT-编码mRNA在链亲和素珠上得到选择,并且随后可以进行逆转录和通过RT-PCR进行扩增。

[0079] 在另一个实施方式中,所述靶蛋白包括逆转录酶辅助酶(helper enzyme),例如解旋酶、焦磷酸酶、持续因子(processivity factor)、RNA结合蛋白或其它在逆转录酶存在下能够改善逆转录反应的蛋白。在该实施方式中,所述核酸加工酶包括逆转录酶,所述逆转录酶是导入油包水乳液的水相微滴中的辅酶。在选择编码靶蛋白的核酸的步骤(e)中,选择通过逆转录酶的作用产生的并且编码逆转录酶辅助酶的cDNA。水相中逆转录酶辅助酶的存在促进了编码所述辅助子的mRNA的逆转录。因此,被逆转录的mRNA形成编码所述辅助子的cDNA,并且可以被PCR扩增。

[0080] 在另一实施方式中,所述靶蛋白包括RNA酶抑制剂。在该实施方式中,所述核酸加工酶包含RNA酶。选择编码靶蛋白的核酸的步骤(e)包括选择未被RNA酶降解的mRNA。在该实施方式中,将RNA酶作为辅酶导入油包水乳液的水相微滴中。一旦反应条件允许酶具有活性,任何不含有有效RNA酶抑制剂的微滴会显示出RNA酶活性,由此会将mRNA降解。因此,编码在所用的反应条件下有效的RNA酶抑制剂的mRNA会继续存在。一般而言,对RNA酶抑制剂的突变体文库进行体外翻译,并将纯化的核糖体三元复合物稀释并在具有合适RNA酶的反应缓冲液中进行乳化。在替代性方案中,可以通过乳化微滴稍后递送RNA酶。对仅编码具有活性(或更稳定)的RNA酶抑制剂的mRNA进行纯化和通过RT-PCR进行扩增。

[0081] 区室化核糖体展示还可用于体外区室化中的反应缓冲液交换,所述外区室化中选择缓冲液与体外翻译混合物不相容,并且底物到产物的转化必须在严格受控的反应条件下进行。

[0082] 如本文讨论,基于与靶蛋白相关的酶活性而选择的编码靶蛋白的核酸可以是DNA或RNA。可以对所述阵列进行转化或扩增,从而形成DNA或RNA。在优选的方案中,对所述阵列进行转化或扩增,从而形成所述方法的步骤(a)中的mRNA的阵列,并进行一个或多个另外的步骤(b)～(e)的循环,从而进一步富集具有更大量的编码靶蛋白的mRNA的阵列。

[0083] 使水相微滴处于允许酶具有活性的反应条件的步骤(d)为选择步骤(e)提供了基础,所述步骤(e)中选择编码靶蛋白的那些核酸。在步骤(d)中可以使用各种反应条件以提供选择压力。在一个实例中,反应条件包括在野生型酶的最适温度之上的温度。这些反应条件可以用于选择比野生型酶具有更高热稳定性的突变酶或在该温度下反应速率更高的突变酶或温度-活性曲线发生改变的突变酶。突变酶可能需要以比野生型酶更高的敏感性运行,因为水相微滴中mRNA的浓度是约400pM。突变酶还必须更精确地运行。所有这些选择压力对于逆转录酶特别重要。像生理条件一样,反应条件可以包括缓冲液、诸如金属离子等其它因子的浓度和pH的变化。

[0084] 在更好的逆转录酶的CRD选择中,可以施加许多更加不同的选择压力:1)选择较不易于聚集的溶解度更高的酶—必须将三元复合物(乳化之前)与疏水性材料预温育,从而消除具有表面暴露的疏水性残基的蛋白;2)选择非常快速的酶—逆转录反应时间在选择循环期间需要逐渐减少;3)选择合成长cDNA的酶—逐渐延伸CRD中使用的mRNA文库,结果合成更长的cDNA;4)选择能够通过二级结构转录的酶—必须将二级结构形成序列导入到CRD中使

用的mRNA文库中;5)选择在不同于RT缓冲液的缓冲液中起作用的酶(例如在PCR中,一步RT-PCR缓冲液或具有变性剂的缓冲液)—必须在我们选择的缓冲液中进行CRD选择;6)选择能够导入核苷酸类似物的酶—必须在具有生物素标记的核苷酸类似物的RT缓冲液中进行选择,并随后在链亲和素珠上进行cDNA纯化。

[0085] 区室化核糖体展示(CRD)还适合于许多荧光激活细胞分选(FACS)应用。必须以核糖体展示形式展示目的蛋白。可选的是,应该对反应缓冲液中混合有非荧光底物(S)的经纯化(或仅仅稀释许多倍)的包含mRNA-核糖体-蛋白(tRNA)的三元复合物进行乳化,产生水包油包水双乳液(Bernath等,2004;Mastrobattista等,2005)。区室化酶的活性变体会将底物(S)转化成荧光产物(P),从而能够以FACS区分荧光微滴(内部为活性酶)和"暗"微滴(内部为失活酶)。与酶促反应必须在转录/翻译混合物中进行的之前公开的实例相反,CRD允许在更天然(为所需要的)的条件下进行完全缓冲液交换和活性酶选择(图10)。

[0086] 还可以使用CRD选择和进化热稳定性DNA聚合酶(图11)。必须以核糖体展示形式展示目的聚合酶。可选的是,可以使用经纯化(或仅仅稀释许多倍)的包含mRNA-核糖体-聚合酶的三元复合物来制备在PCR缓冲液中的具有逆转录酶(辅助酶)、dNTP和引物组的反应混合物。应该对反应液进行乳化,产生油包水乳液。在第一步RT中-逆转录酶必须合成cDNA,该cDNA随后用作第二步PCR的靶-利用核糖体展示的DNA聚合酶进行cDNA扩增。PCR中所用的引物之一可以具有生物素和非互补性5'末端,所述生物素用于可选的使用链亲和素珠的随后纯化。将RT-PCR乳液破乳后,新合成的DNA片段可通过生物素进行纯化,并使用新引物组再次扩增,该新引物组含有的一个引物的序列与cDNA不互补、但与第一次扩增反应中所用引物的5'部分相同(在cDNA背景下的DNA选择性扩增)。相对于活性较低的变体,DNA聚合酶的活性更高的变体将得到富集,并可用于进一步分析或下一轮选择(图11)。

[0087] 区室化核糖体展示技术的重要特征是:

[0088] 1).通过mRNA文库维持基因型,这对于选择RNA加工酶特别有用;

[0089] 2).选择单元是mRNA-核糖体-蛋白(tRNA)三元复合物,并且可以通过超速离心、凝胶过滤、亲和标签纯化和其它简单方法对其容易地进行纯化;

[0090] 3).可以交换反应缓冲液,由此能够将选择单元转移(经过纯化或未经纯化)到新的反应混合物,并且同时被稀释100～200倍(从而将乳化后核糖体复合物的数量调整到低于1个分子/反应区室);

[0091] 4).CRD的mRNA文库多样性仅受体外区室化的多样性指数限制,并且低于$10^{10}$的不同变体;

[0092] 5).一旦乳化,可以在4℃～94℃的宽范围的温度下进行选择反应,原因在于在这些温度下乳液是稳定的;

[0093] 6).如果在较高温度(30℃以上)进行选择,三元复合物会解离,但保持区室化从而不丧失基因型-表型联系,释放mRNA和体外翻译的蛋白。

[0094] 已经发现本发明的区室化核糖体展示方法对于促进新型逆转录酶的进化起特别良好的作用。作为实例,M-MuLV逆转录酶(Gerard等,1986,pRT601)可用于选择(该酶包括用于纯化的N末端His标签)。该逆转录酶的活性最适的温度是42℃,并且在高达50℃的温度下仍有活性。可以在本发明方法的步骤(d)中使编码M-MuLV逆转录酶的mRNA的阵列处于温育温度为50℃～60℃并且包括cDNA合成所需的引物和dNTP的反应条件下。在这些较高的温度

下，在4℃下稳定的贮存的核糖体复合物迅速解离，释放逆转录酶底物(mRNA)和酶到溶液中。

[0095] 在一个实施方式中，从核糖体复合物释放的体外翻译的逆转录酶M-MuLV具有作为间隔子的与在核糖体展示构建中所用的λ噬菌体外表面蛋白D融合的C末端，从而保留在核糖体隧道(Matsuura和Pluckthun，2003)和共价结合的tRNA中，原因在于翻译未正确地终止。蛋白D是非常良好表达的可溶且稳定的蛋白，解折叠转变温度为～57℃(Forrer和Jaussi，1998)，因此是用于选择热稳定性逆转录酶的良好融合伴侣。

[0096] 在区室化核糖体展示选择方法中，只有具有完全活性的逆转录酶的变体可以执行cDNA的合成，该cDNA编码相同活性的酶。在1小时内，逆转录反应完成后，将乳液破乳，对cDNA进行纯化并通过巢式PCR进行扩增。由于CRD的选择性质，只有编码能够执行全长cDNA合成的逆转录酶的活性变体的cDNA会被扩增，并且必要的话可被转移到下一轮选择。为了消除T7聚合酶启动子区域、核糖体结合位点(RBS)和蛋白D序列中不期望的突变，可以将仅编码M-MuLV序列的扩增DNA连接到天然5'和3'末端片段，从而恢复最初的核糖体展示构建体，并可以进行下一轮选择。

[0097] 在5轮选择中已经鉴定出在50°下具有比活性的酶变体，与用于文库制备的原始酶的活性相比，该变体的活性高2倍至4倍。一些蛋白更快速，一些蛋白具有更高的热稳定性。许多选择的M-MuLV变体具有D524G或D583N突变，这些突变关闭了逆转录酶的RNA酶H活性，改善了cDNA合成以及热稳定性(Gerard等，2002)。许多更多的所选择的M-MuLV逆转录酶的变体具有之前提到和描述(US7056716；US20060094050A1；US7078208；US20050232934A1；WO07022045A2)的其它不同的有益突变(H204R；H638R；T197A；M289V；E302K；T306A；N454K；Y64C；E69G；Q190R；V223M；F309S；L435P；E562K)。除此之外，本发明人还发现了逆转录酶氨基酸序列中的许多新的热点。一些突变的重复频率非常高，并且具有非常重要的意义，这对经纯化的突变体进行分析时可以显示(实施例2)。因此本发明人可以声明，本发明的CRD技术是非常快速而且稳健(fast and robust)的选择方法，通过直接进化和M-MuLV逆转录酶的改进可以确认其有效性。作为原理的验证，我们已选择出在更高温度下工作更好的M-MuLV逆转录酶的变体。

[0098] 在另一方面，本发明提供了通过本文所述方法可获得的逆转录酶。

[0099] 在另一方面，本发明提供了在超过42℃，优选至少50℃，更优选50℃～60℃的温度下具有最佳活性的逆转录酶。通过应用具有较高温度，优选至少50℃的反应条件根据本文所述方法可以选择逆转录酶。在本发明的这一方面，可以选择与野生型酶相比活性-温度曲线迁移的逆转录酶，所述活性-温度曲线迁移使观察到的最佳活性的温度升高。

[0100] 在另一方面，本发明提供了一种逆转录酶，该逆转录酶包含在以下氨基酸位置中的一处或多处具有突变的MMLV逆转录酶氨基酸序列：

[0101]

| | | | | | |
|---|---|---|---|---|---|
| D200, | D653, | L603, | T330, | L139, | Q221, |
| T287, | I49, | N479, | H594, | F625, | H126, |
| A502, | E607, | K658, | P130, | Q237, | N249, |
| A307, | Y344, | Q430, | D449, | A644, | N649, |
| L671, | E673, | M39, | Q91, | M66, | W388, |
| I179 | E302 | L333 | R390 | Q374 和 | E5 |

[0102] 当突变位于D653处时，优选该突变不是D653N。当突变位于L603处时，优选该突变

不是L603A。另外，当突变位于H594处时，优选该突变不是H594A。

[0103] 优选位于上述位置处的突变是点突变。

[0104] 优选的是，所述逆转录酶具有以下突变中的一个或多个突变：

[0105]

| | | | | | |
|---|---|---|---|---|---|
| D200N,A 或 G, | D653N,G,A,H 或 V， | L603W 或 M, | T330P, | L139P, | Q221R, |
| T287A, | I49V 或 T, | N479D, | H594R 或 Q, | F625S 或 L， | H126S 或 R, |
| A502V, | E607K, G 或 A, | K658R 或 Q | P130S, | Q237R, | N249D, |
| A307V, | Y344H, | Q430R, | D449G 或 A, | A644V 或 T， | N649S, |
| L671P, | E673G 或 K, | M39V 或 L, | Q91R 或 L, | M66L, | W388R, |
| I179T 或 V | E302K | L333Q | R390W | Q374R 和 | E5K |

[0106] 发现这些突变中的每个，例如使得突变酶在50℃时与相应的野生型酶相比具有更高的活性。这些突变的进一步详细描述记载于具体实施例中。

[0107] 在本发明特别优选的方面，突变酶具有至少两个突变。在一个实施方式中所述两个突变位于D200和L603处。例如所述突变为D200N和L603W。在替代性实施方式中所述突变位于N479和H594。例如所述突变为N479D和H594R。

[0108] 在另一方面，本发明提供了在大于37℃的温度具有最佳活性的逆转录酶，其中50℃下的活性是37℃下活性的至少120％。优选的是，50℃下的活性是37℃下活性的至少130％，更优选至少160％。

[0109] 在另一方面，本发明提供了50℃下活性为相应野生型酶活性的至少两倍的突变逆转录酶。

[0110] 在另一方面，本发明提供了37℃下比活性为相应野生型酶活性的至少130％的突变逆转录酶。优选的是，所述突变逆转录酶的比活性是相应野生型酶比活性的至少140％，更优选至少150％，特别优选至少160％。如本文所述已经发现部分纯化的野生型酶的37℃的比活性为约200000U/mg。在具体实施例中会更详细地讨论根据本发明获得的具体突变逆转录酶。

[0111] 在另一方面，本发明提供了热稳定性为相应野生型酶热稳定性的至少1.5倍的突变逆转录酶。本申请中根据在50℃下处理5分钟后在37℃下的残留活性来测定热稳定性。优选的是，突变逆转录酶的热稳定性是相应野生型酶热稳定性的至少1.5倍，更优选至少2倍，还更优选至少2.5倍。通常，野生型逆转录酶37℃下的残留活性是未经处理酶的约11％。

[0112] 优选本发明的逆转录酶包括MMLV逆转录酶。

[0113] 在另一方面，本发明提供了编码本文所述逆转录酶的多核苷酸，例如mRNA或DNA。

[0114] 本发明的逆转录酶可以用于各种分子生物学技术，如RT-PCR(qRT-PCR等)。可以提供一种用于RT-PCR的试剂盒，所述试剂盒中的逆转录酶是本发明的逆转录酶。

## 附图说明

[0115] 现在参考仅作为示例的附图和附录，更详细地描述本发明。

[0116] 图1.实施例1的实验流程图。使用两种质粒pET_his_MLV_pD(编码与蛋白D间隔子融合的莫洛尼氏鼠白血病病毒(M-MLV)逆转录酶)和pET_his_del_pD(编码与蛋白D间隔子融合的失活(pol结构域中57个氨基酸缺失)的莫洛尼氏鼠白血病病毒(M-MLV)逆转录酶)来合成PCR片段。PCR片段继续用于转录反应和合成3'末端缺少终止密码子的mRNA。将经纯化的mRNA以1:50＝MLV(活性RT):del(失活RT)的比例混合并用于体外翻译反应。翻译反应期

间，核糖体复合物合成蛋白，并在缺少终止密码子的mRNA的末端停止。通过蔗糖垫超速离心对三元复合物(TC)的混合物进行纯化。使用已经含有与体外翻译的MLV逆转录酶相连的mRNA的经纯化三元复合物(取<$3\times10^9$个分子)来制备逆转录反应混合物，该逆转录反应混合物补充有用于RT反应的外部dNTP组和引物。对冰冷的RT反应混合物进行乳化，产生～$1\times10^{10}$个大小为～2μm的油包水区室。为了进行RT反应，对乳化RT反应混合物(低于一个TC(mRNA+MLV RT)/区室)于42℃下温育1小时。区室化RT反应混合物的温度升高后，大多数TC解离，释放mRNA和逆转录酶。仅在含有活性MLV逆转录酶(MLV_pD)的区室中才进行成功的RT反应，在具有失活逆转录酶(del_pD)的区室中未合成cDNA。随后PCR扩增cDNA，并观察到活性逆转录酶(MLV_pD)基因相对于失活逆转录酶(del_pD)的富集。

[0117] 图2.pET_his_MLV_pD质粒的示意图。

[0118] 图3.实施例1-对CRD选择期间合成的cDNA进行的第一次PCR的琼脂糖凝胶电泳。所用引物:RD_Nde(SEQ ID No:9)和pD_55(SEQ ID No:10)。对于MLV_pD，PCR片段的预期长度是2185bp，对于del_pD是2014bp。在每孔加载10μl PCR混合物的1％琼脂糖凝胶上进行扩增分析。

[0119] 图4.实施例1-对第一次PCR产物进行的用于部分基因扩增的巢式PCR的琼脂糖凝胶电泳。所用引物:M_F(SEQ ID No:11)和M_2R(SEQ ID No:12)。对于MLV_pDa，PCR片段的预期长度是907bp，对于del_pD是736bp。在每孔加载10μl PCR混合物的1％琼脂糖凝胶上进行扩增分析。

[0120] 图5.实施例1-对第一次PCR产物进行的用于全基因扩增的巢式PCR的琼脂糖凝胶电泳。所用引物:M_Esp(SEQ ID No:13)和M_Eri(SEQ ID No:14)。对于MLV_pDa，PCR片段的预期长度是2077bp，对于del_pD是1906bp。在每孔加载10μl PCR混合物的1％琼脂糖凝胶上进行扩增分析。

[0121] 图6.实施例2中CRD选择的实验流程图。使用编码逆转录酶(与蛋白D融合)MLV_pD的突变体文库的PCR片段来合成mRNA。经纯化的mRNA用于体外翻译反应。在翻译混合物中形成mRNA-核糖体-MLV_pD(tRNA)的三元复合物(TC)，并通过低温和高浓度的$Mg^{2+}$离子将其稳定。通过蔗糖垫超速离心对TC的混合物进行纯化。将沉淀的TC溶解于冰冷的缓冲液(50mM $Mg^{2+}$)中并用于制备逆转录反应混合物，该逆转录反应混合物补充有用于RT反应的外部dNTP组和引物。对冰冷的RT反应混合物进行乳化，产生～$1\times10^{10}$个大小为～2μm的油包水区室。MLV RT的最适反应温度是～42℃。为了选择在更高温度下工作更好的逆转录酶变体，使乳化的RT反应混合物(低于一个TC(mRNA+MLV RT)/区室)于50℃下温育1小时。在该温度下在含有更具活性或热稳定性的MLV逆转录酶变体的区室中全长cDNA的成功合成进行地更好。随后的PCR用于扩增全长cDNA，并进行更具活性和热稳定性的逆转录酶基因的富集。通过PCR，将扩增的基因回复到CRD形式，通过连接PCR恢复完整的5'(起始片段-T7聚合酶启动子、SD和his-标签编码序列)和3'(末端片段-gs连接子、蛋白D和第二gs连接子)序列。含有逆转录酶基因的富集文库的重建PCR片段被用于随后的mRNA转录和下一轮CRD选择。在温度越来越高的RT反应中进行各轮选择:50℃(第一轮);52.5℃(第二轮);55℃(第三轮);57.5℃(第四轮)和60℃(第五轮)。

[0122] 图7.新一轮CRD选择之前PCR片段的重建流程图。用Esp3I(NcoI相容末端(compatible end))和EcoRI消化突变的MLV RT文库，并将其与起始片段(244bp)和末端片

段(398bp)连接,从而获得适合CRD选择的PCR片段。通过对最初的983bp起始片段(靶-质粒pET_his_del_pD(SEQ ID No:2)、引物-pro-pIVEX(SEQ ID No:3)和M_1R(SEQ ID No:15))进行PCR扩增和随后用NcoI(识别序列C↓CATGG)消化,构建起始片段(含有T7聚合酶启动子、SD和his-标签编码序列),产生244bp的DNA片段。通过对最初的1039bp末端片段(靶-质粒pET_his_del_pD(SEQ ID No:2),引物-M_3F(SEQ ID No:16)和pD-ter(SEQ ID No:4))进行PCR扩增和随后用EcoRI(识别序列G↓AATTC)消化,构建末端片段(含有gs连接子、蛋白D和第二gs连接子序列),产生398bp的DNA片段。

[0123] 图8.37℃、50℃下测定的突变RT变体的逆转录酶活性和50℃下温育5分钟后在37℃下的残留活性。将37℃下的逆转录酶活性标准化成总是100%并且将其省略。因此仅显示两种类型的柱(50℃下RT活性的百分比和50℃下温育5分钟后在37℃下的残留RT活性的百分比)。作为对照,给出用于突变体文库构建的wt M-MuLV逆转录酶。该起始酶在与RT突变变体相同的载体中表达并以相同的方式进行纯化。对于所有测试突变体50℃下的突变RT活性的平均值约为~92%,超过wt酶(45%)的2倍以上。50℃下预温育5分钟后在37℃下突变RT变体的平均残留活性为12%(wt酶~11%)。

[0124] 图9. 37℃下10分钟测定的部分纯化的wt和突变RT变体的比活性(u/mg蛋白)。

[0125] 图10.所提出的使用FACS进行CRD选择的实验流程图。以核糖体展示形式展示目的蛋白。使包含mRNA-核糖体-蛋白(tRNA)的经纯化(或仅仅稀释许多倍)的三元复合物与非荧光底物(S)在反应缓冲液中混合,并随后进行乳化,产生水包油包水双乳液。区室化酶的活性变体会将底物(S)转化成荧光产物(P),使得FACS得以区分荧光微滴(内部为活性酶)和"暗"微滴(内部为失活酶)。

[0126] 图11.所提出的使用CRD对热稳定性DNA聚合酶进行选择和进化的实验流程图。必须以核糖体展示形式展示目的蛋白。可选经纯化(或仅仅稀释许多倍)的包含mRNA-核糖体-聚合酶的三元复合物可用来制备在PCR缓冲液中的具有逆转录酶(辅助酶)、dNTP和引物组的反应混合物。应该对反应液进行乳化,产生油包水乳液。在第一步RT中-逆转录酶必须合成cDNA,该cDNA随后用作第二步PCR的靶—利用核糖体展示的DNA聚合酶进行cDNA扩增。PCR中所用的引物之一可以具有生物素和非互补性5'末端,所述生物素用于可选的使用链亲和素珠进行的随后纯化。将RT-PCR乳液破乳后,新合成的DNA片段经由生物素进行纯化,并使用新引物组再次扩增,该新引物组含有序列与cDNA不互补、但与第一扩增反应(DNA相对于cDNA背景的选择性扩增)中所用的引物的5'部分相同的一种引物。相对于活性较低的变体,DNA聚合酶的活性更高的变体会得到富集,并可用于进一步分析或下一轮选择。

[0127] 图12.实施例4的实验流程图。在该实验计划中将M-MuLV逆转录酶用作DNA依赖性DNA聚合酶。使用两种质粒,pET_his_MLV_D583N_pD(编码减去RNA酶H的与蛋白D间隔子融合的莫洛尼氏鼠白血病病毒(M-MLV)逆转录酶)和pET_his_del_pD(编码与蛋白D间隔子融合的失活逆转录酶-在pol结构域中57个氨基酸缺失,并且在RNA酶H结构域中存在点突变D583N),来合成PCR片段。PCR片段继续用于转录反应。将经纯化的mRNA以1:20=MLV_D583N_pD(活性RT):del_pD(失活RT)的比例混合,并用于通过使用T4DNA连接酶将dsDNA连接到mRNA混合物上来制备mRNA/dsDNA复合物。mRNA/dsDNA复合物用于体外翻译反应。翻译反应期间,核糖体复合物合成蛋白,并在mRNA的末端(位于mRNA/DNA杂交物的开始)处停止。通过蔗糖垫超速离心对三元复合物(TC)的混合物进行纯化。使用已经含有与体外翻译的聚合酶

(M-MuLV逆转录酶)连接的mRNA/dsDNA的经纯化三元复合物(取<$3\times10^9$个分子)来制备补充有外部生物素-dUTP的延伸反应混合物。对冰冷的反应混合物进行乳化,产生~$1\times10^{10}$个大小为~2μm的油包水区室。为了导入生物素化核苷酸,将乳化延伸反应混合物(低于一个TC(mRNA/dsDNA+聚合酶)/区室)于37℃下温育30分钟。区室化反应混合物的温度升高后,大多数TC解离,释放mRNA/dsDNA复合物和聚合酶。仅在含有活性聚合酶(逆转录酶-MLV_D583N_pD)的区室中才进行成功的导入dsDNA底物中的反应,在具有失活聚合酶(del_pD)的区室中未合成cDNA。将乳液破乳后,使用凝胶过滤微柱除去过量的生物素-dUTP。在链亲和素珠上对生物素化的mRNA/dsDNA复合物进行纯化并将其用于合成cDNA。随后PCR扩增cDNA,观察到活性聚合酶(逆转录酶-MLV_D583N_pD)基因相对于失活聚合酶(del_pD)的富集。

[0128] 图13.确定生物素-dUTP导入到mRNA/dsDNA复合物和导入到自为引物的mRNA中的效率。对导入dTTP或生物素-dUTP后的mRNA/dsDNA(MLV_D583N_pD)和mRNA(del_pD)样品进行的RT-PCR的图像。对于MLV_D583N_pD,预测的扩增子大小为907bp,对于del_pD cDNA为736bp。在每孔加载10μl PCR混合物的1%琼脂糖凝胶上分析PCR产物。

[0129] 图14.通过导入dTTP(生物素-dUTP)和[α-$P^{33}$]dATP进行的mRNA/dsDNA复合物存在的一般对比。在导入最初的dTTP或生物素-dUTP后,应该将放射性dATP导入到dsDNA底物中。A-用溴化乙锭可视化的琼脂糖凝胶(mRNA或mRNA/dsDNA带-2.5kb)。B-滤纸上干燥的与A相同凝胶,并且使用Cyclone Phosphor Imager(Perkin-Elmer,Wellesley,MA)使其可视化。观察到标记的mRNA/dsDNA复合物和/或仅观察到dsDNA带。C-mRNA/dsDNA复合物中dsDNA对应物(counterpart)的结构和序列。

[0130] 图15.实施例4的所得最终RT-PCR片段的分析。对于MLV_D583N_pD,预测的扩增子大小为907bp,对于del_pD cDNA为736bp。在每孔加载10μl PCR混合物的1%琼脂糖凝胶上分析PCR产物。进行链亲和素珠纯化前的RT-PCR样品相当于活性聚合酶基因与失活聚合酶基因的比例为1:20(几乎只有del_pD片段~736bp可见)。在链亲和素珠上纯化后的RT-PCR样品相当于活性聚合酶基因与失活聚合酶基因的一轮选择后的比例为1:1。在该轮选择中观察到~20的富集倍数。

[0131] 图16.用于确定1kb和4.5kb cDNA合成反应的最高温度的碱性琼脂糖凝胶的一些实例。PCR仪-Eppendor Mastercycle Gradient。A-D-1kb cDNA合成(M-MuLV(wt)、D200N、L603W和Q221R);温度梯度41.9℃、43.6℃、45.5℃、47.8℃、50.4℃、53.1℃、55.8℃、58.1℃、60.1℃、62.1℃;尺寸标准(size standart)-DNA Fast Ruler中等范围(Fermentas)。E-G-4.5kb cDNA合成(M2、M3和M4);温度梯度49.8℃、51.5℃、53.4℃、55.7℃、58.3℃、61.0℃、63.7℃、66.1℃、68.0℃、70.0℃;尺寸标准-Zip Ruler Express DNA ladder 2(Fermentas)。

[0132] 附录1.起始MLV RT文库(NcoI和EcoRI限制性位点之间的序列-SEQ ID No.24)中发现的突变的概览。于10个测序基因中发现23个核苷酸突变(1个颠换、20个转换-突变位置加下划线,在序列之上标出突变,2个缺失-加下划线并在序列之上标示为虚线),产生15个氨基酸交换、6个沉默突变、1个终止密码子和2个编码框的移码-平均每个基因1至2个氨基酸取代。

[0133] 附录2.为了与文献中通常所用的氨基酸编号相同,对所有104个蛋白序列进行不带N末端His标签的CLUSTALW比对。表示为MLV(SEQ ID No:25)的野生型序列总是以第一序

列给出(按照显示顺序突变序列表示SEQ ID Nos:26～128)。使用于黑色背景中的白色字体标记突变。发生了某种程度上改善M-MuLV逆转录酶性质和描述于不同的专利申请中的突变处的氨基酸位置在比对中标记为以灰色突出显示的氨基酸(白色字体)的柱。源自我们的选择并且位于灰色柱中的突变表明,我们的选择方法精确地靶向有益的热点或甚至别处所述的确切氨基酸突变。与初始wt M-MuLV相比50℃下的活性实质上更好(与wt活性的45%相比,为70%以上)的所分析蛋白的序列用灰色突出显示。

[0134] 附录3.所有选择的RT变体中发现的突变的列表。按照突变数量的递减对蛋白进行排列。

[0135] 附录4.所选择RT变体的突变频率(降序)。与初始wt M-MuLV相比50℃下的活性实质上更好(与wt活性的45%相比,为70%以上)的所分析蛋白的名称用灰色突出显示。

[0136] 附录5.关于M-MuLV(wt)逆转录酶和单突变体的数据的总结表,所述表包含:蛋白名称;选择频率(具有确切突变的测序突变体的数量,括号中的数字表示选择中发现的具体氨基酸突变的总数);蛋白浓度(mg/ml);37℃下的逆转录酶比活性(u/mg);50℃下的相对活性(%);在50℃下酶温育5分钟后在37℃下的相对残留活性(%);蛋白的RNA酶H比活性(u/mol);相对RNA酶H活性(%)以及1kb cDNA合成反应的最高温度。

[0137] 附录6.关于M-MuLV(wt)逆转录酶和单突变体的数据的总结表,所述表包含:蛋白名称;蛋白浓度(mg/ml);37℃下的逆转录酶比活性(u/mg);50℃下的相对活性(%)以及1kb和4.5kb cDNA合成反应的最高温度。

[0138] 附录7.SEQ ID Nos:1～23相关的序列和信息。

## 具体实施方式

[0139] 实施例1-CRD-原理验证

[0140] 为了提供区室化核糖体展示选择系统的原理验证,进行测试选择。通常的原理验证实验对于活性酶(对于我们的情况是起始MLV逆转录酶的RT-PCR片段)应该给出阳性信号和对于失活酶没有信号(对于失活的MLV逆转录酶没有RT-PCR片段)。更复杂的实验使用编码活性酶和失活酶的基因的规定比例的混合物。作为成功实验的结果,编码活性酶的基因应该相对于编码失活酶的基因富集。

[0141] 实验的一般流程图示于图1。使用两种质粒,pET_his_MLV_pD(编码与蛋白D间隔子融合的莫洛尼氏鼠白血病病毒(M-MLV)逆转录酶)和pET_his_del_pD(编码与蛋白D间隔子融合的失活(pol结构域中57个氨基酸缺失)的莫洛尼氏鼠白血病病毒(M-MLV)逆转录酶)来合成PCR片段。进而PCR片段用于合成mRNA的转录反应,所述mRNA的3'末端缺少终止密码子。将从两个上述PCR片段产生的经纯化的mRNA以1:50=MLV(活性RT):del(失活RT)的比例混合并用于体外翻译反应。翻译反应期间,核糖体复合物合成蛋白,并在缺少终止密码子的mRNA的末端停止。通过用冰冷的含有50mM $Mg^{2+}$的缓冲液进行稀释来终止翻译反应。低温、高浓度的$Mg^{2+}$离子和在mRNA的末端不存在终止密码子使得mRNA-核糖体-蛋白(tRNA)的三元复合物(TC)稳定化。通过蔗糖垫超速离心对三元复合物(TC)的混合物进行纯化。对超速离心进行优化,以使TC(~3.5MDa)沉淀于超速离心管的底部,同时小分子量的分子、蛋白和大多数游离mRNA(～0.9MDa)留在上清液中。将沉淀的TC溶解于冰冷的缓冲液(50mM $Mg^{2+}$)中。使用已经含有与体外翻译的MLV逆转录酶相连的mRNA的经纯化三元复合物(取$<3\times10^{9}$个分

子)来制备逆转录反应混合物,该逆转录反应混合物补充有用于RT反应的外部dNTP组和引物。对冰冷的RT反应混合物进行乳化,产生~1×$10^{10}$个大小为~2μm的油包水区室。为了进行RT反应,对乳化RT反应混合物(低于一个TC(mRNA+MLV RT)/区室)于42℃下温育1小时。区室化RT反应混合物的温度升高后,大多数TC解离,释放mRNA和逆转录酶。仅在含有活性MLV逆转录酶(MLV_pD)的区室中才进行成功的RT反应,在具有失活逆转录酶(del_pD)的区室中未合成cDNA。随后的PCR确保了所合成的cDNA的扩增,并且观察到活性逆转录酶(MLV_pD)基因相对于失活逆转录酶(del_pD)的富集。

[0142] 方法和材料

[0143] 通过对pET型质粒的T7聚合酶启动子和Shine-Dalgarno序列区域进行修饰,并且插入具有N末端His-标签(SEQ ID No:1的258-305)和与甘氨酸-丝氨酸(gs)连接子(SEQ ID No:1的2364-2393)、来自λ噬菌体的蛋白D的部分(pD)(SEQ ID No:1的2394-2669)和第二甘氨酸-丝氨酸(gs)连接子(SEQ ID No:1的2670-2759)融合的C末端的MLV H+逆转录酶编码序列(SEQ ID No:1的306-2363),构建起始质粒pET_his_MLV_pD(SEQ ID No:1和图2)。N末端His-标签用于蛋白表达纯化。在蛋白体外翻译和形成mRNA-核糖体-MLV(tRNA)三元复合物期间C末端融合必须保留在核糖体隧道中。

[0144] M-MuLV逆转录酶具有两种主要的酶促活性:RNA依赖性DNA聚合酶和RNA酶H。导入点突变D583N(在质粒pET_his_MLV_pD,SEQ ID No:1中的2055位处单核苷酸G交换为A)使逆转录酶RNA酶H活性关闭。天冬氨酸583位于RNA酶H活性位点,参与Mg离子结合并且对于RNA酶H活性是关键的。新质粒鉴定为pET_his_MLV_D583N_pD,并进一步用于下一质粒pET_his_del_pD(SEQ ID No:2)的构建,所述下一质粒编码失活的逆转录酶。用限制性内切核酸酶XmaJI(识别序列C↓CTAGG-SEQ ID No:1的1047和1218位)消化质粒pET_his_MLV_D583N_pD。除去长度为171bp的基因片段,使经消化的质粒自连接,产生pET_his_del_pD(SEQ ID No:2),pET_his_del_pD编码缩短171个核苷酸或57个氨基酸的逆转录酶基因,在蛋白翻译框上未发生移码。

[0145] 重要的是使得相同的逆转录酶基因具有以下特征:1)长度缩短(用于容易PCR检测);2)失活(实验上确认聚合酶结构域中的57个氨基酸的缺失使聚合酶活性完全失活,以及突变D583N使RNA酶H活性失活)和3)没有移码(任何移码会导致终止密码子的出现,这与核糖体展示形式不相容)。

[0146] 用于体外转录的PCR片段的制备。在冰上制备PCR混合物:20μl—10X具有KCl的Taq缓冲液(Fermentas);20μl—2mM的各dNTP(Fermentas);12μl—25mM $MgCl_2$(Fermentas);16μl—DMSO(D8418-Sigma);4μl—1u/μl LC(重组)Taq DNA聚合酶(Fermentas);1μl—100μM pro-pIVEX引物(SEQ ID No:3);1μl—100μM pD-ter引物(SEQ ID No:4);122μl水—将混合物分成2x 98μl的两试管。向2x 98μl PCR主混合物(master mix)中添加2μl pET_his_MLV_pD(稀释到~1ng/μl)或2μl pET_his_del_pD(稀释到~1ng/μl)。循环方案为:在94℃下进行起始变性步骤3分钟,30个循环(94℃下45秒钟,53℃下45秒钟和72℃下2分钟)和最后在72℃下延伸5分钟。从2ng质粒(7873bp)至~5μg(50ng/μl)扩增产物(对于pET_his_MLV_pD为2702bp PCR片段;对于pET_his_del_pD为2531bp PCR片段),扩增了~7000倍。

[0147] 制备转录混合物:40μl—5x T7转录缓冲液(pH为7.6的1M HEPES-KOH;150mM乙酸镁;10mM亚精胺;0.2M DTT);56μl—25mM各NTP(Fermentas);8μl—20u/μl T7RNA聚合酶

(Fermentas)；4μl—40u/μl RiboLock RNA酶抑制剂(Fermentas)；52μl不含核酸酶的水—将混合物分成2x 80μl的两试管，并且添加20μl—50ng/μl的MLV_pD(pro-pIVEX//pD-ter)或20μl—50ng/μl的del_pD(pro-pIVEX//pD-ter)PCR混合物。转录在37℃下进行3小时。

[0148] 用冰冷的不含核酸酶的水将两种转录混合物都稀释到200μl，并且添加200μl6M LiCl溶液。将混合物在+4℃下温育30分钟，在+4℃下于冷却的离心机中以最大速度(25'000g)离心30分钟。弃去上清液，用500μl冰冷的75％乙醇洗涤RNA团块。再次将试管在+4℃下于冷却的离心机中以最大速度离心5分钟，并弃去上清液。在室温下对RNA团块进行干燥12分钟，并随后通过在+4℃和1400rpm下振摇15分钟使其再悬浮于200μl不含核酸酶的冰冷水中。再次将试管在+4℃下于冷却的离心机中以最大速度离心5分钟，从而分离未溶解的RNA。将约180μl的上清液移到具有20μl 10X DNA酶I缓冲液($Mg^{2+}$) (Fermentas)；1μl—1u/μl的DNA酶I(不含RNA酶) (Fermentas)的新试管中，并且在37℃下温育20分钟，从而降解DNA。向各试管添加20μl pH为5.0的3M乙酸钠溶液和500μl冰冷的96％乙醇。最后通过在-20℃下温育30分钟，并在+4℃下于冷却的离心机中以最大速度(25'000g)离心30分钟使RNA沉淀。弃去上清液，用500μl冰冷的75％乙醇洗涤RNA团块。再次将试管在+4℃下于冷却的离心机中以最大速度离心5分钟，并弃去上清液。在室温下对RNA团块进行干燥12分钟，并随后通过在+4℃和1400rpm下振摇15分钟使其再悬浮于43μl不含核酸酶的冰冷水中。将RNA溶液等分为4x 10μl，并用液氮冷冻。分光光度法测定mRNA的浓度，并在使用RiboRuler™ RNA Ladder，高范围(Fermentas)的琼脂糖凝胶上进行复核—MLV_pD mRNA～1.2μg/μl；del_pD mRNA～1.2μg/μl。

[0149] 将经纯化mRNA以1:50＝MLV(活性RT)：del(失活RT)的比例混合。将MLV_pD mRNA稀释25倍至～48ng/μl，使1μl(～48ng)与2μl～1.2μg/μl的del_pD mRNA(2.4μg)混合，得到～0.8μg/μl比例为1:50的mRNA混合物。使用两种翻译系统RTS 100大肠杆菌HY试剂盒(03 186 148 001-Roche)和合成型WakoPURE(295-59503-Wako)进行体外翻译。给出蛋白翻译序列，对于MLV_pD为SEQ ID No:6，对于del_pD为SEQ ID No:7。

[0150] 用于RTS HY系统的翻译混合物(25μl)：6μl—大肠杆菌裂解物(Roche)；5μl—反应混合物(Roche)；6μl-氨基酸(Roche)；0.5μl—100mM蛋氨酸(Roche)；0.5μl—40u/μl RiboLock RNA酶抑制剂(Fermentas)；0.4μl—200μM assrA寡核苷酸(SEQ ID No:5)；0.25μl—1M DTT；2.5μl重建缓冲液(Roche)；2.5μl不含核酸酶的水和1.5μl—0.8μg/μl mRNA混合物1:50＝MLV_pD:del_pD(～1200ng)。在30℃下进行体外翻译20分钟。

[0151] 用于WakoPURE系统的翻译混合物(25μl)：12.5μl-A溶液(Wako)；5μl—B溶液(Wako)；0.5μl—40u/μl RiboLock RNA酶抑制剂(Fermentas)；0.4μl—200μMαssrA寡核苷酸(SEQ ID No:6)；0.25μl—1M DTT；5μl不含核酸酶的水和1.5μl—0.8μg/μl mRNA混合物1:50＝MLV_pD:del_pD(～1200ng)。在37℃下进行体外翻译30分钟。

[0152] 通过添加155μl冰冷的终止缓冲液$WBK_{500}$+DTT+triton(25℃下pH为7.5的50mM tris-乙酸盐(tris-acetate)；50mM NaCl；50mM乙酸镁；500mM KCl；10mM DTT；0.1％(v/v)-triton x-100(T8787-Sigma))并在+4℃和25'000g下离心5分钟来终止两个翻译体系(～25μl)。非常仔细地将160μl经离心的翻译混合物移液到840μl35％(w/v)蔗糖在$WBK_{500}$+DTT+triton中的溶液(50mM在25℃下pH为7.5的tris-乙酸盐；50mM NaCl；50mM乙酸镁；500mM KCl；10mM DTT；0.1％(v/v)-triton x-100(T8787-Sigma)；35％(w/v)-蔗糖(84097-

Fluka))的顶部。为了对mRNA-核糖体-蛋白(tRNA)的三元复合物(TC)进行纯化,使用TL-100Beckman超速离心机;TLA100.2固定角转头(Beckman);透明的1ml超速离心管(343778-Beckman)在+4℃和100'000rpm下进行超速离心9分钟。为了将小的透明TC团块保持在超速离心管的底部,小心地处理完整的试管。最初从离心管的最上部除去750μl溶液。然后(非常仔细地)用750μl $WBK_{500}$(50mM在25℃下pH为7.5的tris-乙酸盐;50mM NaCl;50mM乙酸镁;500mM KCl)洗涤管壁。最后从离心管的最上部开始处除去所有溶液,并且将团块溶解于30μl冰冷的终止缓冲液$WBK_{500}$+DTT+triton(50mM在25℃下pH为7.5的tris-乙酸盐;50mM NaCl;50mM乙酸镁;500mM KCl;10mM DTT;0.1%(v/v)-triton x-100(T8787-Sigma))中。

[0153] 如超速离心后使用放射性标记的mRNA所确定,5%至30%的输入的mRNA位于三元复合物团块中。因此预期在30μl缓冲液中具有低于360ng(翻译反应中所用的1200ng mRNA的30%)的mRNA(~12ng/μl或$9\times10^9$个分子/μl的三元复合物)。

[0154] 在冰上制备用于选择的逆转录反应混合物:60μl—5X用于逆转录酶的反应缓冲液(Fermentas);7.5μl-40u/μl RiboLock RNA酶抑制剂(Fermentas);15μl—20μM pD_42寡核苷酸(SEQ ID No:8);188μl不含核酸酶的水—将混合物分成2x 135μl的两试管,添加0.9μl经纯化的TC(<$8\times10^9$个分子)(在Roche-RTS HY试剂盒中翻译)或0.9μl经纯化的TC(<$8\times10^9$个分子)(在Wako-WakoPURE翻译)。将各反应混合物(~135μl)再次分成45μl和90μl的两试管。向第一部分—45μl的RT混合物中添加5μl不含核酸酶的水。认为该样品为阴性选择对照(无dNTP),和必须证明在反应混合物中无DNA污染,并且cDNA合成与来自三元复合物中的MLV RT的逆转录酶功能活性严格关联。向第二部分—90μl的RT混合物中添加10μl—10mM各dNTP混合物(Fermentas),并且将反应混合物再次分为用于选择对照的50μl和用于阳性选择对照的补充有1μl—200u/μl的RevertAid H-M-MuLV逆转录酶(Fermentas)的50μl两试管。根据所述方案,各逆转录反应混合物在50μl体积中含有<$2.7\times10^9$个分子的三元复合物。

[0155] 通过将ABIL EM 90(Goldschmidt)混入矿物油(M5904-Sigma)至终浓度4%(v/v)来制备乳化用油类表面活性剂混合物(Ghadessy和Holliger,2004;US2005064460)。通过将950μl油类表面活性剂与50μl RT混合物混合在+4℃下在5ml冷冻小瓶(cryogenic vial)(430492-Corning)中制备乳液。使用速度控制在~2100rpm的MS-3000磁力搅拌器、具有中心环的Rotilabo®-(3x8mm)磁性随动装置(1489.2-Roth)进行混合;每30秒钟向水相添加10μl等分试样,继续混合另外2分钟(总混合时间—4分钟)。根据光学显微镜数据,本发明乳液中的区室大小为0.5μm~10μm,平均直径为~2μm。因此预期对50μl逆转录反应混合物进行乳化后具有~$1\times10^{10}$个油包水区室,含有低于$2.7\times10^9$个分子的三元复合物(每3至4个区室约1个mRNA和逆转录酶分子)。

[0156] 对所有6种乳液在+42℃下温育一小时,所述6种乳液代表在Roche-RTS HY试剂盒中翻译的具有TC的RT混合物(阴性选择对照、选择对照和阳性选择对照)和在Wako-WakoPURE中翻译的具有TC的RT混合物(阴性选择对照、选择对照和阳性选择对照)。

[0157] 为了回收反应混合物,将乳液移至1.5ml的试管,在室温和25'000g下离心1分钟。除去油相,在试管底部留下浓缩(但仍完整)的乳液,并添加250μl PB缓冲液(Qiagen PCR纯化试剂盒)。最后通过用0.9ml水饱和的醚提取;0.9ml水饱和的乙酸乙酯(为了除去ABIL EM 90洗涤剂)提取;并再次用0.9ml水饱和的醚提取将乳液破乳。在室温下于真空中对水相干

燥5分钟。对所合成的cDNA用Qiagen PCR纯化试剂盒进行纯化，并在30μl EB缓冲液(Qiagen PCR纯化试剂盒)中洗脱。

[0158] 通过巢式PCR进行cDNA的扩增。在冰上制备起始的PCR混合物：16μl—10X具有KCl的Taq缓冲液(Fermentas)；16μl—2mM的各dNTP(Fermentas)；9.6μl—25mM $MgCl_2$(Fermentas)；3.2μl—1u/μl LC(重组)Taq DNA聚合酶(Fermentas)；0.8μl—100μM RD_Nde引物(SEQ ID No:9)；0.8μl—100μM pD_55引物(SEQ ID No:10)；74μl水—将混合物分成6个15μl的样品(6x15μl)和30μl的样品。向6x 15μl PCR主混合物中添加5μl cDNA(RT样品1至6)；向30μl PCR主混合物中添加9μl水，并再次将混合物分成2x 19.5μl的两试管以用于阴性PCR对照(加0.5μl水)和阳性PCR对照(加0.5μl—pET_his_MLV_pD和pET_his_del_pD质粒的1:1混合物(～1ng))。循环方案为：在94℃下进行起始变性步骤3分钟，25个循环(94℃下45秒钟，58℃下45秒钟和72℃下2分钟)和最后在72℃下延伸5分钟。对于MLV_pD，PCR片段的预期长度是2185bp，对于del_pD是2014bp。在每孔加载10μl PCR混合物的1％琼脂糖凝胶上对扩增分进行析。

[0159] 使用两组不同的引物进行巢式PCR，导致部分基因扩增(以更好的分辨RT样品中的MLV:del cDN比例)或全基因扩增(以证明全基因恢复的可能性)。

[0160] 在冰上制备用于部分基因扩增的巢式PCR混合物：28μl—10X具有KCl的Taq缓冲液(Fermentas)；28μl—2mM的各dNTP(Fermentas)；16.8μl—25mM $MgCl_2$(Fermentas)；5.6μl—1u/μl LC(重组)Taq DNA聚合酶(Fermentas)；1.4μl—100μM M_F引物(SEQ ID No:11)；1.4μl—100μM M_2R引物(SEQ ID No:12)；185μl水—将混合物分成2x 19μl和6x 38μl。向2x 19μl的PCR主混合物中添加1μl第一次PCR物的阳性对照或阴性对照(引物组RD_Nde//pD_55)—30个PCR循环扩增；向6x38μl的PCR主混合物中添加2μl第一次PCR(引物组RD_Nde//pD_55)(样品1至6)—将各样品再次分成2x 20μl的两部分以用于23或30个PCR循环扩增。循环方案为：在94℃下进行起始变性步骤3分钟，23或30个循环(94℃下45秒钟，57℃下45秒钟和72℃下1分钟)和最后在72℃下延伸5分钟。对于MLV_pD，PCR片段的预期长度是907bp，对于del_pD是736bp。在每孔加载10μl PCR混合物的1％琼脂糖凝胶上对扩增进行分析(图4)。

[0161] 在冰上制备用于全基因扩增的巢式PCR混合物：28μl—10X具有KCl的Taq缓冲液(Fermentas)；28μl—2mM各dNTP(Fermentas)；16.8μl—25mM $MgCl_2$(Fermentas)；5.6μl—1u/μl LC(重组)Taq DNA聚合酶(Fermentas)；1.4μl—100μM M_Esp引物(SEQ ID No:13)；1.4μl—100μM M_Eri引物(SEQ ID No:14)；185μl水—将混合物分成2x 19μl和6x 38μl。向2x 19μl PCR主混合物中添加1μl第一次PCR物的阳性或阴性对照(引物组RD_Nde//pD_55)—30个PCR循环扩增；向6x 38μl PCR主混合物中添加2μl首次PCR(引物组RD_Nde//pD_55)(样品1至6)—将各样品再次分成2x 20μl的两部分以用于23或30个PCR循环扩增。循环方案为：在94℃下进行起始变性步骤3分钟，23或30个循环(94℃下45秒钟，55℃下45秒钟和72℃下2分钟)和最后在72℃下延伸5分钟。对于MLV_pD，PCR片段的预期长度是2077bp，对于del_pD是1906bp。在每孔加载10μl PCR混合物的琼脂糖凝胶上对扩增进行分析(图5)。

[0162] 结果

[0163] 为了证明区室化核糖体展示(CRD)法的原理验证，使用与蛋白D间隔子融合的编码活性(MLV)和失活(del)逆转录酶的两种mRNA的1:50＝MLV:del起始混合物进行选择。使用

两种不同的翻译系统Roche-RTS 100大肠杆菌HY或Wako-WakoPURE进行体外翻译，以了解在本发明的实验方案中哪一种翻译系统更好。对于各翻译系统进行三种区室化RT反应：无dNTP的阴性选择对照，其需要证明在反应混合物中无DNA污染；选择对照，其需要证明编码活性(MLV)逆转录酶的基因相对于编码失活酶(del)的基因的富集，原因在于仅活性酶可以合成cDNA；以及补充有外部RevertAid H-商业逆转录酶的阳性选择对照，其必须充当在所有区室中合成来自MLV_pD和del_pD mRNA的cDNA的阳性RT对照，以表明不施加选择压力时反应混合物中基因的真实比例。

[0164] 通过巢式PCR对合成的cDNA进行扩增。起始PCR(25个循环)的琼脂糖凝胶电泳图像(图3)仅在两个阳性选择对照(翻译系统-Roche和Wako)的情况下显示了弱的PCR片段带。这是正常的，原因在于这些样品含有外部RT酶，该酶与每区室含有仅一个体外合成的逆转录酶分子的反应相比，合成cDNA的效率高得多。

[0165] 巢式PCR(部分基因扩增)琼脂糖凝胶电泳的图像示于图4。23和30个PCR循环后扩增结果一致：

[0166] 1)在阴性选择对照(w/o dNTP)中无扩增(无DNA污染)；

[0167] 2)在阳性选择对照(外部RT酶)中观察到del_pD cDNA(736bp DNA片段)的非常有效的扩增，并且未见到MLV_pD cDNA的扩增，原因在于MLV_pD与del_pD mRNA的起始比例为1:50；

[0168] 3)在选择对照的情况下观察到cDNA MLV_pD(907bp DNA片段)和del_pD(736bp DNA片段)的扩增；

[0169] 4)在由Roche体外翻译系统所合成的逆转录酶的情况下观察到～1:1的MLV_pD:del_pD的比例，这意味着MLV_pD基因相对于del_pD基因从起始的1:50比例开始富集了～50倍；

[0170] 5)在由Wako体外翻译系统所合成的逆转录酶的情况下观察到～1:3的MLV_pD:del_pD的比例，这意味着MLV_pD基因相对于del_pD基因从起始的1:50比例开始富集了～16倍；

[0171] 巢式PCR(全基因扩增)琼脂糖凝胶电泳的图像示于图5。在23和30个PCR循环的结果之间，以及与用于部分基因扩增的巢式PCR(图5)的结果相比，23和30个PCR循环后的扩增结果一致：

[0172] 1)在阴性选择对照(w/o dNTP)中无扩增(无DNA污染)；

[0173] 2)在阳性选择对照(外部RT酶)中观察到del_pD cDNA(1906bp DNA片段)的非常有效的扩增，并且未见到MLV_pD cDNA的扩增，原因在于MLV_pD与del_pD mRNA的起始比例为1:50；

[0174] 3)在选择对照的情况下观察到cDNA MLV_pD(2077bp DNA片段)和del_pD(1906bp DNA片段)的扩增；

[0175] 4)在全基因扩增的情况下难以确定MLV_pD:del_pD的比例，原因在于2077bp(MLV_pD)和1906bp(del_pD)DNA片段之间的相对差不够大，但是通常比例与用于部分基因扩增的巢式PCR的结果相似。

[0176] 作为该实施例的结果，可以总结出在以CRD形式进行的逆转录反应期间，相对于编码失活酶的基因富集了编码活性MLV逆转录酶的基因，在Roche翻译系统的情况下富集倍数

为50，在用于体外合成酶的Wako翻译系统的情况下富集倍数为16。

[0177] 实施例2-CRD-选择在更高温度下性能改善的逆转录酶

[0178] 为了了解区室化核糖体展示(CRD)选择如何有效地起作用，进行莫洛尼氏鼠白血病病毒逆转录酶(M-MuLV RT)的进化实验。实验的一般流程图示于图6。通过使用核苷酸类似物dPTP和8-氧-dGTP的易错PCR构建逆转录酶的起始突变体文库。进行全基因(～2kb)诱变，每个基因引入2至3个核苷酸或1至2个氨基酸突变。使用编码逆转录酶(与蛋白D融合)MLV_pD的突变体文库的PCR片段来合成mRNA。经纯化的mRNA用于体外翻译反应。在翻译混合物中形成mRNA-核糖体-MLV_pD(tRNA)的三元复合物(TC)，并通过低温和高浓度的$Mg^{2+}$离子将其稳定。通过蔗糖垫超速离心对TC的混合物进行纯化。将沉淀的TC溶解于冰冷的缓冲液(50mM $Mg^{2+}$)中并用于制备逆转录反应混合物，该逆转录反应混合物补充有用于RT反应的外部dNTP组和引物。对冰冷的RT反应混合物进行乳化，产生～$1\times10^{10}$个大小为～2μm的油包水区室。MLV RT的最适反应温度是～42℃。为了选择在更高温度下工作更好的逆转录酶变体，使乳化RT反应混合物(低于一个TC(mRNA+MLV RT)/区室)于50℃下温育1小时。在该温度下在含有更具活性或热稳定性的MLV逆转录酶变体的区室中，全长cDNA的成功合成进行地更好。随后的PCR用于扩增全长cDNA，并进行更具活性和热稳定性的逆转录酶基因的富集。通过PCR，将扩增的基因回复到CRD形式，通过连接-PCR恢复完整的5'(起始片段-T7聚合酶启动子、SD和his-标签编码序列)和3'(末端片段-gs连接子、蛋白D和第二gs连接子)序列。

[0179] 含有逆转录酶基因的富集文库的重建PCR片段用于随后的mRNA转录和下一轮CRD选择。以温度越来越高的RT反应进行各轮选择：50℃(第一轮)；52.5℃(第二轮)；55℃(第三轮)；57.5℃(第四轮)和60℃(第五轮)。

[0180] 将第五轮选择后逆转录酶基因的扩增文库(无C末端pD连接子)克隆到质粒载体。对个体克隆进行测序和分析。使用亲和色谱通过his-标签对进化蛋白以及个体突变体的库进行纯化。确定37℃、50℃下MLV逆转录酶的比活性和在50℃下酶温育5分钟后在37℃下的残留活性。

[0181] 方法和材料

[0182] 使用起始质粒pET_his_MLV_pD(SEQ ID No:1和图2)作为易错PCR的起始材料。使用核苷酸类似物dPTP和8-氧-dGTP引入突变。在冰上制备用于易错PCR的PCR混合物：10μl—10X含有KCl的Taq缓冲液(Fermentas)；10μl—2mM各dNTP(Fermentas)；6μl—25mM $MgCl_2$(Fermentas)；2μl—1u/μl LC(重组)Taq DNA聚合酶(Fermentas)；0.5μl—100μM M_Esp引物(SEQ ID No:13)；0.5μl—100μM M_Eri引物(SEQ ID No:14)；1μl—10μM dPTP(TriLink BioTechnolgies)；5μl—100μM8-氧-dGTP(TriLink BioTechnolgies)；3.75μl—40ng/μl(总共150ng)的pET_his_MLV_pD质粒；61.25μl水。循环方案为：在94℃下进行起始变性步骤3分钟，30个循环(94℃下30秒钟，55℃下30秒钟和72℃下2分钟)和最后在72℃下延伸5分钟。从150ng质粒(7873bp)到～6至12μg的扩增产物(对于pET_his_MLV_pD为2077bp PCR片段)，扩增了150至300倍。使用Qiagen PCR纯化试剂盒对PCR片段进行纯化，用Esp3I(识别序列CGTCTC(1/5))和EcoRI(识别序列G↓AATTC)进行消化，最后使用Qiagen凝胶提取试剂盒从琼脂糖凝胶中进行纯化，得到的DNA浓度为～50ng/μl。

[0183] 通过对亚克隆回用NcoI和EcoRI消化的起始pET_his_MLV_pD质粒中的个体克隆进行测序，从而检查突变发生效率和文库质量。据预期突变随机分布在整个MLV RT基因的扩

增序列中(附录1)。在10个测序基因中,发现23个核苷酸突变(1个颠换、20个转换、2个缺失—在附录1中用红色标记),产生15个氨基酸交换、6个沉默突变、1个终止密码子和2个编码框的移码—平均每个基因1至2个氨基酸取代。

[0184] 将突变文库与起始片段(244bp)和末端片段(398bp)连接,从而获得适合CRD选择的PCR片段(图7)。通过对最初的983bp起始片段(靶-质粒pET_his_del_pD(SEQ ID No:2)、引物—pro-pIVEX(SEQ ID No:3)和M_1R(SEQ ID No:15))进行的PCR扩增和随后用NcoI(识别序列C↓CATGG)进行的消化,构建起始片段(含有T7聚合酶启动子、SD和his-标签编码序列),获得244bp的DNA片段。通过对最初的1039bp末端片段(靶-质粒pET_his_del_pD(SEQ ID No:2)、引物—M_3F(SEQ ID No:16)和pD-ter(SEQ ID No:4))的进行PCR扩增和随后用EcoRI(识别序列G↓AATTC)进行的消化,构建末端片段(含有gs连接子、蛋白D和第二gs连接子序列),获得398bp的DNA片段。

[0185] 在室温下制备连接反应物(150μl):15μl—10X用于T4DNA连接酶的连接缓冲液(Fermentas);15μl—1u/μl T4DNA连接酶(Fermentas);26μl—50ng/μl用Esp3I(NcoI相容末端)和EcoRI消化的突变MLV RT文库(~1300ng或~$5.9\times10^{11}$个分子);9.4μl—35ng/μl用NcoI消化的起始片段(~329ng或~$1.2\times10^{12}$个分子);15.7μl—35ng/μl用EcoRI消化的末端片段(~548ng或~$1.2\times10^{12}$个分子);68.9μl-水。在+4℃下进行连接过夜。将反应混合物用苯酚处理1次,并用氯仿处理2次,使其沉淀并溶解于53μl的水。比较连接混合物和已知量的质粒pET_his_MLV_pD的扩增效率,确定连接产率为约~20%。考虑到20%的连接产率,将MLV RT突变体文库的多样性定为~$1.2\times10^{11}$个分子(50μl)。

[0186] 通过PCR扩增连接的MLV RT文库(1ml-在冰上制备):100μl—10X含有KCl的Taq缓冲液(Fermentas);100μl—2mM各dNTP(Fermentas);60μl—25mM $MgCl_2$(Fermentas);80μl—DMSO(D8418-Sigma);20μl—1u/μl LC(重组)Taq DNA聚合酶(Fermentas);5μl—100μM pro-pIVEX引物(SEQ ID No:3);5μl—100μM pD-ter-引物(SEQ ID No:17);20μl-连接的MLV RT文库(~$5\times10^{10}$个分子);610μl-水。循环方案为:在94℃下进行起始变性步骤3分钟,15个循环(94℃下30秒钟,53℃下30秒钟和72℃下3分钟)和最后在72℃下延伸5分钟。从~$5\times10^{10}$个分子(相当于~150ng)的大小为2702bp的最终连接片段至~30μg(30ng/μl)的扩增产物(2702bp PCR片段),扩增了~200倍。

[0187] 第1轮选择

[0188] 制备转录混合物(100μl):20μl—5x T7转录缓冲液(pH为7.6的1M HEPES-KOH;150mM乙酸镁;10mM亚精胺;0.2M DTT);28μl—25mM各NTP(Fermentas);4μl—20u/μl T7RNA聚合酶(Fermentas);2μl—40u/μl RiboLock RNA酶抑制剂(Fermentas);30μl—30ng/μl突变体文库(pro-pIVEX//pD-ter-)PCR混合物(~900ng或~$3\times10^{11}$个分子);16μl不含核酸酶的水。在37℃下转录进行3小时(文库多样性~$5\times10^{10}$个分子)。

[0189] 用冰冷的不含核酸酶的水将转录混合物稀释到200μl,并添加200μl 6M LiCl溶液。将混合物在+4℃下温育30分钟,在+4℃下于冷却的离心机中以最大速度(25'000g)离心30分钟。弃去上清液,用500μl冰冷的75%乙醇洗涤RNA团块。再次将试管在+4℃下于冷却的离心机中以最大速度离心5分钟,并弃去上清液。在室温下对RNA团块进行干燥12分钟,并随后通过在+4℃和1400rpm下振摇15分钟使其再悬浮于200μl不含核酸酶的冰冷的水中。再次将管在+4℃下于冷却的离心机中以最大速度离心5分钟,从而分离未溶解的RNA。将约180μl

的上清液移到具有20μl 10X DNA酶I缓冲液($Mg^{2+}$) (Fermentas)；1μl—1u/μl的DNA酶I(不含RNA酶) (Fermentas)的新试管中，并且在37℃下温育30分钟，从而降解DNA。向反应混合物添加20μl pH为5.0的3M乙酸钠溶液和500μl冰冷的96％乙醇。最后通过在-20℃下温育30分钟，并在+4℃下于冷却的的离心机中以最大速度(25'000g)离心30分钟使RNA沉淀。弃去上清液，用500μl冰冷的75％乙醇洗涤RNA团块。再次将管在+4℃下于冷却的离心机中以最大速度离心5分钟，并弃去上清液。在室温下对RNA团块进行干燥12分钟，并随后通过在+4℃和1400rpm下振摇10分钟再悬浮于33μl不含核酸酶的冰冷的水中。将RNA溶液等分为3x 10μl，并用液氮冷冻。分光光度法测定mRNA的浓度，并在使用RiboRuler™ RNA Ladder，高范围(Fermentas)的琼脂糖凝胶上进行复核-MLV RT文库mRNA-2.1μg/μl。

[0190] 使用RTS 100大肠杆菌HY(03 186 148 001-Roche)翻译系统进行体外翻译(25μl)：6μl-大肠杆菌裂解物(Roche)；5μl-反应混合物(Roche)；6μl-氨基酸(Roche)；0.5μl—100mM蛋氨酸(Roche)；0.5μl—40u/μl RiboLock RNA酶抑制剂(Fermentas)；0.4μl—200uM assrA寡核苷酸(SEQ ID No:5)；0.25μl—1M DTT；3μl重组缓冲液(Roche)；2.5μl不含核酸酶的水和0.6μl—2.1μg/μl mRNA(-1200ng)。在30℃下使反应混合物温育20分钟。通过添加155μl冰冷的终止缓冲液$WBK_{500}$+DTT+triton(25℃下pH为7.5的50mM tris-乙酸盐；50mM NaCl；50mM乙酸镁；500mM KCl；10mM DTT；0.1％(v/v)-triton x-100(T8787-Sigma))来终止翻译，并在+4℃和25'000g下离心5分钟来终止翻译。非常仔细地将160μl经离心的翻译混合物移液到840μl 35％(w/v)蔗糖在WBKsoo+DTT+triton中的35％(w/v)蔗糖溶液(50mM在25℃下pH为7.5的tris-乙酸盐；50mM NaCl；50mM乙酸镁；500mM KCl；10mM DTT；0.1％(v/v)-triton x-100(T8787-Sigma)；35％(w/v)-蔗糖(84097-Fluka))的顶部。为了对mRNA-核糖体-MLV(tRNA)的三元复合物(TC)进行纯化，使用TL-100Beckman超速离心机、TLA100.2固定角转头(Beckman)、透明的1ml超速离心管(343778-Beckman)在+4℃和100,000rpm下进行超速离心9分钟。为了将小的透明TC团块保持在超速离心管的底部，小心地处理完整的试管。最初从离心管的最上部除去750μl溶液。然后(非常仔细地)用750μl $WBK_{500}$(50mM在25℃下pH为7.5的tris-乙酸盐；50mM NaCl；50mM乙酸镁；500mM KCl)洗涤管壁。最后从离心管的最上部开始除去所有溶液，并且将团块溶解于30μl冰冷的终止缓冲液$WBK_{500}$+DTT+triton(50mM在25℃下pH为7.5的tris-乙酸盐；50mM NaCl；50mM乙酸镁；500mM KCl；10mM DTT；0.1％(v/v)-triton x-100(T8787-Sigma))中。

[0191] 使用放射性标记的mRNA通过实验确定：超速离心在三元复合物团块中得到输入mRNA的5％至30％。因此据预测在30μl缓冲液中具有低于360ng(翻译反应中所用的1200ng mRNA的30％)的mRNA(～12ng/μl或$9\times10^9$个分子/μl的三元复合物)。

[0192] 在冰上制备用于选择的逆转录反应混合物：60μl—5X用于逆转录酶的反应缓冲液(Fermentas)；7.5μl—40u/μl RiboLock RNA酶抑制剂(Fermentas)；15μl—20μM pD_42寡核苷酸(SEQ ID No:8)；186μl不含核酸酶的水和1.8μl经纯化(<$1.8\times10^{10}$个分子)TC(在Roche-RTS HY试剂盒中翻译)。将反应混合物分成45μl和225μl的两试管。向第一部分—45μl的RT混合物中添加5μl的不含核酸酶的水。该样品作为阴性选择对照(无dNTP)，并且需要证明在反应混合物中无DNA污染，并且cDNA合成与来自三元复合物中的MLV RT的逆转录酶功能活性严格关联。向第二部分—225μl的RT混合物中添加25μl—10mM各dNTP混合物(Fermentas)，并且将反应混合物再次分为两试管：用于选择对照的200μl(4x 50μl)(总共<

1.2×$10^{10}$个分子的TC)和用于阳性选择对照的补充有1μl—200u/μl的RevertAid H-M-MuLV逆转录酶(Fermentas)的50μl。根据方案,各逆转录反应混合物在50μl体积中含有<3×$10^{9}$个分子的三元复合物。

[0193] 通过将ABIL EM 90(Goldschmidt)混入矿物油(M5904-Sigma)至终浓度4%(v/v)来制备乳化用油-表面活性剂混合物(Ghadessy和Holliger,2004;US2005064460)。通过将950μl油-表面活性剂混合物与50μl RT混合物混合,从而在+4℃下在5ml冷冻小瓶(cryogenic vial)(430492-Corning)中制备乳液。使用速度控制在~2100rpm的MS-3000磁力搅拌器、具有中心环的Rotilabo®-(3x8mm)磁性随动装置(1489.2-Roth)进行混合;每30秒钟向水相添加10μl等分试样,继续混合另外2分钟(总混合时间-4分钟)。根据光学显微镜数据,本发明乳液中的区室大小为0.5μm~10μm,平均直径为~2μm。因此据预期对50μl逆转录反应混合物进行乳化后具有~1×$10^{10}$个油包水区室,含有低于3×$10^{9}$个分子的三元复合物(每3至4个区室约1个mRNA和逆转录酶分子)。

[0194] 使所有乳液在+50℃下温育1小时,以选择在更高温度下工作更好的逆转录酶变体。为了回收反应混合物,将乳液移至1.5ml的试管,在室温和25'000g下离心10分钟。除去油相,在试管底部留下浓缩(但仍完整)的乳液。通过用0.9ml水饱和的醚提取;0.9ml水饱和的乙酸乙酯(为了除去ABIL EM 90洗涤剂)提取;并再次用0.9ml水饱和的醚提取将乳液破乳。在室温下于真空中对水相进行干燥5分钟,并且添加250μl的PB缓冲液(Qiagen PCR纯化试剂盒)。将四个选择样品合并到两个试管中。用Qiagen PCR纯化试剂盒将所合成的cDNA进一步纯化,并在阴性和阳性选择对照的情况下在30μl EB缓冲液(Qiagen PCR纯化试剂盒)中洗脱,在选择对照的情况下用2x 30μl的EB缓冲液洗脱。

[0195] 通过巢式PCR进行cDNA的扩增。首先进行小PCR扩增,以检查阴性和阳性选择对照,并确定对于选择样品中的cDNA进行有效扩增所需的PCR循环的最小数量。在冰上制备PCR混合物(200μl):20μl—10X含有KCl的Taq缓冲液(Fermentas);20μl—2mM各dNTP(Fermentas);12μl—25mM $MgCl_2$(Fermentas);2.5μl—1u/μl LC(重组)Taq DNA聚合酶(Fermentas);1μl—2.5u/μl Pfu DNA聚合酶(Fermentas);1μl—100μM RD_Nde引物(SEQ ID No:9);1μl—100μM pD_55引物(SEQ ID No:10);50μl-选择对照的纯化cDNA;92μl水。用于2185bp PCR片段的循环方案为:在94℃下进行起始变性步骤3分钟,25个循环(94℃下45秒钟,58℃下45秒钟和72℃下2分钟)和最后在72℃下延伸5分钟。

[0196] 在冰上制备用于全基因扩增的巢式PCR混合物(500μl):50μl—10X含有KCl的Taq缓冲液(Fermentas);50μl—2mM各dNTP(Fermentas);30μl—25mM $MgCl_2$(Fermentas);6.25μl—1u/μl LC(重组)Taq DNA聚合酶(Fermentas);2.5μl—2.5u/μl Pfu DNA聚合酶(Fermentas);2.5μl—100μM M_Esp引物(SEQ ID No:13);2.5μl—100μM M_Eri引物(SEQ ID No:14);50μl第一次PCR物(引物组RD_Nde//pD_55);306μl水。用于2077bp PCR片段的循环方案为:在94℃下进行起始变性步骤3分钟,22个循环(94℃下45秒钟,55℃下45秒钟和72℃下2分钟)和最后在72℃下延伸5分钟。

[0197] 使用Qiagen凝胶提取试剂盒对选择样品的最终PCR片段进行琼脂糖凝胶纯化(在60μl中洗脱~50ng/μl)。在37℃下用EcoRI和Esp3I对经纯化的PCR片段消化1小时,并再次进行琼脂糖-凝胶纯化(在30μl中洗脱~50ng/μl)。

[0198] 将第一轮选择后回收的MLV逆转录酶文库与起始片段和末端片段(在本实施例中

之前已描述构建)连接,从而获得适合第二轮CRD选择(图6)的PCR片段(图7)。在室温下制备连接反应物(40μl):4μl—10X用于T4DNA连接酶的连接缓冲液(Fermentas);2μl—1u/μl T4DNA连接酶(Fermentas);4μl—50ng/μl用Esp3I和EcoRI消化的选择文库(~200ng或$0.9\times10^{11}$个分子);1.1μl—35ng/μl用NcoI消化的起始片段(~35ng或~$1.5\times10^{11}$分子);1.76μl—35ng/μl用EcoRI消化的末端片段(~61ng或~$1.5\times10^{11}$个分子);27.2μl-水。在室温下进行连接1小时。

[0199] 通过PCR对连接的MLV RT文库进行扩增(300μl-在冰上制备):30μl—10X含有KCl的Taq缓冲液(Fermentas);30μl—2mM各dNTP(Fermentas);18μl—25mM $MgCl_2$(Fermentas);24μl—DMSO(D8418-Sigma);3.7μl—1u/μl LC(重组)Taq DNA聚合酶(Fermentas);1.5μl—2.5u/μl Pfu DNA聚合酶(Fermentas);1.5μl—100μM pro-pIVEX引物(SEQ ID No:3);1.5μl—100μM pD-ter-引物(SEQ ID No:17);25.5μl-连接的MLV RT文库(~$.6\times10^{10}$个分子);164.3μl-水。用于2702bp PCR片段的循环方案为:在94℃下进行起始变性步骤3分钟,15个循环(94℃下45秒钟,53℃下45秒钟和72℃下3分钟)和最后在72℃下延伸5分钟。使用Qiagen凝胶提取试剂盒对PCR片段进行琼脂糖凝胶纯化(在30μl中洗脱~100ng/μl)。

[0200] 第2轮选择

[0201] 第2轮选择按照第1轮选择实验流程的一般设置进行,不同之处在于PCR循环、乳化RT反应温度和几个其它细节上的微小改动。

[0202] 所有改动如下给出:

[0203] • 转录—取10μl—100ng/μl(~1000ng)的琼脂糖凝胶纯化的PCR片段,第2轮选择中所用的mRNA的终浓度为1.5μg/μl;

[0204] • 翻译—取0.8μl—1.5μg/μl(~1.2μg)的mRNA;

[0205] • 乳化RT反应在52.5℃进行1小时;

[0206] • 第一次PCR(RD_Nde//pD_55)—进行24个循环;

[0207] • 第二次(巢式)PCR(M_Esp//M_Eri)—进行23个循环;

[0208] • 经消化PCR片段的终浓度—80ng/μl;

[0209] • 连接—取200ng(~$0.9\times10^{11}$个分子)的MLV RT文库;

[0210] • PCR(对连接混合物进行)—取<$0.6\times10^{10}$个分子的经选择文库,并进行15个PCR循环;最终琼脂糖凝胶纯化的PCR片段的浓度为200ng/μl。

[0211] 第3轮选择

[0212] 第3轮选择按照第1轮选择实验流程的一般设置进行,不同之处在于PCR循环、乳化RT反应温度和几个其它细节上的微小改动。

[0213] 所有改动如下给出:

[0214] • 转录—取5μl—200ng/μl(~1000ng)的琼脂糖凝胶纯化的PCR片段,第3轮选择中所用的mRNA的终浓度为1.5μg/μl;

[0215] • 翻译—取0.8μl—1.5μg/μl(~1.2μg)的mRNA;

[0216] • 乳化RT反应在55℃下进行1小时;

[0217] • 第一次PCR(RD_Nde//pD_55)—进行25个循环;

[0218] • 第二次(巢式)PCR(M_Esp//M_Eri)—进行22个循环;

[0219] • 经消化PCR片段的终浓度—70ng/μl；

[0220] • 连接—取200ng(～$0.9\times10^{11}$个分子)的MLV RT文库；

[0221] • PCR(对连接混合物进行)—取<$0.6\times10^{10}$个分子的所选择文库，并进行15个PCR循环；最终琼脂糖凝胶纯化的PCR片段的浓度为100ng/μl。

[0222] 第4轮选择

[0223] 第4轮选择按照第1轮选择实验流程的一般设置进行，不同之处在于PCR循环、乳化RT反应温度和几个其它细节上的微小改动。

[0224] 所有改动如下给出：

[0225] • 转录—取10μl—100ng/μl(～1000ng)的琼脂糖凝胶纯化的PCR片段，第4轮选择中所用的mRNA的终浓度为1.8μg/μl；

[0226] • 翻译—取0.67μl—1.8μg/μl(～1.2μg)的mRNA；

[0227] • 乳化RT反应在57.5℃下进行1小时；

[0228] • 第一次PCR(RD_Nde//pD_55)—进行25个循环；

[0229] • 第二次(巢式)PCR(M_Esp//M_Eri)—进行24个循环；

[0230] • 经消化PCR片段的终浓度—50ng/μl；

[0231] • 连接—取200ng(～$0.9\times10^{11}$个分子)的MLV RT文库；

[0232] • PCR(对连接混合物进行)—取<0.6×1010个分子的所选择文库，并进行15个PCR循环；最终琼脂糖凝胶纯化的PCR片段的浓度为100ng/μl。

[0233] 第5轮选择

[0234] 第5轮选择按照第1轮选择实验流程的一般设置进行，不同之处在于PCR循环、乳化RT反应温度和分析最终阶段上的一些改动。

[0235] 所有改动如下给出：

[0236] • 转录—取10μl—100ng/μl(～1000ng)的琼脂糖凝胶纯化的PCR片段，第5轮选择中所用的mRNA的终浓度为1.1μg/μl；

[0237] • 翻译—取1.1μl—1.1μg/μl(～1.2μg)的mRNA；

[0238] • 乳化RT反应在60℃下进行1小时；

[0239] • 第一次PCR(RD_Nde//pD_55)—进行25个循环；

[0240] • 第二次(巢式)PCR(M_Esp//M_Eri)—进行33个循环；

[0241] 在冰上制备用于全基因扩增的巢式PCR混合物(500μl)：50μl—10X含有KCl的Taq缓冲液(Fermentas)；50μl-2mM各dNTP(Fermentas)；30μl-25mM $MgCl_2$(Fermentas)；6.25μl-1u/μl LC(重组)Taq DNA聚合酶(Fermentas)；2.5μl-2.5u/μl Pfu DNA聚合酶(Fermentas)；2.5μl-100μM M_Esp引物(SEQ ID No:13)；2.5μl-100μM M_Hind3+引物(SEQ ID No:18)；50μl的第一次PCR物(引物组RD_Nde//pD_55)；306μl水。用于2077bp PCR片段的循环方案为：在94℃下进行起始变性步骤3分钟，22个循环(94℃下45秒钟，55℃下45秒钟和72℃下3分钟)和最后在72℃下延伸5分钟。

[0242] 使用Qiagen凝胶提取试剂盒对选择样品的最终PCR片段进行琼脂糖凝胶纯化(在60μl中洗脱～60ng/μl)。经纯化的PCR片段在37℃下用HindIII和Esp3I消化1小时，并再次进行琼脂糖凝胶纯化(在40μl中洗脱～50ng/μl)。

[0243] 将第5轮选择后将回收的MLV逆转录酶文库连接到从用NcoI和HindIII消化的pET_

his_MLV_pD(SEQ ID No:1和图2)制备的质粒载体中,得到编码MLV RT的新的7474bp质粒pET_his_MLV(SEQ ID No:19),该MLV RT具有N末端his-标签,并且在C末端上没有pD融合以用于使用亲和色谱的快速蛋白纯化。

[0244] 将第5轮选择后的连接的MLV RT文库电穿孔到T7表达株ER2566中。对个体克隆进行测序和分析。使在相同构建中进化蛋白的库以及个体突变体和初始wt M-MuLV逆转录酶在200ml LB中生长至A590为～0.7,并使用2ml Qiagen-Ni-NTA Superflow树脂利用his标签进行纯化(根据供应商建议在天然条件下进行所有纯化)。在1ml EB(50mM—$NaH_2PO_4$,300mM—NaCl,pH为8.0的250mM—咪唑,10mM—β-巯基乙醇和0.1%triton X-100)中进行洗脱。以50倍过量的贮存缓冲液(50mM Tris-HCl(在25℃下pH 8.3),0.1M NaCl,1mM EDTA,5mM DTT,0.1%(v/v)Triton X-100和50%(v/v)甘油)对所有蛋白进行透析。在SDS-PAGE上检查蛋白纯度(通常～40%至80%的靶蛋白)。使用基于Bredford法的Bio-Rad蛋白检测(500-0006)确定蛋白浓度。

[0245] 在37℃、50℃下测定MLV逆转录酶的比活性(将酶稀释于特定的稀释缓冲液:30mM 25℃下pH为8.3的Tris-HCl,10mM DTT,0.5mg/ml BSA),以及在50℃下使酶温育5分钟后在37℃下测定残留活性(将酶稀释,并在1x RT反应缓冲液:50mM25℃下pH为8.3的Tris-HCl,4mM $MgCl_2$,10mM DTT,50mM KCl中测定其稳定性)。在以下最终混合物中检测所有情况下的酶活性:50mM Tris-HCl(25℃下pH为8.3),6mM $MgCl_2$,10mM DTT,40mM KCl,0.5mM dTTP,0.4MBq/ml[3H]-dTTP,0.4mM polyA·寡$(dT)_{12-18}$。确定活性单位,测定在特定反应温度下在10分钟内dTMP在多核苷酸级分(在DE-81上吸收)中的导入,并且与已知量的商业酶比较。

[0246] 结果

[0247] 在选择数据进行的分析的期间,我们已经累积了104个表达全长M-MuLV逆转录酶的序列。为了与文献中通常所用的具有相同的氨基酸编号相同,使将对所有蛋白(附录2)进行的全部总CLUSTALW比对以编译成无不带N末端His标签的方式汇编记。表示为MLV的野生型序列总是作为第一序列给出。使用于黑色背景中的白色字体标记突变(附录2)。其处发生的突变某种程度上改善M-MuLV逆转录酶性质并和描述于不同的专利申请中的突变处的氨基酸位置,在比对中标记为以灰色突出显示的氨基酸(白色字体)的柱。源自我们的选择并且位于灰色柱中的突变作为证明表明,我们的选择方法精确地靶向有益的热点或甚至别处所述的准确确切氨基酸突变。从104个测序克隆中我们发现了98个独特序列和1个wt(L5_87)序列。5个序列重复两次(L5_21和L5_111;L5_43和L5_112;L5_49和L5_63;L5_64和L5_93;L5_85和L5_96)。我们总共随机表达了55个蛋白。从在55个所表达蛋白中,根据SDS-PAGE,将40个M-MuLV逆转录酶的40个酶促活性突变变体(包括对照wt)利用SDS-PAGE成功纯化到40%至80%均匀性均一性(homogeneity)。纯化RT样品中的总蛋白浓度为0.6mg/ml-至5.5mg/ml。在37℃,50℃下测试突变RT变体在37℃、50℃下的逆转录酶活性和在50℃下温育5分钟后在37℃下的残留活性(图8)。将37℃下的逆转录酶活性标准化成100%并且在图8中省略。因此仅显示两种类型的柱(50℃下RT活性的百分比和50℃下温育5分钟后在37℃下的残留RT活性的百分比)。作为对照,显示了提出用于突变体文库构建的wt M-MuLV逆转录酶。该起始酶在与RT突变变体相同的载体中表达并以相同的方式进行纯化。50℃下wt酶RT活性的平均值为37℃下活性的约45%。除了少量例外,几乎所有测试蛋白在50℃下具有高于45%的活性。对于所有测试突变体50℃下的RT活性的平均值约为～92%,超过wt酶(45%)

的2倍。一些变体在50℃下的活性是37℃下活性的100％或甚至更高：20、23、L5_16、L5_24、L5_30、L5_35、L5_37、L5_43、L5_46、L5_47、L5_49、L5_52、L5_55、L5_64、L5_65、L5_68、L5_72。发现的最佳突变体在50℃下具有的RT活性是约140％以上(比wt的45％高3倍)：20(165％)、L5_37(162％)、L5_43(156％)、L5_46(135％)、L5_47(179％)、L5_52(137％)、L5_64(142％)和L5_68(153％)。

[0248] 即使大多数的突变体在50℃下具有非常高的RT活性，它们仍然不是热稳定的。Wt对照的50℃下温育5分钟后在37℃的残留RT活性为～11％。所选择酶的相同平均残留活性为类似的～12％。虽然一些所测试的RT变体实质上热稳定性更高，并且具有的残留活性比wt酶(11％)高2至3倍：L5_8(25％)、L5_43(32％)、L5_46(27％)、L5_64(28％)、L5_65(25％)、L5_68(31％)。

[0249] 部分经纯化wt酶的比活性(u/mg蛋白)是～200’000u/mg(图9)。以各种方式表达和纯化所选择的RT变体，并且平均比活性(～155’000u/mg)稍微低于wt对照(图9)。比活性在某些情况下减少，在某些情况下增加(20—～274’000u/mg；L5_11—～273’000u/mg；L5_28—～230’000u/mg；L5_30—～224’000u/mg；L5_35—～316’000u/mg；L5_43—～328’000u/mg；L5_46—～304’000u/mg；L5_52—～310’000u/mg；L5_64—～256’000u/mg；L5_65—～247’000u/mg)。

[0250] 显然我们的选择系统运行良好。使用RT反应的温度升高作为选择压力因子，我们成功地进化了更快(50℃下的RT比活性更高)和热稳定性(50℃下预温育5分钟后在37℃的残留RT活性)更高的逆转录酶。

[0251] 有价值信息的来源是所选择蛋白序列的比对(附录2)。与初始wt M-MuLV相比50℃下的活性实质上更好(与wt活性的45％相比，为70％以上)的所分析蛋白的序列用灰色突出显示(附录2)。发生突变的氨基酸数量为0(wt或L5_87)～12(L5_9)。在所有所选择RT变体中发现的突变的列表如附录3所示。蛋白按突变数量递减排序。大多数突变体(104个中的53个)具有4至6个突变/序列。显然逆转录酶序列具有一些热点，这些热点通常对于RT反应和对于酶的热稳定性是重要和有益的。作为特定位置处的突变集合，这些热点可以在多序列比对(附录2)中容易鉴定。特别重要的是M-MuLV逆转录酶的运行更好的变体中发现的突变(序列以灰色突出显示—附录2)。关于频率最高的突变的总结信息(以降序)示于附录4。如之前附录一样，在50℃下具有实质上更高活性的突变蛋白用灰色突出显示。如果相同突变重复许多次并且具有该突变的测试逆转录酶在50℃下运行更好，这意味着该突变以某种方式对逆转录反应有益。

[0252] 根据突变的发现频率，可以将它们分成5类：21至31个重复；14至18个重复；4至7个重复；2至3个重复和1个重复。第一组频率最高的突变包括四个氨基酸D524(31个重复)；D200(30个重复)；D653(23个重复)和D583(21个重复)。已知两个氨基酸(D524和D583)在核糖核酸酶H结构域的活性中心络合镁离子。使用突变D524G、D583N和E562Q来关闭M-MuLV逆转录酶的RNA酶H活性(Gerard等，2002)，这改善cDNA的合成。我们的选择结果惊人地相似。在98个序列中的31个内发现天冬氨酸524的突变。另外，D524N取代发现1次，D524A发现10次以及最后D524G发现20次。因此我们的选择不仅精确靶向重要的氨基酸，还靶向已知为最佳的相同的氨基酸取代。完全相同的情况是天冬氨酸583的突变，其在104个所选择蛋白中的21个中重复。取代D583E发现1次，D583A发现3次，D583G发现7次以及最后D583N发现10次。再

次，选择频率最高的是已知为最佳的相同氨基酸和相同取代(D583N)。来自Invitrogen的商业酶SUPERSCRIPT II具有3个突变：D524G，D583N和E562Q(WO2004024749)。令人感兴趣的事情是第3个氨基酸取代E562的突变在我们选择中仅发现一次(在L5_71中的E562K)。该结果表明，很有可能谷氨酸562不如天冬氨酸524和583那样重要，或出于某些原因该氨基酸的交换可能引起某些副作用，并且对于在较高温度(>50℃)下进行的RT反应不利。

[0253] 对经选择蛋白序列的进一步分析可以鉴定更多热点氨基酸位置，这些热点氨基酸的突变在改善的M-MuLV逆转录酶的其它专利申请中有所描述：H204R—7个重复(US7078208)；H638R—4个重复(US20050232934A1)；T197A-2个重复(US7056716)；M289V(L)、T306A(M)—2个重复(US7078208)；E302K、N454K—2个重复(WO07022045A2)；E69G、L435P—1个序列(WO07022045A2)；Y64C、Q190R、V223M、F309S—1个序列(US7056716)；E562K—1个序列(US7078208)。还存在两个经选择的逆转录酶序列，所述序列具有文献中所述的3个氨基酸取代的组合(30—D200N、T306M、D524N、D583G；L5_28—T306A、F309S、D524A、H594R、F625S)。

[0254] 除了已知的突变之外，我们还鉴定出许多经常发生突变的其它氨基酸位置：D200N(A,G)—30个重复；D653N(G、A、H、V)—23个重复；L603W(M)—18个重复；T330P—15个重复；L139P—14个重复；Q221R—6个重复；T287A—6个重复；I49V(T)—5个重复；N479D—5个重复；H594R(Q)—5个重复；F625S(L)—5个重复；P65S—4个重复；H126S(R)—4个重复；L333Q(P)—4个重复；A502V—4个重复；E607K(G,A)—4个重复；K658R(Q)—4个重复；H8P(R)—3个重复；P130S—3个重复；E233K—3个重复；Q237R—3个重复；N249D—3个重复；A283D(T)—3个重复；A307V—3个重复；Y344H—3个重复；P407S(L)—3个重复；M428L—3个重复；Q430R—3个重复；D449G(A)—3个重复；A644V(T)—3个重复；N649S—3个重复；L671P—3个重复；E673G(K)—3个重复；N678I—3个重复(附录4)。

[0255] 表现最佳的RT变体通常具有被修饰最频繁的氨基酸的突变：

[0256] 20(50℃-123%)-D200N(30个重复)，L603W(18个重复)和稍微修饰的C末端-N678I，S679P，R680A；

[0257] L5_35(50℃-125%)-D200N(30个重复)，T330P(15个重复)，N479D(5个重复)；

[0258] L5_37(50℃-162%)-H123S(4个重复)，L149F(1个序列)，D200N(30个重复)，N454K(2个重复)，D583N(21个重复)；

[0259] L5_43(50℃-160%)-D200N(30个重复)，Q237R(3个重复)，T330P(15个重复)，D524G(31个重复)，F625S(5个重复)，D653N(23个重复)；

[0260] L5_46(50℃-135%)-D200N(30个重复)，T330P(15个重复)，D583N(21个重复)，T644T(3个重复)；

[0261] L5_47(50℃-179%)-N107S(1个重复)，H126R(4个重复)，T128A(1个重复)，II79V(2个重复)，D200N(30个重复)，H642Y(2个重复)，D653N(23个重复)；

[0262] L5_52(50℃-137%)-D200N(30个重复)，T330P(15个重复)，Q374R(2个重复)，D583N(21个重复)；

[0263] L5_64(50℃-142%)-D200N(30个重复)，D216G(2个重复)，D524G(31个重复)，E545G(2个重复)；

[0264] L5_65(50℃-127%)-D200N(30个重复)，Q238H(1个重复)，L570I(1个重复)，L603W

(18个重复)；

[0265] L5_68(50℃-153%)-M39V(2个重复)，I49V(2个重复)，Q91R(2个重复)，H204R(7个重复)，T287A(6个重复)，N454K(2个重复)，F625L(5个重复)，D653H(23个重复)。

[0266] 突变蛋白的所测定RT活性和序列比对分析的组合数据组允许我们确定M-MuLV逆转录酶序列中的许多有益突变(和其组合)。

[0267] 实施例3-莫洛尼氏鼠白血病病毒逆转录酶突变体的分析

[0268] 实施例2中所述的体外进化实验是非常有效的。利用温度逐渐增加的逆转录反应作为选择压力并且产生M-MuLV RT的大量不同突变变体。与最初的酶相比，它们中的大多数在升高的温度下能够表现地更好。所进化的逆转录酶的序列分析指明了热点和负责酶性质的复杂改善的最重要的氨基酸位置(置换)。为了阐明不同突变的个体影响，构建M-MuLV RT的单突变体和多突变体，并对其部分纯化和进行分析。突变体构建的起点是编码M-MuLV RT的7474bp质粒pET_his_MLV(SEQ ID No:19)，具有用于使用亲和色谱进行快速蛋白纯化的N-末端标签。确定37℃下M-MuLV逆转录酶的比活性、50℃下的相对活性和在50℃下使酶温育5分钟后在37℃下的相对残留活性。在某些情况下检查RNA酶H活性，并在不同温度下对1kb或4.5kb RNA进行cDNA合成反应。

[0269] 方法和材料

[0270] 使用起始质粒pET_his_MLV(SEQ ID No:19)作为用于突变发生PCR的起始材料。使用突变发生引物引入突变。对个体克隆进行测序和分析。在T7表达菌株ER2566中表达M-MuLV RT突变体。使相同构建中个体蛋白和起始wt M-MuLV逆转录酶在200ml LB中生长至A590为～0.7，并通过使用2ml Qiagen-Ni-NTA Superflow树脂的亲和色谱利用his-标签进行纯化(根据供应商建议在天然条件下进行所有纯化)。在1ml EB(50mM—$NaH_2PO_4$，300mM—NaCl，pH为8.0的250mM—咪唑，10mM—β-巯基乙醇和0.1%triton X-100)中进行洗脱。以50倍过量的贮存缓冲液(50mM Tris-HCl(在25℃下pH 8.3)，0.1M NaCl，1mM EDTA，5mM DTT，0.1%(v/v)Triton X-100和50%(v/v)甘油)对所有蛋白进行透析。在SDS-PAGE上检查蛋白纯度(通常为靶蛋白的～40%至80%)。使用Bradford试剂(Fermentas#R1271)确定蛋白浓度。

[0271] 在37℃、50℃下测定MLV逆转录酶的活性(将酶稀释于特定的稀释缓冲液中：30mM 25℃下pH为8.3的Tris-HCl，10mM DTT，0.5mg/ml BSA)，以及在50℃下使酶温育5分钟后在37℃下测定残留活性(将酶稀释，并在1x RT反应缓冲液：50mM25℃下pH为8.3的Tris-HCl，4mM $MgCl_2$，10mM DTT，50mM KCl中测定其稳定性)。在以下最终混合物中检测所有情况下的酶活性：50mM Tris-HCl(25℃下pH为8.3)，6mM $MgCl_2$，10mM DTT，40mM KCl，0.5mM dTTP，0.4MBq/ml[3H]-dTTP，0.4mM polyA·寡$(dT)_{12-18}$。确定活性单位，测定在特定反应温度下在10分钟内dTMP在多核苷酸级分(在DE-81上吸收)中的导入，并且与已知量的商业酶比较。根据美国专利US5405776测定M-MuLV逆转录酶变体的RNA酶H活性。在含有pH为8.3的50mM Tris-HCl，2mM $MnCl_2$，1mM DTT和[3H](A)n*(dT)n(5μM[3H](A)n，35cpm/pmol；20μM(dT)n)的反应混合物(50μl)中检测经纯化酶的RNA酶H活性。使反应在37℃下进行温育10分钟，并通过添加10μl的tRNA(1mg/ml)和20μl的冷50%TCA而终止。在冰上10分钟后，在Eppendorf离心机中使混合物离心10分钟(25000g)。在LSC-通用鸡尾酒式混合物(LSC-universal cocktail)(Roth-Rotiszint eco plus)中对40μl上清液进行计量。1单位的RNA酶H活性为：

[3H](A)n*(dT)n中，在37℃下在10分钟内使1摩尔[3H](A)n溶解所需的酶的量。

[0272] 使用"RevertAid™ First Strand cDNA Synthesis Kit"(#K1622-Fermentas)及其对照1.1kb具有3'-poly(A)尾的RNA和寡(dT)$_{18}$引物来检查经纯化逆转录酶在不同温度下合成cDNA的能力。作为选择，使用4.5kb RNA(从Eco31I线性化的pTZ19R质粒合成，其另外含有λ噬菌体DNA 5505-8469bp的碎片)来测试逆转录反应。按照所提供的策略使用试剂盒组分1μg的合成RNA，在20μl反应体积中进行1小时的cDNA合成，不同之处在于仅一些小的改动(没有在37℃下预温育5分钟)。在Eppendorf Mastercycler gradient PCR仪上使用相应的温度梯度在96孔PCR板上进行逆转录反应。通过碱性琼脂糖凝胶电泳(用溴化乙锭染色)分析所合成的cDNA。在碱性琼脂糖凝胶上cDNA合成分析的样品示于图16。

[0273] 结果

[0274] 根据M-MuLV逆转录酶进化期间的突变发现的频率，将突变分成5类：21至31个重复；14至18个重复；4至7个重复；2至3个重复和1个重复。通常根据该信息进行个体逆转录酶突变体的构建。首先测试发现频率最高的突变。确定37℃下逆转录酶的比活性、50℃下的相对活性和在50℃下使酶温育5分钟后在37℃下的相对残留活性。在某些情况下检查RNA酶H活性，并在不同温度下对1kb或4.5kb RNA进行cDNA合成反应。关于个体突变体的所有实验数据示于附录5。第二列("选择频率")表示具有确切突变的测序突变体的数量，括号中的数字表示在选择中发现的特定氨基酸突变的总数。例如D200N-25(30)表示：该天冬氨酸200替换成天冬酰胺在全部30个D200突变中发现25次。37℃下测定的逆转录酶比活性的以单位/mg蛋白给出。50℃下的相对酶活性和在50℃下温育5分钟后在37℃下的相对残留RT活性以相对于37℃下测定的相同酶的比活性(100%)标准化的百分比给出。作为对照(第1行)给出用于突变体文库构建的wt M-MuLV逆转录酶。该起始酶在与RT突变变体相同的载体中表达并以相同的方式进行纯化。Wt酶在37℃下的比活性为约200'000u/mg，在50℃下的相对活性(相对于37℃下的活性)-45-50%(90'000u/mg至100'000u/mg)，在50℃下温育5分钟后在37℃下的相对残留RT活性(相对于37℃下的活性)为约11%(～22'000u/mg)。野生型酶的RNA酶H活性为约160u/mol至200u/mol，并且在48℃下可以合成全长1kb cDNA。已知的是，M-MuLV逆转录酶通过与模板-引物底物结合而避免热失活，反之酶仅在溶液中的热稳定性较差(Gerard等，2002)。在50℃下温育5分钟后在37℃下的相对残留RT活性直接表明在不含底物的溶液中的酶的热稳定性。同时50℃下的相对活性表示在具有RNA/DNA底物的复合物中酶的热稳定性和cDNA合成的速度。逆转录酶的速度快的突变变体会在50℃下产生数量增加的聚合酶单位，即使其热稳定性与野生型酶相同。cDNA合成(在我们的情况下为1kb或4.5kb)的最高温度是最全面的参数，其表示酶于较高的温度下合成cDNA的一般能力。在37℃下的比活性(≥220'000u/mg，200'000u/mg-wt)、相对于37℃下的突变活性的在50℃下的相对活性(≥54%，45%至50%-wt)或相对于37℃下的突变活性的在50℃下温育5分钟后在37℃下的相对残留活性(≥13%，11%-wt)增加至少10%的逆转录酶突变体，以灰色阴影表示，并认为是显著改善的酶。能够在高于48℃的温度下合成全长1kb cDNA的突变体也以灰色阴影表示。

[0275] 37℃下具有更高比活性(≥220'000u/mg)的逆转录酶是(附录5)：

[0276] D200(D200N-254'000u/mg；D200G-276'000u/mg；D200H-234'000u/mg)，

[0277] T330(T330N-223'000u/mg；T330D-240'000u/mg)，

[0278] Q221(Q221R-268’000u/mg)，

[0279] H594(H594K-270’000u/mg;H594Q-231’000u/mg)，

[0280] D449(D449E-224’000u/mg;D449N-221’000u/mg)，

[0281] M39(M39N-349’000u/mg)，

[0282] M66(M66L-237’000u/mg;M66V-227’000u/mg;M66I-240’000u/mg)，

[0283] H126(H126R-227’000u/mg)，

[0284] W388(W388R-266’000u/mg)，

[0285] I179(I179V-251’000u/mg)．

[0286] 50℃下具有更高相对活性(≥54％)的逆转录酶(与37℃下的活性相比)是(附录5)：

[0287] D200(D200N-84％;D200A-87％;D200Q-103％;D200E-79％;D200V-131％;D200W-103％;D200G-88％;D200K-102％;D200R-68％;D200H-54％)，

[0288] L603(L603W-105％;L603F-104％;L603Y-95％;L603M-77％)，

[0289] D653(D653N-93％;D653K-106％;D653A-99％;D653V-98％;D653Q-93％;D653L-83％;D653H-116％;D653G-90％;D653W-93％;D653E-80％)，

[0290] T330(T330P-80％;T330N-69％;T330D-55％;T330V-65％;T330S-67％)，

[0291] Q221(Q221R-94％;Q221K-77％;Q221E-64％;Q221M-58％;Q221Y-77％)，

[0292] E607(E607K-84％;E607A-98％;E607G-72％;E607D-69％)，

[0293] L139(L139P-59％)，

[0294] T287(T287S-68％)，

[0295] N479(N479D-81％)，

[0296] H594(H594R-69％;H594K-80％;H594Q-75％;H594N-61％)，

[0297] D449(D449G-79％;D449E-77％;D449N-75％;D449A-99％;D449V-83％)，

[0298] M39(M39V-54％;M39N-71％)，

[0299] M66(M66L-79％;M66V-73％;M66I-80％)，

[0300] L333(L333Q-54％)，

[0301] H126(H126R-58％)，

[0302] P130(P130S-70％)，

[0303] Q91(Q91R-56％)，

[0304] W388(W388R-72％)，

[0305] R390(R390W-64％)，

[0306] Q374(Q374R-56％)，

[0307] E5(E5K-67％)．

[0308] 在50℃下温育5分钟后在37℃下具有增加的相对残留RT活性(≥13％)的逆转录酶(与37℃下的活性相比)是(附录5)：

[0309] D200(D200N-15％;D200A-18％;D200Q-23％;D200R-27％;D200H-27％)，

[0310] L603(L603W-23％;L603Y-13％;L603P-15％)，

[0311] D653(D653N-21％;D653K-15％;D653A-18％;D653V-16％;D653Q-18％;D653H-13％;D653G-13％;D653W-13％;D653E-19％)，

[0312] T330(T330P-21%;T330N-13%;T330D-16%;T330S-15%),

[0313] T287(T287A-13%;T287F-13%),

[0314] H594(H594R-14%;H594Q-13%),

[0315] D449(D449G-13%),

[0316] M39(M39V-13%),

[0317] M66(M66L-13%),

[0318] Y344(Y344H-13%),

[0319] Q91(Q91R-13%),

[0320] N649(N649S-16%),

[0321] W388(W388R-14%).

[0322] 能够在高于48℃的温度下合成全长1kb cDNA的突变体是(附录5):

[0323] D200(D200N-50.4℃;D200H-50.4℃),

[0324] L603(L603W-53.1℃;L603F-50.4℃;L603Y-47.8℃至50.4℃),

[0325] D653(D653N-50.4℃至53.1℃;D653K-50.4℃至53.1℃;D653A-50.4℃;D653V-50.4℃;D653Q-50.4℃;D653L-50.4℃;D653H-50.4℃至53.1℃;D653G-50.4℃;D653W-50.4℃),

[0326] Q221(Q221R-50.4℃),

[0327] E607(E607K-47.8℃至50.4℃),

[0328] H594(H594K-47.8℃至50.4℃;H594Q-47.8℃至50.4℃)

[0329] 在碱性琼脂糖凝胶上进行的1kb cDNA合成分析的样品示于图16A-D。

[0330] 根据收集的生化数据,M-MuLV逆转录酶序列中可以影响较高温度下cDNA合成的最重要的位置是:D200、L603、D653、T330、Q221、E607、L139、T287、N479、H594、D449、M39、M66、L333、H126、Y344、P130、Q91、N649、W388、R390、1179、Q374、E5。

[0331] 通常,可以对目的突变进行组合,进一步改善具有或不具有底物的M-MuLV逆转录酶热稳定性结合、速率、持续性和较高温度下合成cDNA的总体能力。表明通过组合途径改善酶的一些数据示于附录6。单突变D200N和L603W在50℃下具有的相对活性为84%和105%。1kb cDNA合成的最高温度为50.4℃和53.1℃。双重突变体D200N;L603W在50℃下的相对活性为131%,并且可以在56℃下合成1kb cDNA。三重突变体D200N;L603W;T330P(在50℃下相对活性为80%;在47.8℃下合成1kb cDNA)得到进一步改善并在50℃下具有175%的相对活性,并且可以在56℃至58℃下合成1kb cDNA。四重突变体D200N;L603W;T330P;E607K(在50℃下相对活性为84%;在47.8℃至50.1℃合成下1kb cDNA)在50℃下具有174%的相对活性,并且可以在60℃至62℃下合成1kb cDNA。五重突变体D200N;L603W;T330P;E607K;L139P(在50℃下相对活性为59%;在47.8℃下合成1kb cDNA)在50℃下具有176%的相对活性,并且可以在62℃下合成1kb cDNA,该温度比野生型M-MuLV逆转录酶(在47.8℃下合成1kb cDNA)的温度高约14℃。在以下情况下也观察到热稳定性的加和特征:

[0332] N479D,H594R突变体(D200N;L603W-50℃下相对活性为131%,在56℃下合成1kb cDNA,而D200N;L603W;N479D;H594R-在50℃下相对活性为182%,在56℃至58℃下合成1kb cDNA,在56℃至58℃下合成4.5kb cDNA),

[0333] T330P突变体(D200N;L603W;D653N;D524G-在50℃下相对活性为155%,在58℃至

60℃下合成1kb cDNA，而D200N;L603W;D653N;D524G;T330P-在50℃下相对活性为180％，在60℃至62℃下合成1kb cDNA)。

[0334] 在碱性琼脂糖凝胶上进行的4.5kb cDNA合成分析的样品示于图16E-G。

[0335] 实施例4-CRD的改进-使用生物素dUTP选择DNA依赖性DNA聚合酶活性(原理证明)

[0336] 该实施例阐释了作为DNA依赖性DNA聚合酶的逆转录酶的基于活性的选择策略，该酶能够将修饰的核苷酸引入到DNA-DNA底物中。选择的原理性示意图图示于图12。使用两个质粒pET_his_MLV_D583N_pD(编码减去RNA酶H的与蛋白D间隔子融合的莫洛尼氏鼠白血病病毒)及其衍生物pET_his_del_pD(编码失活的逆转录酶;pol结构域中57个氨基酸缺失，并且RNA酶H结构域中存在突变D583N，实施例1，SEQ ID No:2)作为该实施例的起始材料。使用质粒pET_his_MLV_D583N_pD和pET_his_del_pD作为靶，在两个单独的聚合酶链式反应中合成起始的编码活性逆转录酶和失活逆转录酶的DNA片段。所合成的PCR片段用于合成mRNA的转录反应，所述mRNA的3'末端缺乏终止密码子。将经纯化的mRNA以1:20＝MLV(活性RT):del(失活RT)的比例混合。通过T4DNA连接酶将双链DNA衔接体(选择DNA依赖性DNA聚合酶活性所需)连接到3'mRNA。该mRNA/dsDNA复合物用于体外翻译反应。沿着mRNA移动的核糖体在RNA-DNA杂交的位点停止翻译(Tabuchi等，2001)。通过用冰冷的含有50mM $Mg^{2+}$的缓冲液稀释翻译混合物，从而使核糖体-mRNA/dsDNA-蛋白复合物稳定化(与常规的核糖体展示一样)。通过蔗糖垫超速离心对三元复合物(TC)的混合物进行纯化。使用含有与体外翻译的M-MuLV(RNA酶H-)逆转录酶连接的mRNA-dsDNA的经纯化的三元复合物(取<$3\times10^9$个分子)来制备反应混合物，该反应混合物另外补充有生物素-dUTP和反应缓冲液。对冰冷的RT反应混合物进行乳化，产生～$1\times10^{10}$个大小为～2μm的油包水区室。在37℃下使乳化的反应混合物(小于一个TC，核糖体-mRNA/dsDNA-蛋白/区室)温育30分钟。区室化反应混合物的温度升高后，大多数TC解离，释放mRNA/dsDNA和逆转录酶。仅在含有活性M-MuLV(RNase H-)逆转录酶的区室中可以发生成功的引入反应，导致mRNA/dsDNA复合物的生物素化。使生物素化的复合物选择性地固定于链亲和素包被的磁珠上，并通过RT-PCR特异性扩增。作为成功实验的结果，编码活性酶的基因(在我们的情况下为MLV_D583N_pD逆转录酶的RT-PCR片段)应该相对于编码失活酶(del_pD)的基因富集。

[0337] 方法和材料

[0338] mRNA/dsDNA复合物的制备

[0339] (1)连接效率的确定。通过使用ddC-Long2引物(SEQ ID No:22)引物作为夹板序列(splint)将MLV_pD mRNA(从pET_his_VlLV_pD质粒合成，实施例1)与引物Long+Tb(SEQ ID No:23)连接，从而确定连接反应的效率。预先使用T4多核苷酸激酶(Fermentas)将Long+Tb的5'末端磷酸化。

[0340] 通过将8.5pmol(～10μg)的经纯化MLV_pD mRNA与摩尔数4倍过量的Long+Tb(SEQ ID No:23)以及摩尔数4.2倍过量的ddC-Long2(SEQ ID No:22)于不含核酸酶的水中混合，制备36μl的退火混合物。在70℃下使混合物温育5分钟，然后冷却至室温20分钟。添加连接反应组分之前，将混合物移到冷却台2分钟。

[0341] 向36μl的退火混合物中添加4.5μl的10x连接缓冲液和4.5μl的T4DNA连接酶(5v/μl)(Fermentas)。在37℃下使连接反应进行30分钟，随后用等体积的Roti®苯酚/氯仿(ROTH)提取1次，并用等体积的氯仿提取两次。假设mRNA的量接近于为连接反应而取的起始量，利

用不含核酸酶的水将～5μg的连接产物混合物稀释到43μl。在固定于Dynabeads M-280链亲和素珠(Dynabeads® kilobase BINDER™ Kit(DYNAL Biotech))上之前，留出2μl所得混合物的等分试样以用于在琼脂糖凝胶上的分析。

[0342] 将10μl再悬浮的Dynabeads转移到1.5ml微量离心管中，用50μl试剂盒中所提供的结合溶液洗涤。将试管置于磁体上1分钟至2分钟直到珠于所述试管中沉淀，并且除去溶液。通过吹吸在38μl的补充有2μl tRNA(来自酵母的tRNA(Roche))水溶液(1μg/μl)的结合缓冲液中温和地使Dynabead再悬浮，从而使得非特异性mRNA结合最小化。向含有连接产物混合物的溶液(～40μl)中添加40μl在结合溶液中的Dynabead。在+22℃下在热混合器(Eppendorf)中振摇60分钟，对试管进行温育。与所连接的mRNA/dsDNA结合后，除去上清液，并将所述珠用50μl洗涤液(试剂盒中提供)洗涤3次。使得所收集的具有固定物(连接mRNA/dsDNA复合物)的Dynabead再悬浮于26μl不含核酸酶的水，随后用40μl的Roti®苯酚/氯仿(ROTH)提取1次，并用40μl体积的氯仿提取两次，从而使mRNA/dsDNA复合物从磁珠释放。将1、2和5μl最终混合物与进行固定之前留出的样品(2μl)和Mass Ruler™高范围DNA ladder一起在琼脂糖凝胶上分析。确定在链亲和素珠上纯化的mRNA/DNA复合物中mRNA的量，比较固定前后的mRNA的量。mRNA的回收率为～60％，这意味着至少60％的mRNA与DNA双链成功连接，产生mRNA/dsDNA复合物。

[0343] (2)确定生物素-dUTP导入到mRNA/dsDNA复合物和导入到自引导mRNA中的效率。如第一次mRNA与dsDNA连接实验所示，连接反应效率为～60％以上。连接混合物中剩下的游离mRNA可以自引导，并参与使用M-MuLV逆转录酶和生物素-dUTP的延伸反应。进行该实验，从而证明mRNA/dsDNA复合物(连接产物)是比游离的自引导mRNA更好的底物。

[0344] 按照上述方法将MLV_D583N_pD mRNA与long+寡核苷酸(SEQ ID No:21)连接(起始质粒pET_his_MLV_D583N_pD和pET_his_del_pD的构建详细地描述于实施例1)。将～12,5ng制备的(MLV_D583N_pD mRNA/long+)底物与预先在70℃下温育5分钟的～12,5ng ng的del_pD mRNA混合，然后在不含核酸酶的水中冷却至室温20分钟，总体积为12.5μl。制备用于通过逆转录酶导入dTTP或生物素-dUTP的第二混合物：8μl 5x用于逆转录酶的反应缓冲液；1μl—40u/μl RiboLock™ RNA酶抑制剂(Fermentas)；18.6μl不含核酸酶的水；0.4μl—200u/μl RevertAid™ Minus M-MuLV逆转录酶(Fermentas)。将所制备的混合物分成2x 15μl两试管，并添加1μl的1mM dTTP(Fermentas)或1μl的1mM生物素-dUTP(Fermentas)。随后向含有dTTP和生物素-dUTP的混合物中添加5μl底物(来自第一混合物)。在37℃下使反应进行60分钟。然后向两个样品添加1μl 0.5M EDTA(pH 8.0)，对反应混合物用等体积的Roti®苯酚/氯仿(ROTH)提取1次，用等体积的氯仿提取1次，然后在G-50微柱(GE Healthcare)上进行纯化。留下2μl所得反应产物用于直接RT反应(无进行链亲和素珠纯化)，并将剩余部分的溶液用于生物素化mRNA-dsDNA复合物在Dynabeads M-280链亲和素珠(Dynabeads® kilobase BINDER™ Kit(DYNAL Biotech))上的固定。将10μl再悬浮的Dynabead转移到1.5ml微量离心管，用25μl所提供的结合缓冲液洗涤。将试管置于磁体上1分钟至2分钟直到珠沉淀在所述试管的底部，并且除去溶液。通过吹吸在90μl结合缓冲液中温和地使Dynabead再悬浮，并添加2μl tRNA(来自酵母的tRNA(Roche))水溶液(1μg/μl)，从而使得非特异性mRNA结合最小化。向含有延伸有通过dTTP和生物素-dUTP延伸的RNA-DNA片段的溶液(～40μl)中添加40μl

在结合溶液中的40μl Dynabead。通过振摇(1400rpm)，使得试管在+22℃下在热混合器(Eppendorf)中温育40分钟。结合步骤后，除去上清液，并用50μl洗涤液(试剂盒中提供)在+22℃下振摇(1400rpm)5分钟，对所述珠进行洗涤3次。使所收集的具有固定物(延伸mRNA-dsDNA复合物)的延伸mRNA-dsDNA复合物的Dynabead再悬浮于逆转录反应混合物中。在冰上制备逆转录反应混合物：20μl-5X用于逆转录酶的反应缓冲液(Fermentas)；10μl-10mM dNTP(Fermentas)；2.5μl-40u/μl RiboLock RNA酶抑制剂(Fermentas)；5μl-20μM pD_42寡核苷酸(SEQ ID No:8)；2,5μl RevertAid™ Minus M-MuLV逆转录酶(200u/μl)(Fermentas)；55μl不含核酸酶的水。将所制备的混合物分成5个19μl(5x19μl)的等分试样。:两个用于再悬浮使具有固定的延伸的mRNA/dsDNA复合物或mRNA的Dynabead再悬浮，将另两个转移到具有未进行链亲和素珠纯化的所留样品的试管，向剩下的19μl等分试样中添加1μl不含核酸酶的水-阴性反应对照，用于证明反应混合物未受到DNA污染。通过在+42℃下在热混合器(Eppendorf)中振摇(1000rpm)1小时，使对所有反应混合物进行温育，直到cDNA合成反应结束。

[0345] 通过巢式PCR进行cDNA的扩增。在冰上制备起始PCR混合物：14μl-10X含有KCl的Taq缓冲液(Fermentas)；14μl-2mM各dNTP(Fermentas)；8.4μl-25mM $MgCl_2$(Fermentas)；2.8μl-1u/μl LC(重组)Taq DNA聚合酶(Fermentas)；0.7μl-100μM M_F引物(SEQ ID No:11)；0.7μl-100pM M_2R引物(SEQ ID No:12)；92.4μl水—将混合物分成7个19μl的样品(7x19μl)。向5x 19μl的PCR主混合物中添加1μl的cDNA(RT样品1至5)；向第6管PCR主混合物中添加1μl的水-阴性PCR对照；向第7管(阳性PCR对照)-添加1μl的pET_his_MLV_D583N_pD质粒(～1ng)。循环方案为：在94℃下进行起始变性步骤3分钟，25个循环(94℃下45秒钟，57℃下45秒钟和72℃下1分钟)和最后在72℃下延伸3分钟。

[0346] 对于MLV_D583N_pD，预测的扩增子大小为907bp，对于del_pD cDNA，预测的大小为736bp。在每孔加载10μl PCR混合物的1％琼脂糖凝胶上分析PCR产物(图13)。

[0347] 如所预期的，仅在使用生物素-dUTP的延伸反应中观察到有效的cDNA扩增。在使用dTTP的延伸反应中可以检测到所扩增cDNA的非常微弱的带，并且可以解释为mRNA与链亲和素珠的弱的非特异性结合。在链亲和素珠上纯化后，编码MLV_D583N_pD基因(907bp)的DNA相对于del_pD(736bp)的DNA富集。这意味着mRNA/dsDNA复合物通过生物素-dUTP的延伸比自引导del_pD mRNA要有效的多。

[0348] (3)mRNA混合物(MLV_D583N_pD:del_pD＝1:20)和RNA/dsDNA复合物的制备。

[0349] 用于体外转录的PCR片段的制备。在冰上制备PCR混合物：20μl-10X含有KCl的Taq缓冲液(Fermentas)；20μl-2mM的各dNTP(Fermentas)；12μl-25mM $MgCl_2$(Fermentas)；16μl-DMSO(D8418-Sigma)；4μl-1u/μl LC(重组)Taq DNA聚合酶(Fermentas)；1μl-100μM pro-pIVEX引物(SEQ ID No:3)；1μl-100μM pD-ter引物(SEQ ID No:20)；122μl水—将混合物分成2x 98μl的两试管。向2x 98μl的PCR主混合物中添加2μl pET_hisJViLV_D583N_pD(稀释至～1ng/μl)或2μl pET_his_del_pD(稀释至～1ng/μl)(起始质粒pET_his_MLV_D583N_pD和pET_his_del_pD的构建详细地描述于实施例1)。循环方案为：在94℃下进行起始变性步骤3分钟，30个循环(94℃下45秒钟，53℃下45秒钟和72℃下2分钟)和最后在72℃下延伸5分钟。从2ng质粒(7873bp)至～5μg(50ng/μl)扩增产物(来自pET_his_MLV_pD的2702bp PCR片段MLV_D583N_pD；来自pET_his_del_pD的2531bp PCR片段)，扩增效率为～7000倍。

[0350] 制备转录混合物:80μl-5x T7转录缓冲液(1M HEPES-KOH pH 7.6;150mM乙酸镁;10mM亚精胺;0.2M DTT);56μl-112mM各NTP(Fermentas);16μl-20u/μl T7RNA聚合酶(Fermentas);8μl-40u/μl RiboLock RNA酶抑制剂(Fermentas);114μl不含核酸酶的水—将混合物分成2x 165μl两试管,并且添加35μl-20ng/μl的PCR片段MLV_D583N_pD(未经纯化的PCR混合物)或35μl-20ng/μl的PCR片段del_pD PCR(未经纯化的PCR混合物)。转录在37℃下进行2小时。

[0351] 用冰冷的不含核酸酶的水将两种转录混合物都稀释到200μl,并且添加200μl的6M LiCl溶液。将混合物在+4℃下温育25分钟,在+4℃下于冷却的离心机中以最大速度(25'000g)离心25分钟。弃去上清液,用500μl冰冷的75%乙醇洗涤RNA团块。再次将试管在+4℃下以最大速度离心5分钟,并弃去上清液。在室温下对RNA团块干燥5分钟,并随后通过在+4℃和1400rpm下振摇15分钟使其再悬浮于400μl不含核酸酶的冰冷的水中。再次将试管在+4℃下以最大速度离心5分钟,以分离未溶解的RNA。将约380μl的上清液移到具有42μl 10X DNA酶I缓冲液($Mg^{2+}$)(Fermentas);3μl-1u/μl的DNA酶I(不含RNA酶)(Fermentas)的新试管中,并且在37℃下温育20分钟,从而降解DNA。对反应混合物用等体积的Roti®苯酚/氯仿(ROTH)提取1次,用等体积的氯仿提取2次,从而除去DNA酶I。向各试管添加43μl pH为5.0的3M乙酸钠溶液和1075μl冰冷的96%乙醇。最后通过在-20℃下温育30分钟,并在+4℃下以最大速度(25'000g)离心25分钟使RNA沉淀。弃去上清液,用500μl冰冷的75%乙醇洗涤RNA团块。再次将试管在+4℃下以最大速度离心4分钟,并弃去上清液。在室温下对RNA团块干燥5分钟,并随后通过在+4℃下振摇(1400rpm)15分钟使其再悬浮于150μl不含核酸酶的冰冷的水中。将RNA溶液等分为10μl,并用液氮冷冻。用分光光度法测定mRNA浓度,并在与RiboRuler™ RNA Ladder,高范围(Fermentas)一同进行的琼脂糖凝胶上复核。

[0352] 通过使用ddC-Long2寡脱氧核苷酸作为夹板将long+寡脱氧核苷酸与mRNA连接,从而产生mRNA/dsDNA复合物。使用T4多核苷酸激酶(Fermentas)使Long+的5'末端预先磷酸化。寡脱氧核苷酸ddC-Long2具有3'末端修饰(ddC),从而防止逆转录酶对其天然RNA-DNA底物的3'末端延伸的可能性。

[0353] 通过将17pmol比例为1:20=MLV_D583N_pD(活性RT):del_pD(失活RT)的经纯化mRNA混合物与在不含核酸酶的水中的摩尔数4.3倍过量的long+和摩尔数4.1倍过量的ddC-Long2混合,制备50μl的退火混合物。在70℃下使所述混合物温育5分钟,然后冷却至室温20分钟。添加连接反应组分之前,将混合物移到冷却台2分钟。

[0354] 向40μl的退火混合物中添加5μl的10x连接缓冲液和5μl的T4DNA连接酶(5v/μl)(Fermentas)。使用在无T4DNA连接酶的1x连接缓冲液中的相同退火混合物进行阴性连接反应。在37℃下使制备的连接反应混合物温育30分钟。为了终止连接,向两个试管添加1μl的0.5M EDTA(pH 8.0),并对反应混合物用等体积的Roti®苯酚/氯仿(ROTH)提取1次,并用等体积的氯仿提取两次,然后在30℃下在真空浓缩器5301(Eppendorf)中对反应产物浓缩10分钟。使用illiustra ProbeQuant G-50微柱(GE Healthcare)进行脱盐。通过利用RiboRuler™ RNA Ladder,高范围(Fermentas)进行的琼脂糖凝胶确定连接产物的浓度-与long+/ddC-Long2寡脱氧核苷酸连接的mRNA混合物(MLV_D583N_pD:del_pD=1:20)~0.24μg/μl(具有T4DNA连接酶的样品),和具有long+/ddC-Long2寡脱氧核苷酸的简单mRNA混合物~0.06μg/μl(无T4DNA连接酶的样品)。

[0355] (4)通过引入[α-$P^{33}$]dATP获得的mRNA/dsDNA复合物的一般对照。测试所制备的mRNA/dsDNA复合物(底物)的通过逆转录酶进行的dTTP(或生物素-dUTP)以及随后[α-$P^{33}$]dATP导入。反应混合物:16μl 5x用于逆转录酶的反应缓冲液;4μl-40u/μl RiboLock™ RNA酶抑制剂(Fermentas);2μl[α-$P^{33}$]dATP(10mCi)/ml,SRF-203(Hartmann Analytic));35μl不含核酸酶的水;1μl-200u/μl RevertAid™ Minus M-MuLV逆转录酶(Fermentas)。将所制备的混合物分成2x 28μl的两试管,并添加2μl 1mM dTTP(Fermentas)或2μl 1mM生物素-dUTP(Fermentas)。将所得混合物分成2x 15μl的两试管。向第一管添加1.25μl(~0.3μg)连接产物和3.75μl不含核酸酶的水。向第二管添加5μl(~0.3μg)阴性连接反应产物。使反应混合物在37℃下温育30分钟。然后,向所有试管添加1μl 0.5M EDTA(pH 8.0),对反应混合物用等体积的Roti®苯酚/氯仿(ROTH)提取1次,用等体积的氯仿提取1次,然后在illiustra G-50微柱(GE Healthcare)上进行纯化。在利用RiboRuler™ RNA Ladder,高范围(Fermentas)进行的琼脂糖凝胶上分析反应产物(图14A)。在所有样品(具有或不具有连接酶)中我们可以看到del_pD mRNA的离散的带(~2500b)(比在mRNA混合物中存在的不能被区分的MLV_D583N_pD mRNA~2700b的量小20倍)。随后在滤纸上对琼脂糖凝胶进行干燥,并且仅在阳性连接样品(具有连接酶)的情况下检测放射性标记的mRNA/dsDNA复合物(在与mRNA相同的位置),而在阴性连接样品(不具有连接酶)的情况下不进行检测(图14B)。

[0356] (5)使用生物素-dUTP对DNA依赖性DNA聚合酶活性进行的选择。将之前制备的mRNA/dsDNA复合物(mRNA混合物MLV_D583N_pD(活性RT):del_pD(失活RT)=1:20)用于使用合成WakoPURE系统(295-59503-Wako)的体外翻译。用于WakoPURE的翻译混合物(25μl):12,5μl-A溶液(Wako);5μl-B溶液(Wako);0,5μl-40u/μl RiboLock RNA酶抑制剂(Fermentas);0.25μl-1M DTT;1,75μl不含核酸酶的水和5μl-0.24μg/μl mRNA/dsDNA底物(~1200ng)。在37℃下进行体外翻译120分钟。通过添加155μl冰冷的终止缓冲液$WBK_{500}$+DTT+triton(25℃下pH为7.5的50mM tris-乙酸盐;50mM NaCl;50mM乙酸镁;500mM KCl;10mM DTT;0.1%(v/v)-Triton x-100(T8787-Sigma))终止翻译物(~25μl),并在+4℃和25'000g下离心5分钟。非常仔细地将160μl经离心的翻译混合物转移到透明的1ml超速离心管(343778-Beckman)中至840μl 35%的蔗糖在$WBK_{500}$+DTT+Triton X-100中的溶液(50mM在25℃下pH为7.5的Tris-乙酸盐;50mM NaCl;50mM乙酸镁;500mM KCl;10mM DTT;0.1%(v/v)-Triton x-100(T8787-Sigma);35%(w/v)-蔗糖(84097-Fluka))的顶部。通过在+4℃下以100,000rpm在TL-100Beckman超速离心机中用TLA100.2固定角转头(Beckman)进行超速离心9分钟,对由mRNA/dsDNA-核糖体-蛋白(tRNA)组成的三元复合物(TC)进行纯化。最初从离心管的最上部移出750μl溶液。然后(非常仔细地使位于超速离心管底部的TC的小透明团块保持完整)用750μl $WBK_{500}$(50mM在25℃下pH为7.5的tris-乙酸盐;50mM NaCl;50mM乙酸镁;500mM KCl)洗涤管壁。最后从离心管的最上部开始移出所有溶液,并且将团块溶解于30μl冰冷的终止缓冲液$WBK_{500}$+DTT+triton(50mM在25℃下pH为7.5的tris-乙酸盐;50mM NaCl;50mM乙酸镁;500mM KCl;10mM DTT;0.1%(v/v)-Triton x-100(T8787-Sigma))中。

[0357] 如超速离心后使用放射性标记的mRNA所确定,5%至30%输入的mRNA位于三元复合物团块中。因此据预测在30μl缓冲液中具有低于360ng(翻译反应中所用的1200ng mRNA的30%)的mRNA(~12ng/μl或$9\times10^9$个分子/μl的三元复合物)。

[0358] 在冰上通过将以下成分混合而制备导入经修饰核苷酸的反应混合物:5μl-5X用于

逆转录酶的反应缓冲液(Fermentas);1.25μl-40u/μl RiboLock RNA酶抑制剂(Fermentas);2.5μl-1mM生物素-dUTP(Fermentas);40.95μl不含核酸酶的水和0.3μl经纯化的(<2.7×$10^9$个分子)TC。根据所述方案,50μl的核苷酸导入反应混合物含有<2.7×$10^9$个分子的三元复合物。

[0359] 通过将ABIL EM 90(Goldschmidt)混入矿物油(M5904-Sigma)至终浓度4%(v/v)来制备油-表面活性剂混合物(Ghadessy和Holliger,2004;US2005064460)。通过将950μl油-类表面活性剂混合物与50μl RT混合物混合,在+4℃下在5ml冷冻小瓶(cryogenic vial)(430492-Corning)中制备乳液。使用速度控制在~2100rpm的MS-3000磁力搅拌器、具有中心环的Rotilabo®-(3x8mm)磁棒(1489.2-Roth)进行混合;每30秒钟向水相添加10μl等分试样,继续混合另外2分钟(总混合时间-4分钟)。根据光学显微镜数据,所制备乳液中的区室大小为0.5μm~10μm,平均直径为~2μm。因此据预期对含有低于2.7×$10^9$个分子的三元复合物的50μl逆转录反应混合物进行乳化后具有~1×$10^{10}$个油包水区室(约1个mRNA-dsDNA复合物和逆转录酶分子/3至4个区室)。

[0360] 使所制备的乳液在+37℃下温育30分钟。

[0361] 为了从乳液中回收反应混合物,向乳液添加20μl 0.1M EDTA,搅拌10秒钟,然后添加50μl苯酚/氯仿混合物,并搅拌另外10秒钟。然后,将乳液转移到1.5ml微量离心管,添加0.5ml水饱和的醚,通过涡旋混合,并在室温下以16'000g离心10分钟。除去油-醚相,在试管底部留下浓缩(但仍完整)的乳液。最后通过用0.9ml水饱和的醚提取;0.9ml水饱和的乙酸乙酯(为了除去ABIL EM 90洗涤剂)提取;用0.9ml水饱和的醚提取2次将乳液破乳。在室温下对水相真空干燥12分钟,然后在illiustra ProbeQuant G-50微柱(GE Healthcare)上除去导入的核苷酸。留出所得混合物的2μl等分试样用于直接RT反应(未进行链亲和素珠纯化),并将溶液的其余部分用于生物素化mRNA-dsDNA复合物在Dynabeads M-280链亲和素珠(DYNAL Biotech)上的固定。

[0362] 根据所提供的产品说明书,使用Dynabeads® kilobase BINDER™ Kit(DYNAL Biotech)来分离生物素化mRNA-dsDNA复合物。将5μl再悬浮的Dynabead转移到1.5ml微量离心管,用20μl所提供的结合缓冲液洗涤。将试管置于磁体上1分钟至2分钟直到珠在所述试管中沉淀,并且除去溶液。通过移液在50μl结合缓冲液中温和地使Dynabead再悬浮,并添加1μl tRNA(来自酵母的tRNA(Roche))水溶液(1μg/μl),从而使得非特异性mRNA结合最小化。向含有生物素化RNA-DNA片段的溶液(~50μl)中添加50μl在结合缓冲液中的Dynabead。使试管在+22℃下在热混合器(Eppendorf)中振摇温育1小时。mRNA/dsDNA结合后,从所述珠中除去上清液,并在+22℃下用50μl的洗涤液(试剂盒中提供)振摇(1400rpm)5分钟对所述珠洗涤2次,并在+22℃下用50μl的洗涤液振摇(1400rpm)12分钟对所述珠洗涤1次。

[0363] 使所收集的具有所固定的生物素化mRNA-dsDNA复合物的Dynabead再悬浮于逆转录反应混合物中。

[0364] 在冰上制备用于所选择的mRNA/dsDNA复合物的逆转录反应混合物:12μl-5X用于逆转录酶的反应缓冲液(Fermentas);6μl-10mM dNTP(Fermentas);1.5μl-40u/μl RiboLock RNA酶抑制剂(Fermentas);0.3μl-20μM pD_42寡核苷酸(SEQ ID No:8);1.5μl RevertAid™ Minus M-MuLV逆转录酶(200u/μl)(Fermentas);35.7μl不含核酸酶的水。将所制备的混合物分成3个19μl的等分试样:一个用于使具有固定的生物素化mRNA/dsDNA复合

物的Dynabead再悬浮，将另一个19μl等分试样转移到具有留出的未进行链亲和素珠纯化的延伸mRNA/dsDNA复合物样品的试管，向剩下的19μl等分试样中添加1μl不含核酸酶的水(阴性反应对照-用于证明反应混合物未受到污染)。通过+42℃下在热混合器(Eppendorf)中振摇(1000rpm)1小时，使所有反应混合物进行温育。

[0365] 通过巢式PCR进行cDNA的扩增。在冰上制备起始PCR混合物：10μl-10X含有KCl的Taq缓冲液(Fermentas)；10μl-2mM各dNTP(Fermentas)；6μl-25mM $MgCl_2$(Fermentas)；2μl-1u/μl LC(重组)Taq DNA聚合酶(Fermentas)；1μl-2.5u/μl PfuDNA聚合酶(Fermentas)；0.5μl-100μM RD_Nde引物(SEQ ID No:9)；0.5μl-100pM pD_55引物(SEQ ID No:10)；65μl水--将混合物分成5个19μl的样品(5x19μl)。向3x 19μl的PCR主混合物中添加1μl cDNA(RT样品1至3)；向一个具有19μl的PCR主混合物的试管中添加1μl水-阴性PCR对照。对于阳性PCR对照，添加1μl的pET_his_MLV_pD质粒(～1ng)。循环方案为：在94℃下进行起始变性步骤3分钟，30个循环(94℃下45秒钟，58℃下45秒钟和72℃下3分钟)和最后在72℃下延伸5分钟。

[0366] 在冰上制备用于部分基因扩增(为了更好地分辨RT样品中MLV_D583N_pD:del_pD cDNA的比例)的巢式PCR混合物：20μl-10X含有KCl的Taq缓冲液(Fermentas)；20μl-2mM各dNTP(Fermentas)；12μl-25mM $MgCl_2$(Fermentas)；0.9μl-5u/μl Taq DNA聚合酶(Fermentas)；1.0μl-100μM M_F引物(SEQ ID No:11)；1.0μl-100μM M_2R引物(SEQ ID No:12)；135.1μl水-将混合物分成5x 38μl。向所制备的巢式PCR混合物添加2μl第一次PCR(引物组RD_Nde//pD_55)产物。将Master mix再次混合，并分成两试管(2x 20μl)以用于30个或35个PCR循环扩增。循环方案为：在94℃下进行起始变性步骤3分钟，30或35个循环(94℃下45秒钟，57℃下45秒钟和72℃下1分钟)和最后在72℃下延伸3分钟。对于MLV_D583N_pD，PCR片段的预期长度是907bp，对于del_pD是736bp。在每孔加载10μl PCR混合物的1％琼脂糖凝胶上对扩增进行分析(图15)。

[0367] 结果

[0368] 1.使用T4DNA连接酶将双链(dsDNA)衔接体成功地连接到mRNA。如通过dsDNA-生物素衔接体的连接所确定的，连接效率是约60％。可以在链亲和素珠上对mRNA/dsDNA复合物进行特异性纯化，提供了分辨生物素标记的底物和未标记的底物的机会。与mRNA/dsDNA相比，游离mRNA对于DNA依赖性DNA聚合酶是差得多的底物。因此，dsDNA与mRNA的60％连接效率足够良好，并且所述底物可以成功地用于进化流程。

[0369] 2.进行使用mRNA/dsDNA复合物(mRNA混合物MLV_D583N_pD:del_pD＝1:20)的一般选择实验。使用WakoPURE蛋白翻译系统进行体外翻译，并且进行区室化生物素-dUTP导入到dsDNA的反应，以证明编码活性逆转录酶的基因(MLV_D583N_pD)相对于编码失活酶的基因(del_pD)的富集。根据选择流程(图12)，生物素-dUTP的导入反应应该仅发生在含有活性(MLV_D583N_pD)逆转录酶的水性区室中，导致mRNA/dsDNA复合物的生物素化。通过使生物素化的复合物与固定在磁珠上的链亲和素结合，从而选择DNA依赖性DNA聚合酶，然后通过RT-PCR对所选择的基因进行扩增。编码活性酶的基因(在我们的情况下是MLV_D583N_pD逆转录酶的RT-PCR片段)相对于编码失活酶的基因富集(图15)。基因MLV_D583N_pD:del_pD的起始比例为1:20，最终比例(富集后)是～1:1。个别地，在该特定实验中富集倍数是～20倍。不同实验中计算的富集倍数是5～200。并且经确认可以应用区室化核糖体展示(CRD)法的，

通过经修饰核苷酸的导入选择DNA依赖性DNA聚合酶。可以立即进行常规DNA依赖性DNA聚合酶的选择。生物素-dUTP可以与不同的的目核苷酸类似物(包括具有3'修饰的核苷酸类似物)交换。将所述核苷酸类似物导入DNA链后,3'末端被封闭,不能延长,并且会终止延伸反应。该方法用于边合成边测序(SBS)方案,并且利用区室化核糖体展示(CRD)技术可以容易地对适合于SBS的DNA聚合酶进行进化。

[0370] 附录1

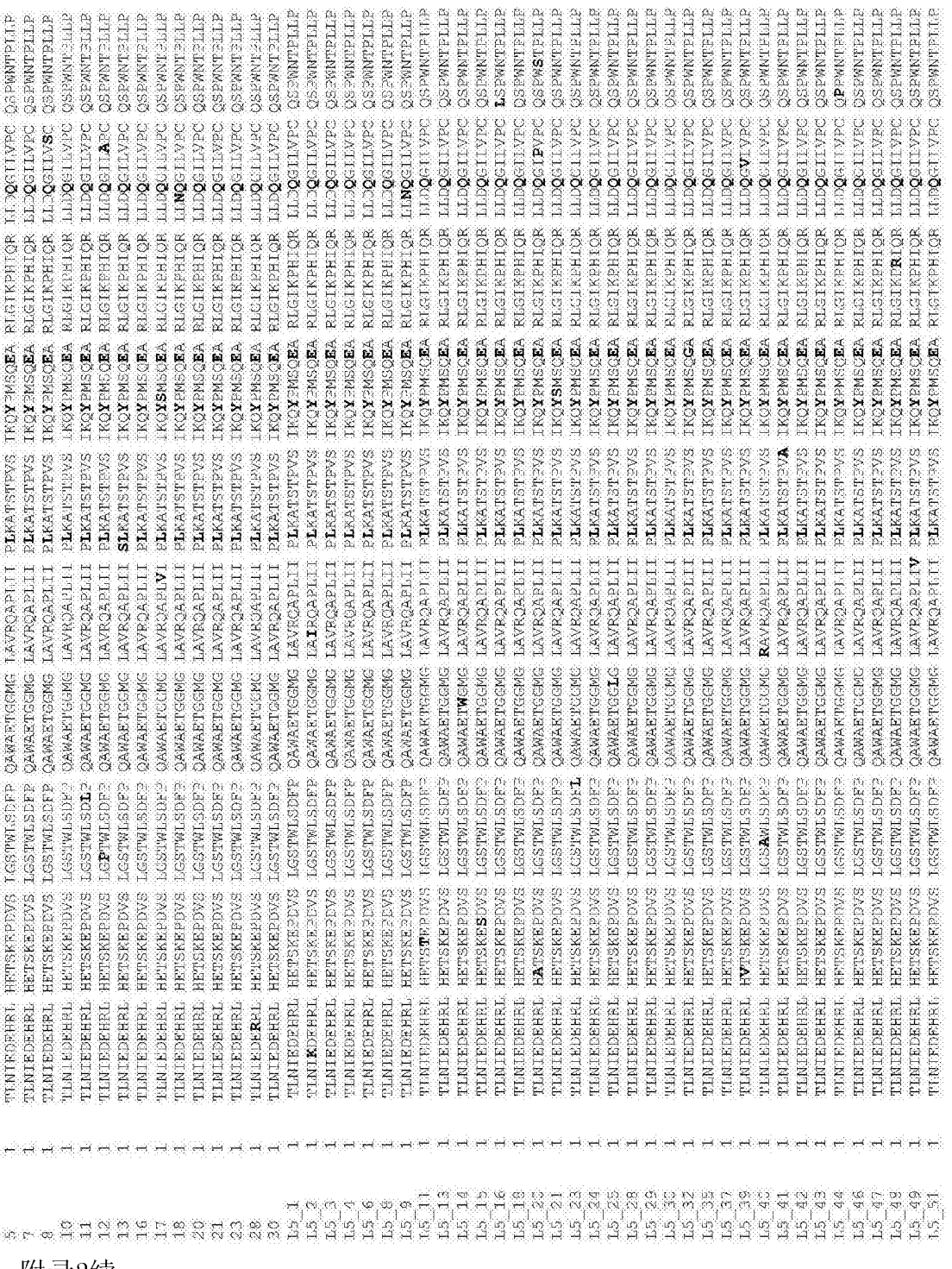
NcoI

```
ccatgggcatgaccctaaatatagaagatgagcatcggctacatgagacctcaaaagagc    60
cagatgtttctctagggtccacatggctgtctgattttcctcaggcctgggcggaaaccg   120
ggggcatgggactggcagttcgccaagctcctctgatcatacctctgaaagcaacctcta   180
ccccccgtgtccataaaacaatacccccatgtcacaagaagccagactgggggatcaagcccc  240
                                  g                   g
acatacagagactgttggaccagggaatactggtaccctgccagtcccccctggaacacgc   300
                                        a
ccctgctacccgttaagaaaccagggactaatgattataggcctgtccaggatctgagag    360
                        g                   g
aagtcaacaagcgggtggaagacatccaccccaccgtgcccaaccccttacaacctcttga   420
                  c
gcgggctcccaccgtcccaccagtggtacactgtgcttgatttaaaggatgccttttttct   480
            c
gcctgagactccacccccaccagtcagcctctcttcgcctttgagtggagagatccagaga   540
tgggaatctcaggacaattgacctggaccagactcccacagggtttcaaaaacagtcccca   600
                           c                              t  t
ccctgtttgatgaggcactgcacagagacctagcagacttccggatccagcacccagact    660
              g
tgatcctgctacagtacgtggatgacttactgctggccgccacttctgagctagactgcc    720
aacaaggtactcgggccctgttacaaaccctagggaacctgggtatcgggcctcggcca     780
             g                    c
agaaaagcccaaatttgccagaaaacaggtcaagtatctggggtatcttctaaaagagggtc   840
agagatggctgactgaggccagaaaaagagactgtgatgggggcagcctactccgaagaccc   900
                                                    a
ctcgacaactaagggagttcctagggacggcaggcttctgtcgcctctggatccctgggt    960
ttgcagaaatggcagccccccttgtacctctcaccaaaaacgggggactctgtttaattgggg 1020
gccccagaccaacaaaaggcctatcaagaaatcaagcaagctcttctaactgccccagccc  1080
                                    c
tggggttgccagatttgactaagccctttgaactctttgtcgacgagaagcagggctacg   1140
ccaaaggtgtcctaacgcaaaaactgggacctggcgtcggccggtggcctacctgtcca    1200
aaaagctagacccagtagcagctgggtggccccctgcctacggatggtagcagccattg    1260
ccgtactgacaaaggatgcaggcaagctaaccatgggacagccactagtcattctggccc   1320
  ~~~~~~~                    ~~~~~~~~~~~~~~~~
ccccatgcagtagaggcactagtcaaacaaccccccgaccgctggctttccaacgcccgga  1380
tgactcactatcaggccttgcttttggacacggaccgggtccagttcggaccggtggtag   1440
ccctgaacccggctacgctgctcccactgcctgaggaagggctgcaacacaactgccttg   1500
             g
atatcctggccgaagcccacggaacccgacccgacctaacggaccagccgctcccagacg   1560
                                  c       g
ccgaccacacctggtacacggatggaagcagtctcttacaagagggacagcgtaaggcgg   1620
gagctgcggtgaccaccgagaccgaggtaatctgtggctaaagccctgccagccgggacat   1680
ccgctcagcgggctgaactgatagcactcacccaggccctaaagatggcagaaggtaaga   1740
                                                        g
agctaaatgtttatactgatagccgttatgcttttgctactgcccatatccatggagaaa   1800
tatacagaaggcgtgggttgctcacatcagaaggcaaaagagatcaaaaataaagacgaga   1860
tcttggccctactaaaagccctctttctgcccaaaagacttagcataatccattgtccag   1920
               a                    g
gacatcaaaaggagacacagcgccgaggctagaggcaaccggatggctgaccaagcggccc   1980
gaaaggcagccatcacagagactccagacacctctaccctcctcatagaaaattcatcac   2040
ccaattcccgcttaattaatgaattc
                    EcoRI
```

[0371]

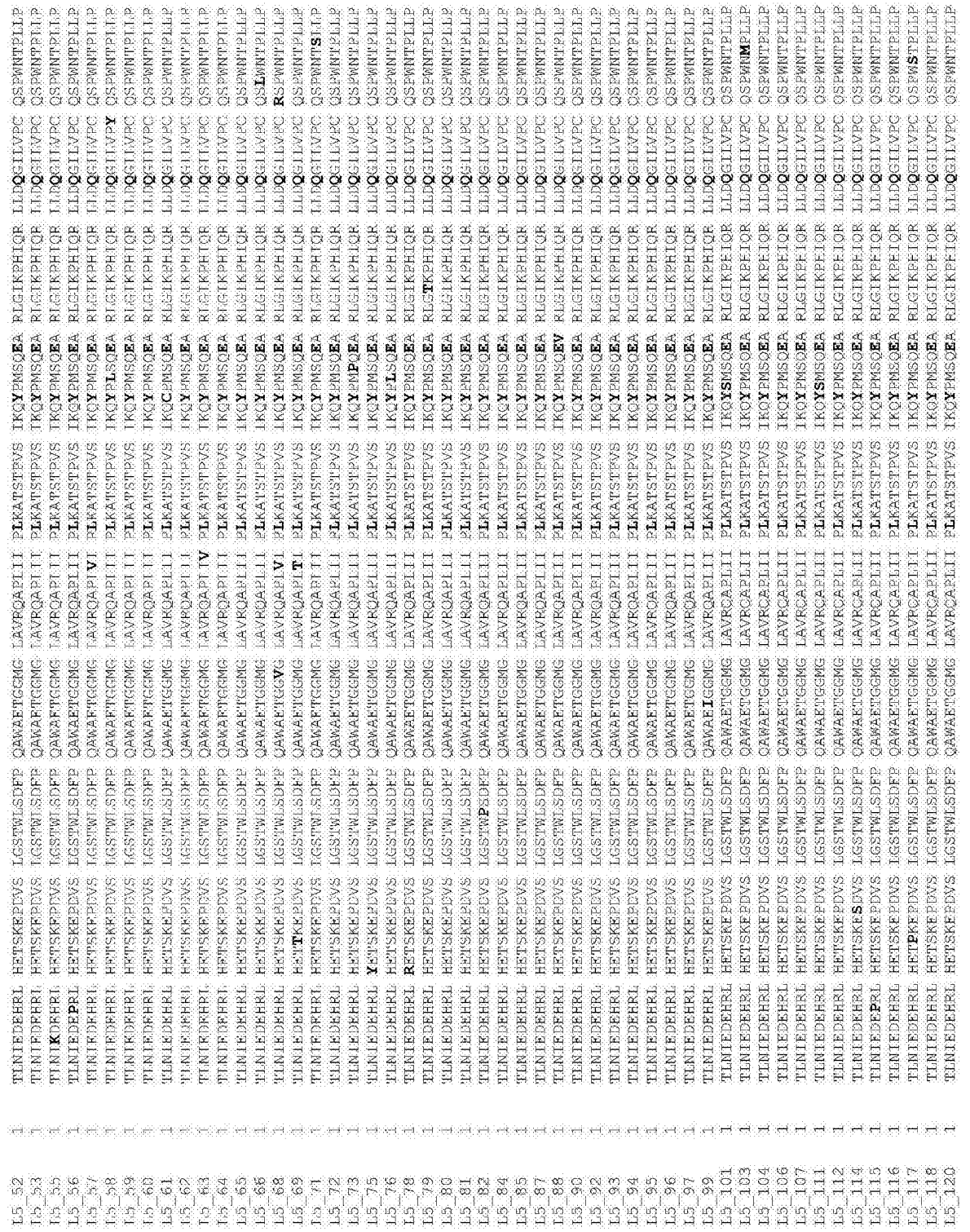
[0372] 附录2

[0373]



[0374] 附录2续

[0375]



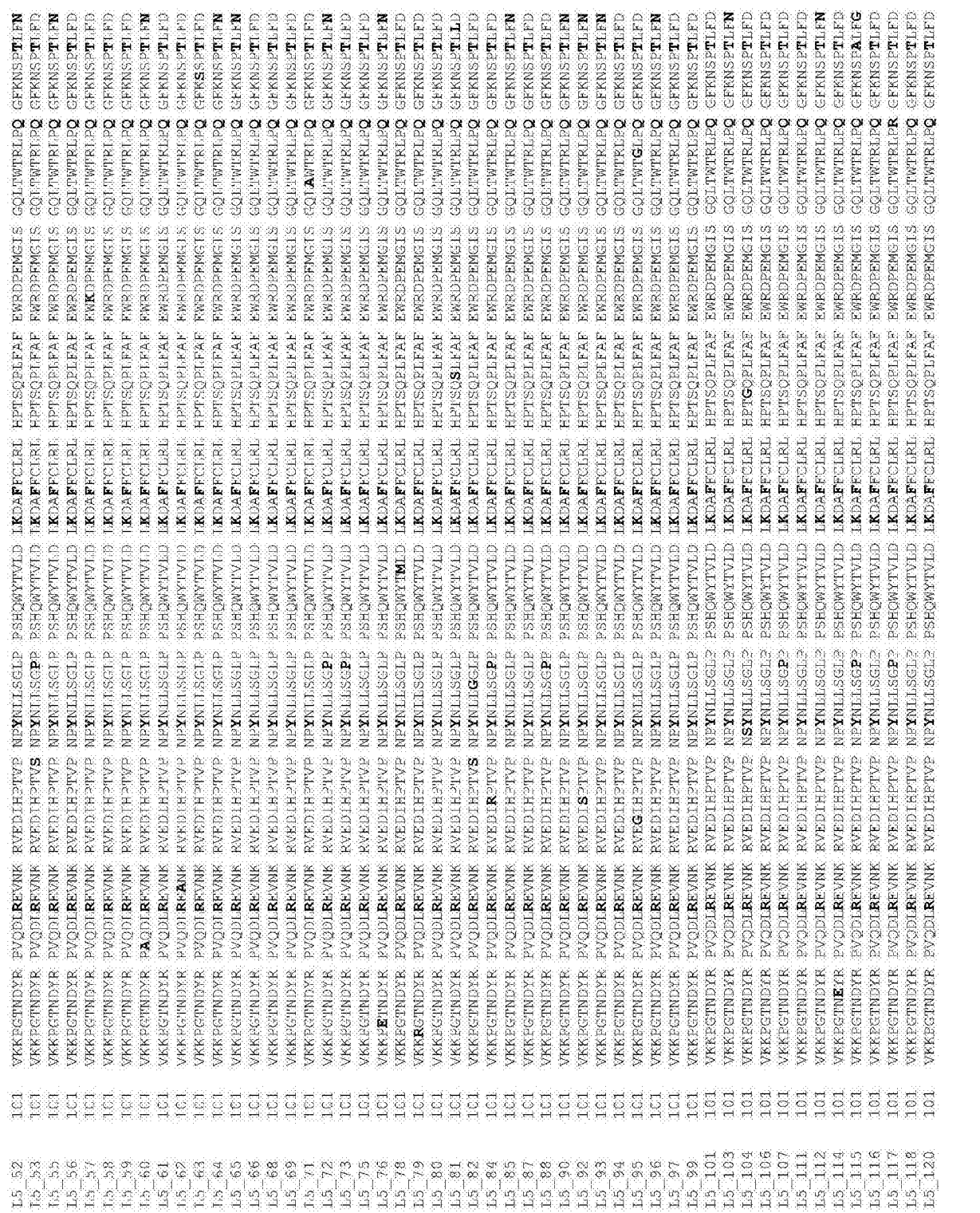
[0376] 附录2续

[0377]



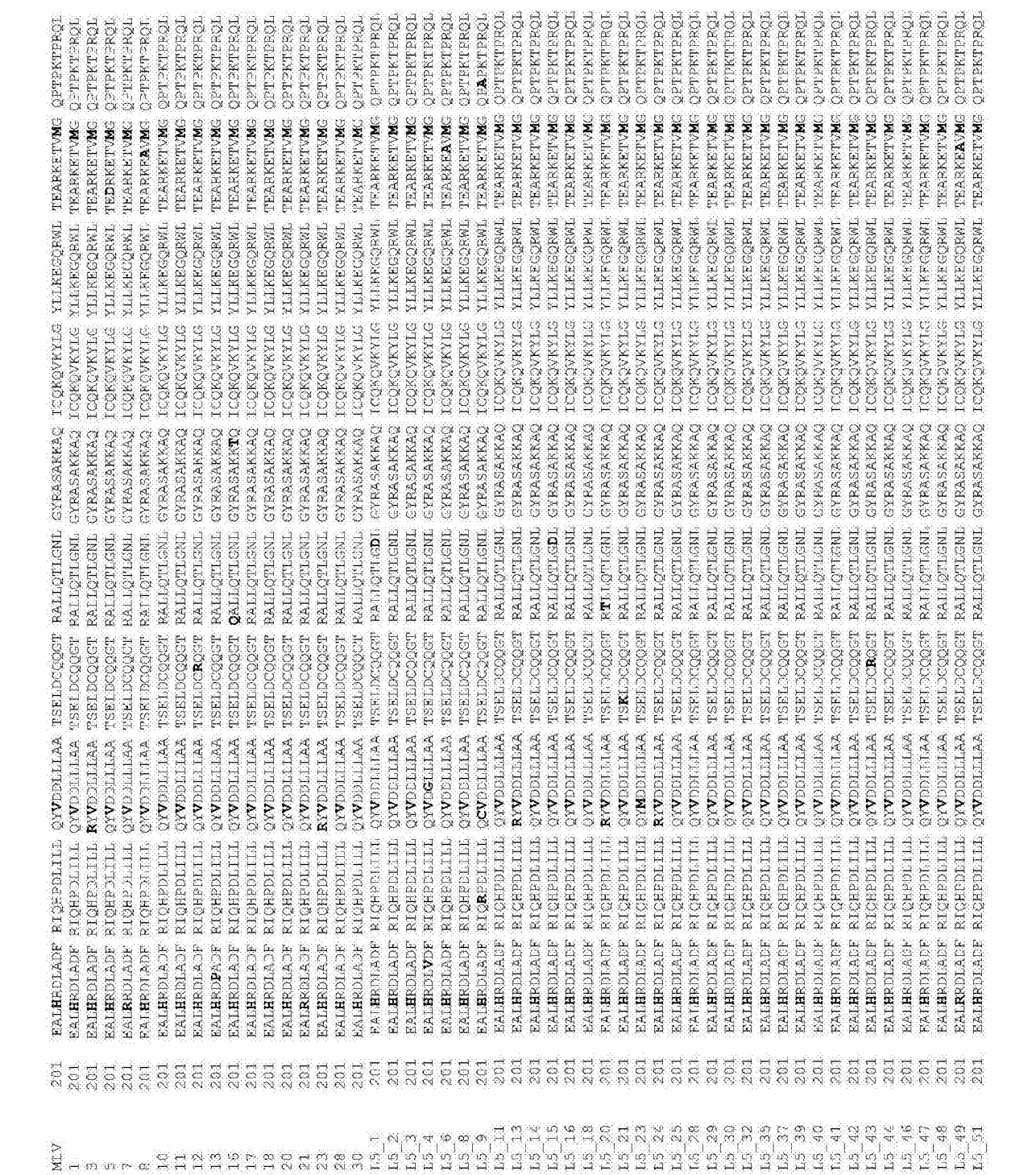
[0378] 附录2续

[0379]



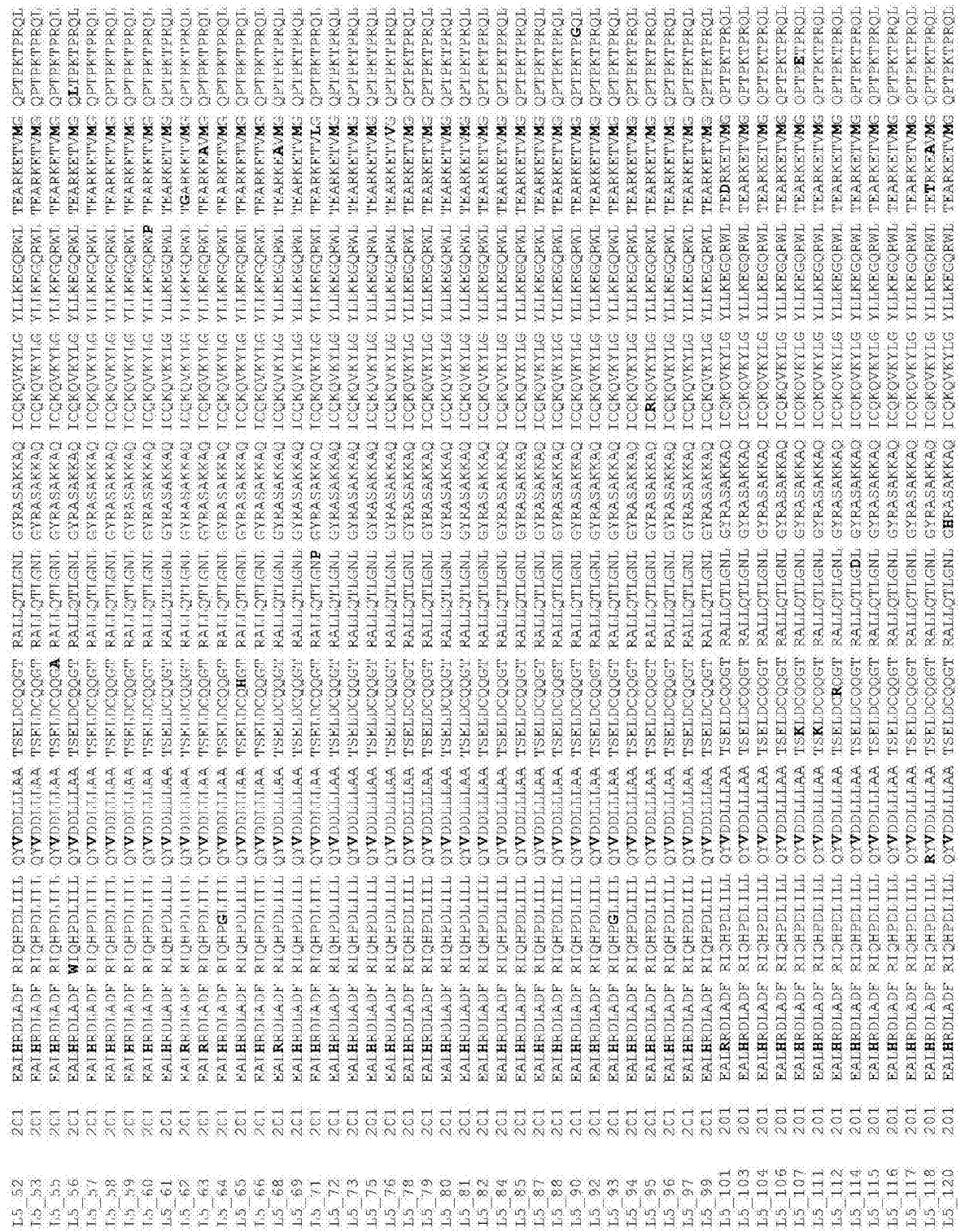
[0380] 附录2续

[0381]



[0382] 附录2续

[0383]



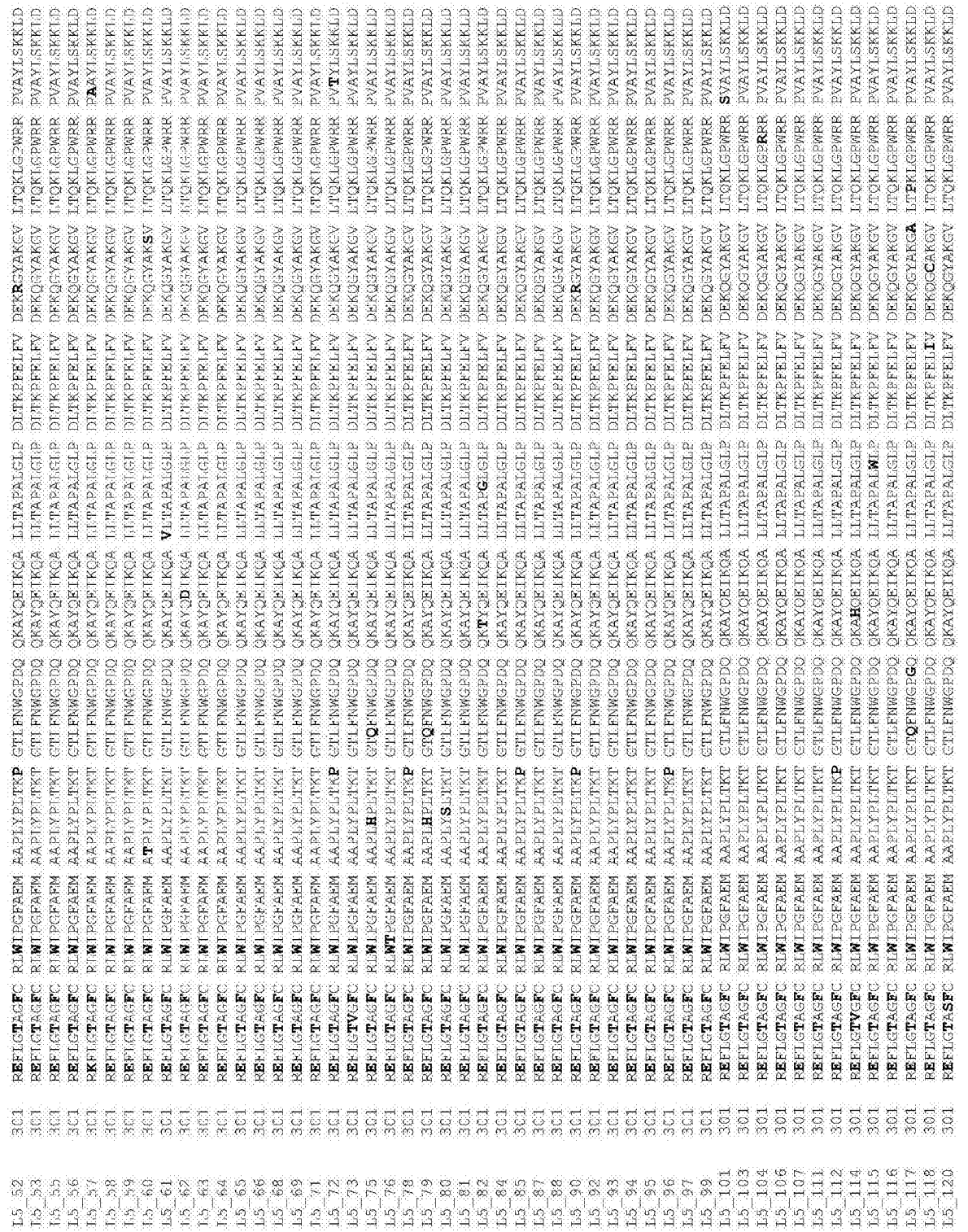
[0384] 附录2续

[0385]

| | MLV | 301 | REFLGTAGFC RLWIPGFAEM AAPLYPLTKT GTLFNWGPDQ QKAYQEIKQA LLTAPAIGLP DLTKPFELFV DEKQGYAKGV LTQKLGPWRR PVAYLSKKLD |
|---|---|---|---|
| | 1 | 301 | REFLGTAGFC RLWIPGFAEM AAPLYPLTKT GTLFNWGPDQ QKAYQEIKQA LLTAPAIGLP DLTKPFELFV DEKQGYAKGV LTQKLGPWRR PVAYLSKKLD |
| | 3 | 301 | REFLGTAGFC RLWIPGFAEM AAPLYPLTKT GTLFNWGPDQ QKAYQEIKQA LLTAPALGLP DLTKPFELFV DEKQGYAKGV LTQKLGPWRR PVAYLSKKLD |
| | 5 | 301 | REFLGTAGFC RLWIPGFAEM AAPLYPLTKT GTLFNWGPDQ QKAYQEIKQA LLTAPALVLP DLTKPFELFV DEKQGYAKGV LTQKLGPWRR PVAYLSKKLD |
| | 7 | 301 | REFLGTAGFC RLWIPGFAEM AAPLYPLTKT GTLFNWGPDQ QKAYQEIKQA LLTAPALGLP DLTKPFELFV DEKQGYAKGV LTQKLGPWRR PVAYLSKKLD |
| | 8 | 301 | REFLGTAGFC RLWIPGFAEM AAPLYPLTKP GTLFNWGPDQ QKAYQEIKQA LLTAPALGLP DLTKPFELFV DEKQGYAKGV LTQKLGPWRR PVAYLSKKLD |
| | 10 | 301 | REFLGTAGFC RLWIPGFAEM AA2LYPLTKT GTLFNWGPDQ QKAYQEIKQA LLTAPALGLP DLTKPFELFV DEKQGYAKGV LTQKLGPWRR PVAYLSKKLD |
| | 11 | 301 | REFLGTAGFC RLWIPGFAEM AA2LYPLTKT GTLFNWGPDQ QKAYQEIKQA LLTAPALGLP DLTKPFELFV DEKQGYAKGV LTQKLGPWRW PVAYLSKKLD |
| | 12 | 301 | REFLGTAGFC RLWIPGFAEM AA2LYPLTKP GTLFNWGPDQ QKAYQEIKQA LLTAPALGLP DLTKPFELFV DEKQGYAKGV LTQKLGPWRR PVAYLSKKLD |
| | 13 | 301 | REFLGTAGFC RLWIPGFAEM AA2LYPLTKT GTLFNWGPDQ QKAYQEIKQA LLTAPALGLP DLTKPFELFV DEKQGYAKGV LTQKLGPWRR PVAYLSKKLD |
| | 16 | 301 | REFLGTAGFC HLWIPGFAEM AA2LYPLTKT GTLFNWGPDQ QKAYQEIKQA LLTAPALGLP DLTKPFELFV DEKQGYAKGV LTQKLGPWRR PVAYLSKKLD |
| | 17 | 301 | REFLGTAGFC RLWIPGFAEM AA2LYPLTKT GTLFNWGPDQ QKAYQEIKQA LLTAPALGLP DLTKPFELFV DEKQGYAKGV LTQKLGPWRR PVAYLSKKLD |
| | 18 | 301 | REFLGTAGFC RLWIPGFAEM AA2LYPLTKT GTLFNWGPDQ QKAYQEIKQA LLTAPAIGLP DLTKPFELFV DEKQGYAKGV LTQKLGPWRR PVAYLSKKLD |
| | 20 | 301 | REFLGTAGFC RLWIPGFAEM AA2LYPLTKT GTLFNWGPDQ QKAYQEIKQA LLTAPALGLP DLTKPFELFV DEKQGYAKGV LTQKLGPWRR PVAYLSKKLD |
| | 21 | 301 | REFLGTAGFC RLWIPGFAEM AALLYPLTKT GTLFNWGPDQ QKAYQEIKQA LLAAPALGLP DLTKPFELFV DEKQGYAKGV LTQKLGPWRR PVAYLSKKLD |
| | 23 | 301 | REFLGTAGFC RLWIPGFAEM AA2LYPLTKT GILFNWGPDQ QKAYQEIKQA LLTAPALGLP DLTKPFELFV DEKQGYAKGV LTQKLGPWRR PVAYLSKKLD |
| | 28 | 301 | REFLGTAGFC RLWIPGFAEM AA2LYPLTKT GTLFNWGPDQ QKAYQEIKQA LLTACALGLP DLTKPFELFV DEKQGYAKGV LTQKLGPWRR PVAYLSKKLD |
| | 30 | 301 | REFLSMAGFC RLWIPGFAEM AA2LYPLTKT GTLFNWGPDQ QKAYQEIKQA LLTAPALGLP DLTKPFELFV DEKQGYAKGV LTQKLGPWRR PVAYLSKKLD |
| L5_1 | | 301 | REFLGTAGFC RLWIPGFAEM AAPLYPLTKT GTLFNWGPDQ QKAYQEIKQA LLTAPALGLP DLTKPFELFV DEKQGYAKGV LTQKLGPWRR PVAYLSKKLD |
| L5_2 | | 301 | REFLGTAGFC RLWIPGFAEM AAPLYPLTKT GTLFNWGPDQ QKAYQEIKQA LLTAPALGLP DLTKPFELFV DEKQGYAKGV LTQKLGPWRR LVAYLSKKLD |
| L5_3 | | 301 | REFLGTAGFC RLWIPGFAEM AAPLYPLTKT GTLFNWGPDQ QKAYQEIKQA LLTAPALGLP DLTKPFELFV DEKQGYAKGA LTQKLGPWRR PVAYLSKKLD |
| L5_4 | | 301 | REFLGTAGFC RLWIPGFAEM AAPLYPLTKT GTLFNWGPDQ QKAYQEIKQA LLTAPALGLP DLTKPFELFV DEKQGYAKGV LTQKLGPWRR PVAYLSKKLD |
| L5_6 | | 301 | REFLGTAGFC RLWIPGFAEM AAPLYPLTKT GTLFNWGPDQ QKAYQEIKQA LLTAPALGLP DLTKPFELFV DEKQGYAKGV LTQKLGPWRR PVAYLSKKLD |
| L5_9 | | 301 | REFLGTAGFC RLWIPGFAEM AAPLYPLTKT GTLFNWGPDQ QKAYQEIKQA LLTAPALGLP DLTKPFELFV DEKQGYAKGV LTQKLGPWRR PVAYLSKKLD |
| L5_9 | | 301 | REFLGTAGFC RLWIPGFAEM AAPLYPLTKT GTLFNWGPDQ QKAHQEIKQA LLTAPALGLP DLTKPFELFV DEKQGYAKGV LTQKLGPWRR PVAYLSKKLD |
| L5_11 | | 301 | REFLGTAGFC RLWIPGFAEM AAPLYPLTKT GTLFNWGPDQ QKAYQEIKQA LLTAPALGLP DLTKPFELFV DEKQGYAKGV LTQKLGPWRR PVAYLSKKLD |
| L5_13 | | 301 | REFLGTAGFC RLWIPGFAEM AAPLYPLTKP GTLFNWGPDQ QKAYQEIKQA LLTAPALGLP DLTKPFELFV DEKQGYAKGV LTQKLGPWRR PVAYLSKKLD |
| L5_14 | | 301 | REFLGTAGFC RLWIPGFAEM AAPLYPLTKT GTLFNWGPDQ QKAYQEIKQA LLTAPALGLP DLTKPFELFV DEKQGYAKGV LTQKLGPWRR PVAYLSKKLD |
| L5_15 | | 301 | REFLGTYGFC RLWIPGFAEM AAPLYPLTKT GTLFNWGPDQ QKAHQEIKQA LLTAPALGLP DLTKPFELFV DEKQGYAKGV LTQKLGPWRR PVAYLSKKLD |
| L5_16 | | 301 | REFLGTAGFC RLWIPGFAEM AAPLYPLTKT GTLFNWGPDQ QKAYQEIKQA LLTAPALGLP DLTKPFELFV DEKQGYAKGV LTQKLGPWRR PVAYLSKKLD |
| L5_18 | | 301 | REFLGTAGFC RLWIPGFAEM AAPLYPLTKT GTLFNWGPDQ QKAYQEIKQA LLTAPALGLP DLTKPFELFV DEKQGYAKGV LTQKLGPWRR PVAYLSKKLD |
| L5_20 | | 301 | REFLGTAGFC RLWIPGFAEM AAPLYPLTKT GTLFNWGPDQ QKAYQEIKQA LLTAPALGLP ETTKPFELFV DEKQGYAKGV LTQKLGPWRR PVAYLSKKLD |
| L5_21 | | 301 | REFLGTAGFC RLWIPGFAFM AAPLYPLTKT GTLFNWGPDQ QKAYQEIKQA LLTAPALGLP DLTKPFELFV DEKQGYAKGV LTQKLGPWRR PVAYLSKKLD |
| L5_23 | | 301 | REFLGTAGFC RLWIPGFAEM AAPLYPLTKT GTLFNWGPDQ QKAYQEIKQA LLTAPALGLP DLTKPFELFV DEKQGYAKGV LTQKLGPWRR PVAYLSKKLD |
| L5_24 | | 301 | REFLGTAGFC RLWICGFAEM AAPLYPLTKT GTLFNWGPDQ QKAYQEIKQA LLTAPALGLP DLTKPFELFV DEKQGYAKGV LTQKLGPWRR PVAYLSKKLD |
| L5_25 | | 301 | RKFLGTAGFC RLWIPGFAEM AAPLYPLTKT GTLFNWGPDQ QKAYQEIKQA LLTAPALGLP DLTKPFELFV DEKQGYAKGV LTQKLGPWRR PVAYLSKKLD |
| L5_28 | | 301 | REFLGAAGSC RLWIPGFAEM AAPLYPLTKT GTLFNWGPDQ QKAYQEIKQA LLTAPALGLP DLTKPFELFV DEKQGYAKGV LTQKLGPWRR PVAYLSKKLD |
| L5_29 | | 301 | REFLGTAGFC RPWIPGFAEM AAPLYPLTKP GTLFNWGPDQ QKAYQEIKQA LLTAPALGLP DLTKPFELFV DEKQGYAKGV LTQKLGPWRR PVAYLSKKLD |
| L5_30 | | 301 | REFLGTAGFC RLWIPGFAEM AAPLYPLTKP GTLFNWGPDQ QKAYQEIKQA LLTAPALGLP DLTKPFELFV DEKQGYAKGV LTQKLGPWRR PVAYLSKKLD |
| L5_32 | | 301 | REFLGTAGFC RLWIPGFAEM AAPLYPLTKT GTLFNWGPDQ QKAYQEIKQA LLTAPALGLP DLTKPFELFV DEKQGYAKGV LTQKLGPWRR PVAYLSKKLD |
| L5_35 | | 301 | REFLGTAGFC RLWIPGFAEM AAPLYPLTKP GTLFNWGPDQ QKAYQEIKQA LLTAPALGLP DLTKPFELFV DEKQGYAKGV LTQKLGPWRR PVAYLSKKLD |
| L5_37 | | 301 | REFLGTAGFC RLWIPGFAEM AAPLYPLTKT GTLFNWGPDQ QKAYQEIKQA LLTAPALGLP DLTKPFELFV DEKQGYAKGV LTQKLGPWRR PVAYLSKKLD |
| L5_39 | | 301 | REFLGTAGFC HLWIPGFAEM AAPLYPLTKT GTLFNWGPDQ QKAYQEIKQA LLTAPALGLP DLTKPFELFV DEKQGYAKGV LTQKLGPWRR PVAYLSKKLD |
| L5_40 | | 301 | REFLGTAGFC RLWIPGFAEM AAPLYPLTKP GTLFNWGPDQ QKAYQEIKQA LLTAPALGLP DLTKPFELFV DEKQGYAKGV LTQKLGPWRR PVAYLSKKLD |
| L5_41 | | 301 | REFLGTAGFC RLWIPGFAEM AAPLYPLTKT GTTFNWGPDQ QKAYQEIKQA LLTAPALGLP DLTKPFELFV DEKQGYAKGV LTQKLGPWRR PVAYLSKKLD |
| L5_42 | | 301 | REFLGTAGFC RLWIPGFAEM AAPLYPLTKT GTLFNWGPDQ QKAYQEIKQA LLTAPALGLP DLTKPFELFV DEKQGYAKGV LTQKLGPWRR PVAYLSKKLD |
| L5_43 | | 301 | REFLGTAGFC RLWIPGFAEM AAPLYPLTKP GTLFNWGPDQ QKAYQEIKQA LLTAPALGLP DLTKPFELFV DEKQGYAKGV LTQKLGPWRR PVAYLSKKLD |
| L5_44 | | 301 | REFLGTAGFC RLWIPGFAEM AAPLYPLTKT GTLFNWGPDQ QKAYQEIKQA LLTAPALGLP DLTKPFELFV DEKQGYAKGV LTQKLGPWRR PVAYLSKKLD |
| L5_46 | | 301 | REFLGTAGFC RLWIPGFAEM AAPLYPLTKP GTLFNWGPDQ QKAYQEIKQA LLTAPALGLP DLTKPFELFV DEKQGYAKGV LTQKLGPWRR PVAYLSKKLD |
| L5_47 | | 301 | REFLGTAGFC RLWIPGFAEM AAPLYPLTKT GTTFNWGPDQ QKAYQEIKQA LLTAPALGLP DLTKPFELFV DEKQGYAKGV LTQKLGPWRR PVAYLSKKLD |
| L5_48 | | 301 | REFLGTAGFC RLWIPGFAEM AAPLYPLTKT GTLFNWGPDQ QKAYQEIKQA LLTAPALGLP DLTKPFELFV DEKQGYAKGV LTQKLGPWRR PVAYLSKKLD |
| L5_49 | | 301 | REFLGTAGFC RLWIPGFAEM AAPLYPLTKT GTLFNWGPDQ QKAYQEIKQA LLTAPALGLP DLTKPFELFV DEKQGYAKGV LTQKLGPWRR PVAYLSKKLD |
| L5_51 | | 301 | REFLGTAGFC RLWIPGFAPM AAPLYPLTKT GTLFNWGPDQ QKAYQEIKQA LLTAPALGLP DLTKPFELFV DEKQGYAKGV LTQKLGPWRR PVAYLSKKLD |

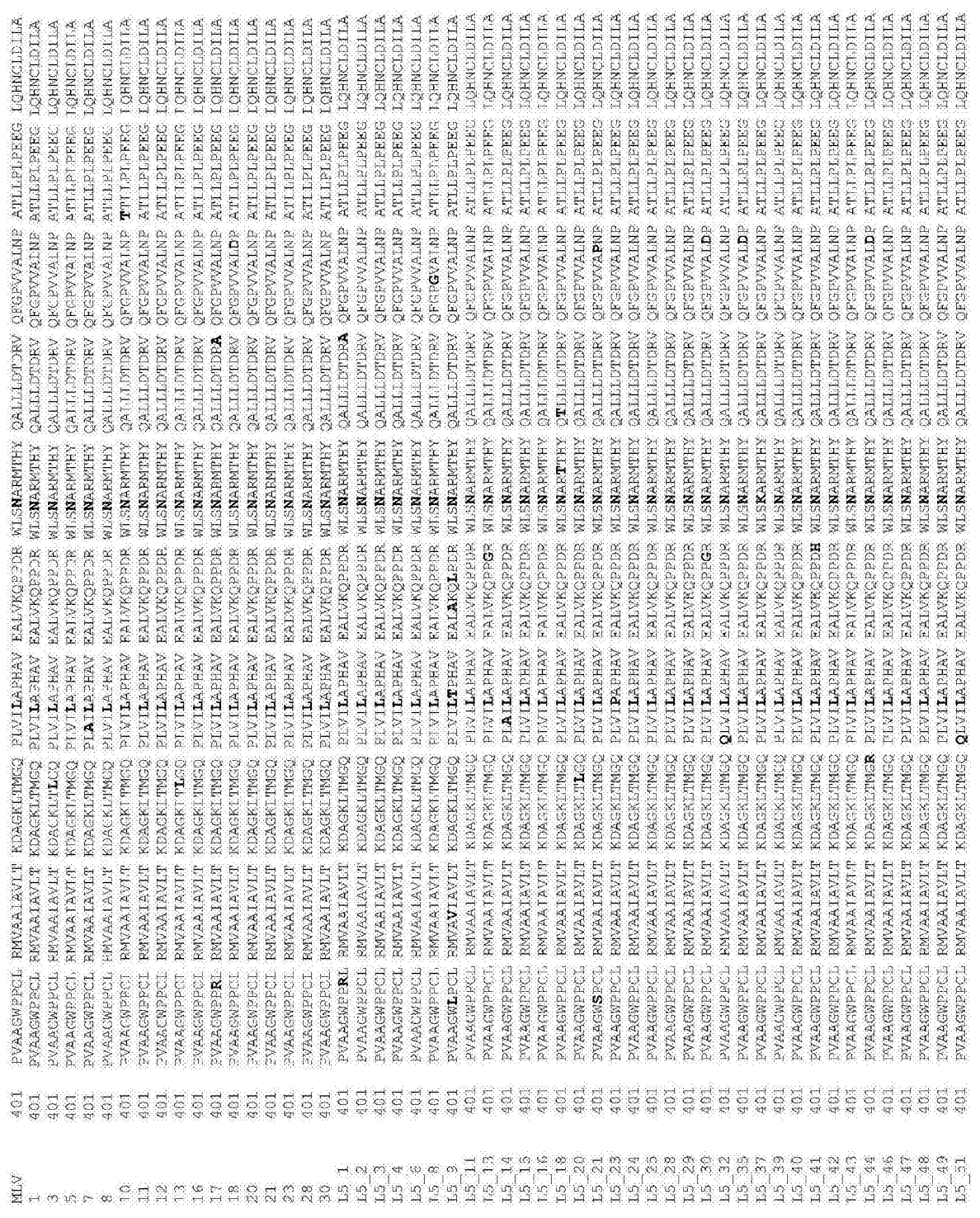
[0386] 附录2续

[0387]



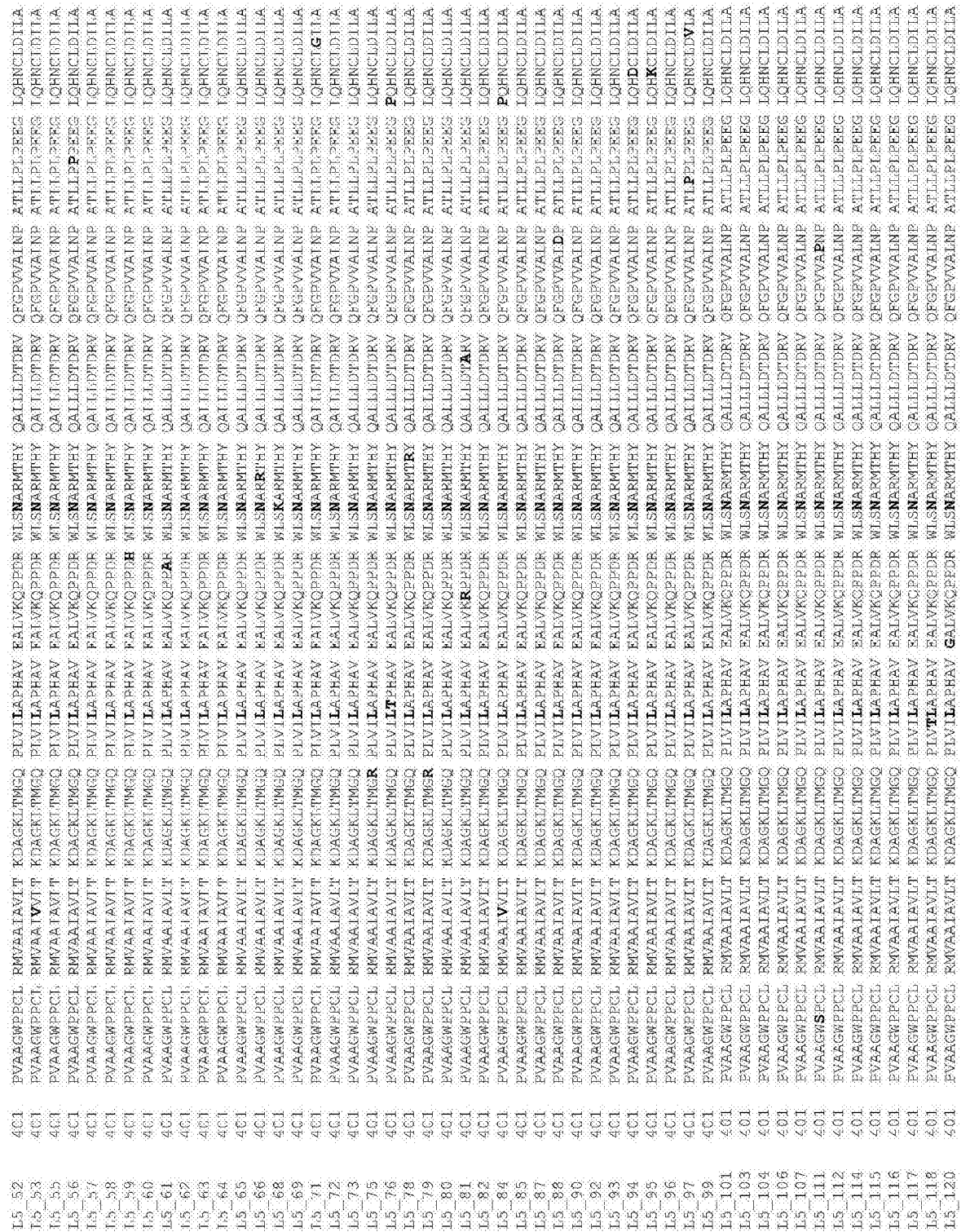
[0388] 附录2续

[0389]



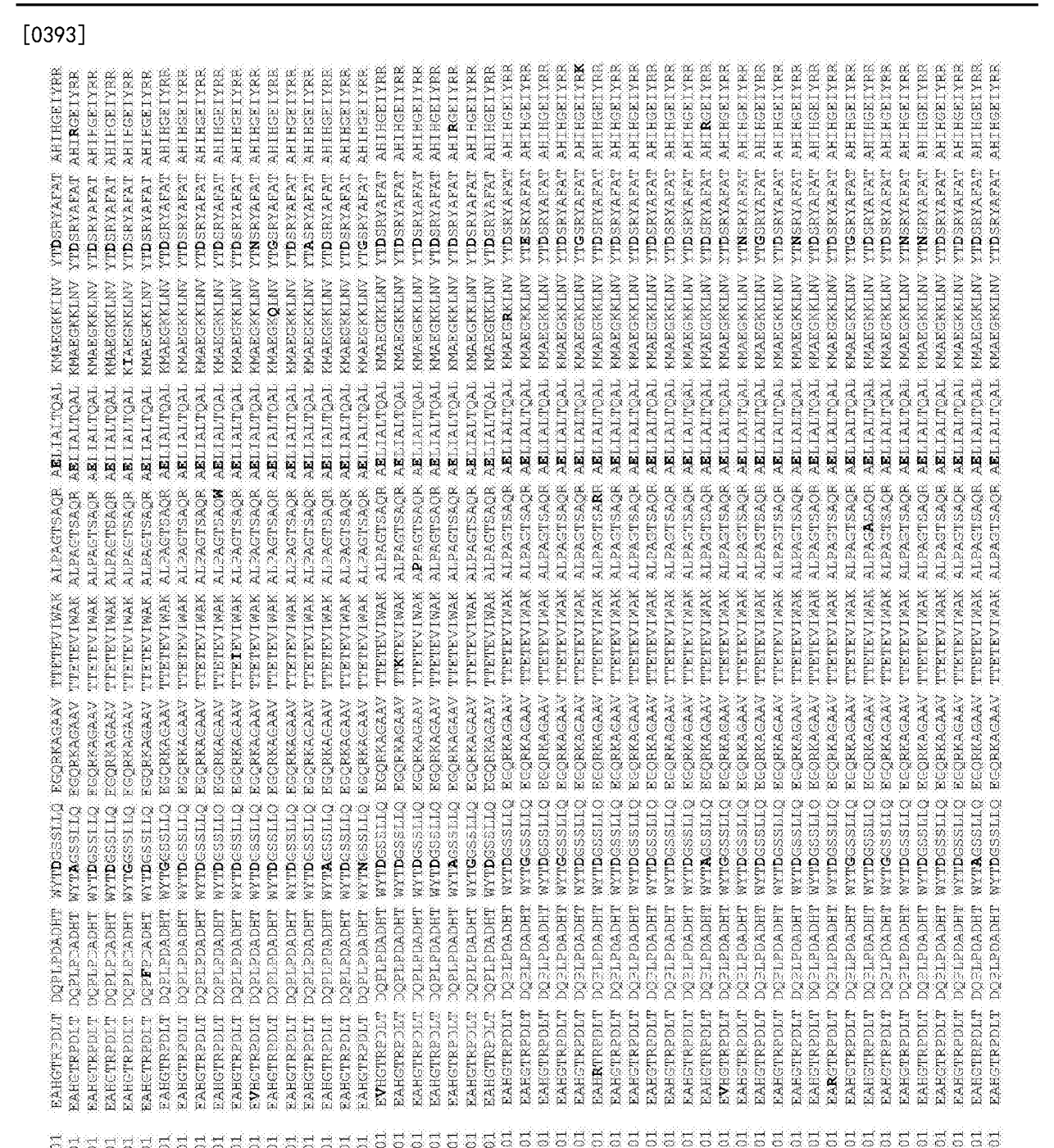
[0390] 附录2续

[0391]



[0392] 附录2续

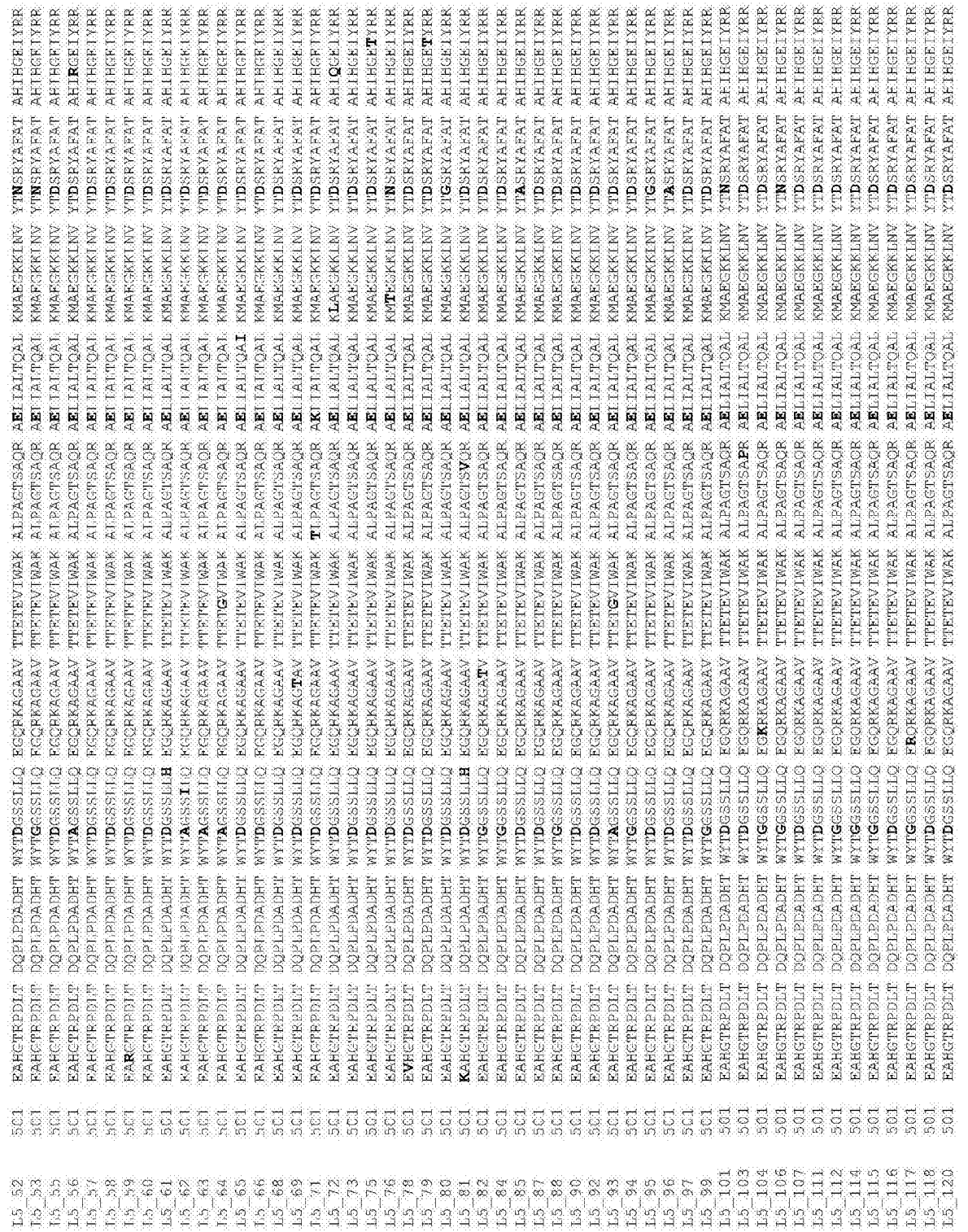
[0393]



| MLV | 501 | EAHGTRPDLT | DQPLPDADHT | WYTDGSSLLQ | EGQRKAGAAV | TTETEVIWAK | ALPAGTSAQR | AELIALTQAL | KMAEGKKLNV | YTDSRYAFAT | AHIHGEIYRR |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 501 | EAHGTRPDLT | DQPLPDADHT | WYTASSILQ | EGQRKAGAAV | TTETEVIWAK | ALPAGTSAQR | AELIALTQAL | KMAEGKKLNV | YTDSRYAFAT | AHIRGEIYRR |
| 3 | 501 | EAHGTRPDLT | DQPLPDADHT | WYTDGSSILQ | EGQRKAGAAV | TTETEVIWAK | ALPAGTSAQR | AELIALTQAL | KMAEGKKLNV | YTDSRYAFAT | AHIHGEIYRR |
| 5 | 501 | EAHGTRPDLT | DQPLPDADHT | WYTDGSSILQ | EGQRKAGAAV | TTETEVIWAK | ALPAGTSAQR | AELIALTQAL | KMAEGKKLNV | YTDSRYAFAT | AHIHGEIYRR |
| 7 | 501 | EAHGTRPDLT | DQPLPDADHT | WYTGGSSILQ | EGQRKAGAAV | TTETEVIWAK | ALPAGTSAQR | AELIALTQAL | KIAEGKKLNV | YTDSRYAFAT | AHIHGEIYRR |
| 8 | 501 | EAHGTRPDLT | DQPFPDADHT | WYTDGSSILQ | EGQRKAGAAV | TTETEVIWAK | ALPAGTSAQR | AELIALTQAL | KMAEGKKLNV | YTDSRYAFAT | AHIHGEIYRR |
| 10 | 501 | EAHGTRPDLT | DQPLPDADHT | WYTGGSSLLQ | EGQRKAGAAV | TTETEVIWAK | ALPAGTSAQR | AELIALTQAL | KMAEGKKLNV | YTDGRYAFAT | AHIHGEIYRR |
| 11 | 501 | EAHGTRPDLT | DQPLPDADHT | WYTDGSSLLQ | EGQRKAGAAV | TTETEVIWAK | ALPAGTSAQR | AELIALTQAL | KMAEGKKLNV | YTDSRYAFAT | AHIHGEIYRR |
| 12 | 501 | EAHGTRPDLT | DQPLPDADHT | WYTDGSSLLQ | EGQRKAGAAV | TTETEVIWAK | ALPAGTSAQR | AELIALTQAL | KMAEGKKLNV | YTDSRYAFAT | AHIHGEIYRR |
| 13 | 501 | EAHGTRPDLT | DQPLPDADHT | WYTDGSSLLQ | EGQRKAGAAV | TTETEVIWAK | ALPAGTSAQW | AELIALTQAL | KMAEGKKLNV | YTDSRYAFAT | AHIHGEIYRR |
| 16 | 501 | EAHGTRPDLT | DQPLPDADHT | WYTDGSSLLQ | EGQRKAGAAV | TTEIEVIWAK | ALPAGTSAQR | AELIALTQAL | KMAEGKKLNV | YTDSRYAFAT | AHIHGEIYRR |
| 17 | 501 | EVHGTRPDLT | DQPLPDADHT | WYTDGSSLLQ | EGQRKAGAAV | TTETEVIWAK | ALPAGTSAQR | AELIALTQAL | KMAEGKKLNV | YTESRYAFAT | AHIHGEIYRR |
| 18 | 501 | EAHGTRPDLT | DQPLPDADHT | WYTDGSSLLQ | EGQRKAGAAV | TTETEVIWAK | ALPAGTSAQR | AELIALTQAL | KMAEGKKLNV | YTDSRYAFAT | AHIHGEIYRR |
| 20 | 501 | EAHGTRPDLT | DQPLPDADHT | WYTDGSSLLQ | EGQRKAGAAV | TTETEVIWAK | ALPAGTSAQR | AELIALTQAL | KMAEGKKLNV | YTDSRYAFAT | AHIHGEIYRR |
| 21 | 501 | EAHGTRPDLT | DQPLPDADHT | WYTDGSSLLQ | EGQRKAGAAV | TTETEVIWAK | ALPAGTSAQR | AELIALTQAL | KMAEGKKLNV | YTASRYAFAT | AHIHGEIYRR |
| 23 | 501 | EAHGTRPDLT | DQPLPDADHT | WYTASSLLQ | EGQRKACAAV | TTETEVIWAK | ALPAGTSAQR | AELIALTQAL | KMAEGKKLNV | YTDSRYAFAT | AHIHGEIYRR |
| 28 | 501 | EAHGTRPDLT | DQPLPDADHT | WYTDGSSLLQ | EGQRKACAAV | TTETEVIWAK | ALPAGTSAQR | AELIALTQAL | KMAEGKKLNV | YTDSRYAFAT | AHIHGEIYRR |
| 30 | 501 | EAHGTRPDLT | DQPLPDADHT | WYTNGSSLLQ | EGQRKACAAV | TTETEVIWAK | ALPAGTSAQR | AELIALTQAL | KMAEGKKLNV | YTGSRYAFAT | AHIHGEIYRR |
| L5_1 | 501 | EVHGTRPDLT | DQPLPDADHT | WYTDGSSLLQ | EGQRKAGAAV | TTETEVIWAK | ALPAGTSAQR | AELIALTQAL | KMAEGKKLNV | YTDSRYAFAT | AHIHGEIYRR |
| L5_2 | 501 | EAHGTRPDLT | DQPLPDADHT | WYTDGSSLLQ | EGQRKAGAAV | TTKTEVIWAK | ALPAGTSAQR | AELIALTQAL | KMAEGKKLNV | YTDSRYAFAT | AHIHGEIYRR |
| L5_3 | 501 | EAHGTRPDLT | DQPLPDADHT | WYTDGSSLLQ | EGQRKAGAAV | TTETEVIWAK | APPAGTSAQR | AELIALTQAL | KMAEGKKLNV | YTDSRYAFAT | AHIHGEIYRR |
| L5_4 | 501 | EAHGTRPDLT | DQPLPDADHT | WYTDGSSLLQ | EGQRKAGAAV | TTPTEVIWAK | ALPAGTSAQR | AELIALTQAL | KMAEGKKLNV | YTDSRYAFAT | AHIHGEIYRR |
| L5_6 | 501 | EAHGTRPDLT | DQPLPDADHT | WYTASSILQ | EGQRKAGAAV | TTETEVIWAK | ALPAGTSAQR | AELIALTQAL | KMAEGKKLNV | YTDSRYAFAT | AHIRGEIYRR |
| L5_8 | 501 | EAHGTRPDLT | DQPLPDADHT | WYTGGSSLLQ | EGQRKAGAAV | TTETEVIWAK | ALPAGTSAQR | AELIALTQAL | KMAEGKKLNV | YTDSRYAFAT | AHIHGEIYRR |
| L5_9 | 501 | EAHGTRPDLT | DQPLPDADHT | WYTDGSSLLQ | EGQRKAGAAV | TTETEVIWAK | ALPAGTSAQR | AELIALTQAL | KMAEGKKLNV | YTDSRYAFAT | AHIHGEIYRR |
| L5_11 | 501 | EAHGTRPDLT | DQPLPDADHT | WYTDGSSLLQ | EGQRKAGAAV | TTETEVIWAK | ALPAGTSAQR | AELIALTQAL | KMAEGRKLNV | YTDSRYAFAT | AHIHGEIYRR |
| L5_13 | 501 | EAHGTRPDLT | DQPLPDADHT | WYTGGSSLLQ | EGQRKAGAAV | TTETEVIWAK | ALPAGTSAQR | AELIALTQAL | KMAEGKKLNV | YTDSRYAFAT | AHIHGEIYRR |
| L5_14 | 501 | EAHGTRPDLT | DQPLPDADHT | WYTDGSSLLQ | EGQRKAGAAV | TTETEVIWAK | ALPAGTSAQR | AELIALTQAL | KMAEGKKLNV | YTDSRYAFAT | AHIHGEIYRR |
| L5_15 | 501 | EAHGTRPDLT | DQPLPDADHT | WYTGGSSLLQ | EGQRKAGAAV | TTETEVIWAK | ALPAGTSAQR | AELIALTQAL | KMAEGKKLNV | YTDSRYAFAT | AHIHGEIYRR |
| L5_16 | 501 | EAHGTRPDLT | DQPLPDADHT | WYTDGSSLLQ | EGQRKAGAAV | TTETEVIWAK | ALPAGTSAQR | AELIALTQAL | KMAEGKKLNV | YTGSRYAFAT | AHIHGEIYRK |
| L5_18 | 501 | EAHRTRPDLT | DQPLPDADHT | WYTDGSSLLQ | EGQRKAGAAV | TTETEVIWAK | ALPAGTSARR | AELIALTQAL | KMAEGKKLNV | YTDSRYAFAT | AHIHGEIYRR |
| L5_20 | 501 | EAHGTRPDLT | DQPLPDADHT | WYTDGSSLLQ | EGQRKAGAAV | TTETEVIWAK | ALPAGTSAQR | AELIALTQAL | KMAEGKKLNV | YTDSRYAFAT | AHIHGEIYRR |
| L5_21 | 501 | EAHGTRPDLT | DQPLPDADHT | WYTDGSSLLQ | EGQRKAGAAV | TTETEVIWAK | ALPAGTSAQR | AELIALTQAL | KMAEGKKLNV | YTDSRYAFAT | AHIHGEIYRR |
| L5_23 | 501 | EAHGTRPDLT | DQPLPDADHT | WYTDGSSLLQ | EGQRKAGAAV | TTETEVIWAK | ALPAGTSAQR | AELIALTQAL | KMAEGKKLNV | YTDSRYAFAT | AHIHGEIYRR |
| L5_24 | 501 | EAHGTRPDLT | DQPLPDADHT | WYTDGSSLLQ | EGQRKAGAAV | TTETEVIWAK | ALPAGTSAQR | AELIALTQAL | KMAEGKKLNV | YTDSRYAFAT | AHIRGEIYRR |
| L5_25 | 501 | EAHGTRPDLT | DQPLPDADHT | WYTDGSSLLQ | EGQRKAGAAV | TTETEVIWAK | ALPAGTSAQR | AELIALTQAL | KMAEGKKLNV | YTDSRYAFAT | AHIHGEIYRR |
| L5_29 | 501 | EVHGTRPDLT | DQPLPDADHT | WYTGGSSLLQ | EGQRKAGAAV | TTETEVIWAK | ALPAGTSAQR | AELIALTQAL | KMAEGKKLNV | YTDSRYAFAT | AHIHGEIYRR |
| L5_30 | 501 | EAHGTRPDLT | DQPLPDADHT | WYTDGSSLLQ | EGQRKAGAAV | TTETEVIWAK | ALPAGTSAQR | AELIALTQAL | KMAEGKKLNV | YTNSRYAFAT | AHIHGEIYRR |
| L5_32 | 501 | EAHGTRPDLT | DQPLPDADHT | WYTDGSSLLQ | EGQRKAGAAV | TTETEVIWAK | ALPAGTSAQR | AELIALTQAL | KMAEGKKLNV | YTGSRYAFAT | AHIHGEIYRR |
| L5_33 | 501 | EAHGTRPDLT | DQPLPDADHT | WYTDGSSLLQ | EGQRKAGAAV | TTETEVIWAK | ALPAGTSAQR | AELIALTQAL | KMAEGKKLNV | YTDSRYAFAT | AHIHGEIYRR |
| L5_37 | 501 | EAHGTRPDLT | DQPLPDADHT | WYTDGSSLLQ | EGQRKAGAAV | TTETEVIWAK | ALPAGTSAQR | AELIALTQAL | KMAEGKKLNV | YTNSRYAFAT | AHIHGEIYRR |
| L5_39 | 501 | EAHGTRPDLT | DQPLPDADHT | WYTDGSSLLQ | EGQRKAGAAV | TTETEVIWAK | ALPAGTSAQR | AELIALTQAL | KMAEGKKLNV | YTDSRYAFAT | AHIHGEIYRR |
| L5_40 | 501 | EARGTRPDLT | DQPLPDADHT | WYTDGSSLLQ | EGQRKAGAAV | TTETEVIWAK | ALPAGTSAQR | AELIALTQAL | KMAEGKKLNV | YTDSRYAFAT | AHIHGEIYRR |
| L5_41 | 501 | EAHGTRPDLT | DQPLPDADHT | WYTGGSSLLQ | EGQRKAGAAV | TTETEVIWAK | ALPAGTSAQR | AELIALTQAL | KMAEGKKLNV | YTGSRYAFAT | AHIHGEIYRR |
| L5_42 | 501 | EAHGTRPDLT | DQPLPDADHT | WYTDGSSLLQ | EGQRKAGAAV | TTETEVIWAK | ALPAGAAQAR | AELIALTQAL | KMAEGKKLNV | YTDSRYAFAT | AHIHGEIYRR |
| L5_43 | 501 | EAHGTRPDLT | DQPLPDADHT | WYTGGSSLLQ | EGQRKAGAAV | TTETEVIWAK | ALPAGTSAQR | AELIALTQAL | KMAEGKKLNV | YTDSRYAFAT | AHIHGEIYRR |
| L5_44 | 501 | EAHGTRPDLT | DQPLPDADHT | WYTDGSSLLQ | EGQRKAGAAV | TTETEVIWAK | ALPAGTSAQR | AELIALTQAL | KMAEGKKLNV | YTNSRYAFAT | AHIHGEIYRR |
| L5_46 | 501 | EAHGTRPDLT | DQPLPDADHT | WYTDGSSLLQ | EGQRKAGAAV | TTETEVIWAK | ALPAGTSAQR | AELIALTQAL | KMAEGKKLNV | YTNSRYAFAT | AHIHGEIYRR |
| L5_47 | 501 | EAHGTRPDLT | DQPLPDADHT | WYTDGSSLLQ | EGQRKAGAAV | TTETEVIWAK | ALPAGTSAQR | AELIALTQAL | KMAEGKKLNV | YTDSRYAFAT | AHIHGEIYRR |
| L5_48 | 501 | EAHGTRPDLT | DQPLPDADHT | WYTDGSSLLQ | EGQRKAGAAV | TTETEVIWAK | ALPAGTSAQR | AELIALTQAL | KMAEGKKLNV | YTDSRYAFAT | AHIHGEIYRR |
| L5_49 | 501 | EAHGTRPDLT | DQPLPDADHT | WYTASSLLQ | EGQRKAGAAV | TTETEVIWAK | ALPAGTSAQR | AELIALTQAL | KMAEGKKLNV | YTDSRYAFAT | AHIHGEIYRR |
| L5_51 | 501 | EAHGTRPDLT | DQPLPDADHT | WYTDGSSLLQ | EGQRKAGAAV | TTETEVIWAK | ALPAGTSAQR | AELIALTQAL | KMAEGKKLNV | YTDSRYAFAT | AHIHGEIYRR |

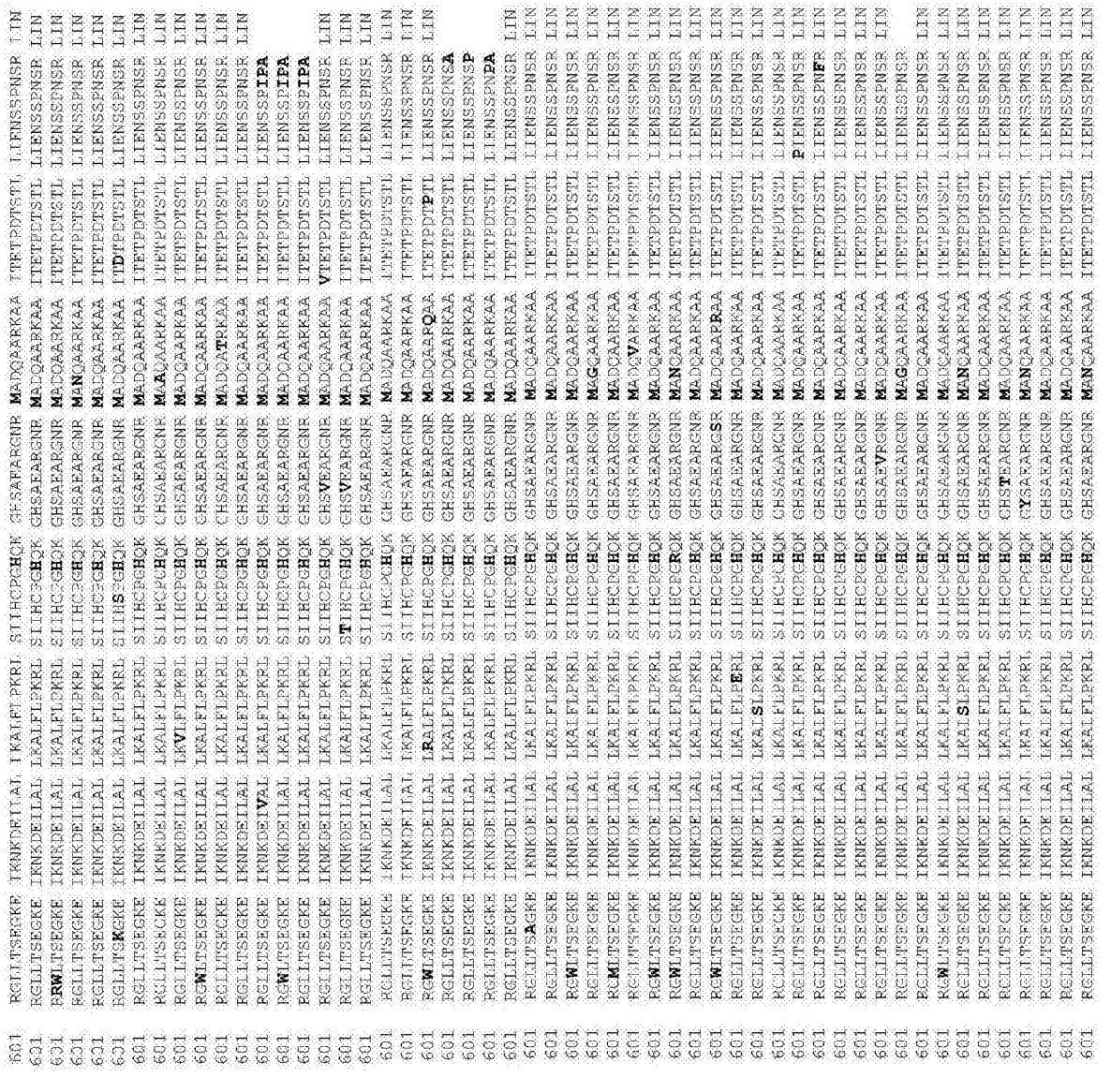
[0394] 附录2续

[0395]



[0396] 附录2续

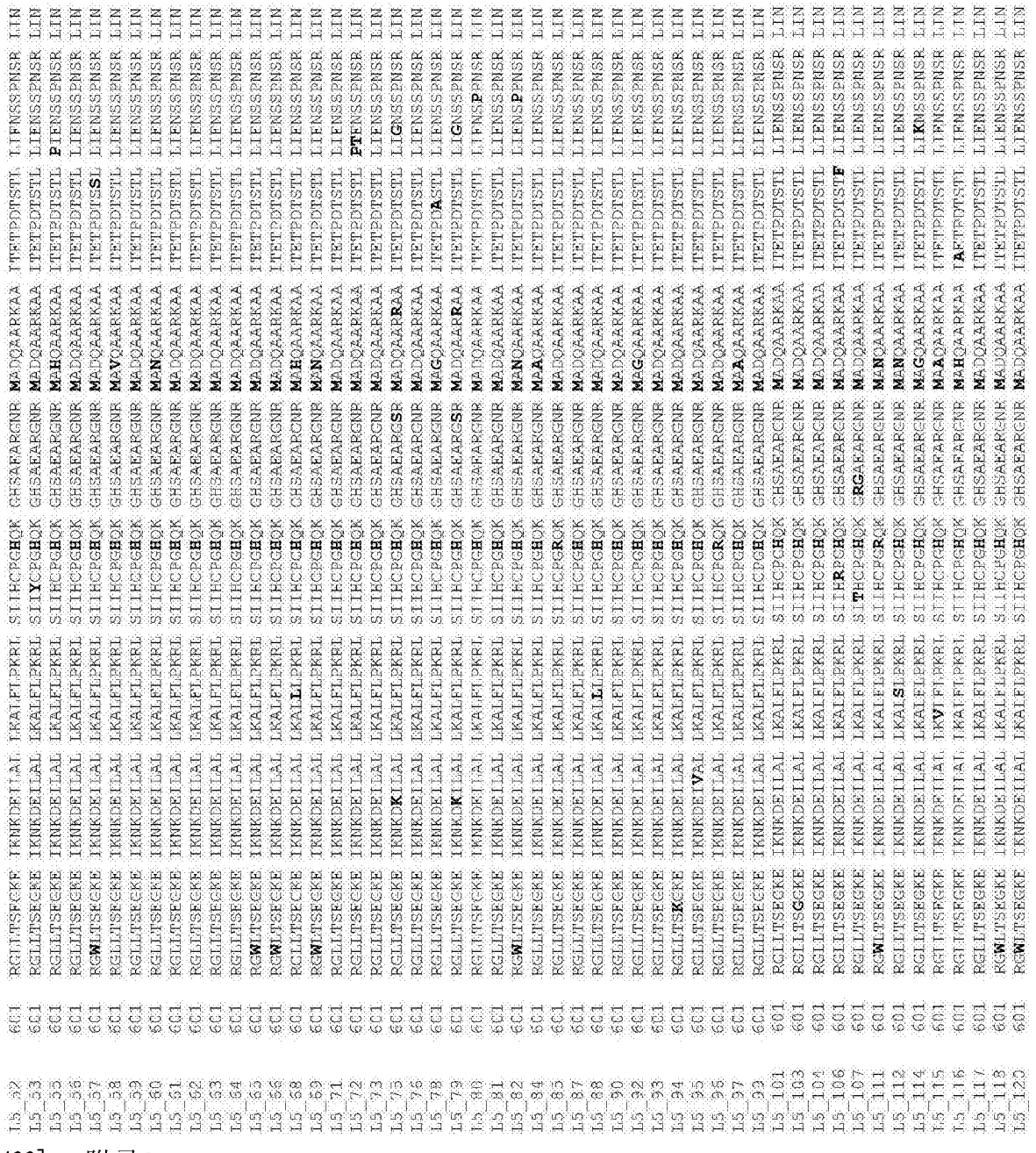
[0397]



| MTV | 601 | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 601 | RGLLTSEGKE | IKNKDEILAL | LKALFLPKRL | SIIHCPGHQK | GHSAEARGNR | MADQAARKAA | ITETPDTSTL | LIENSSPNSR | LIN | | | | |
| 2 | 601 | RGLLTSEGKE | IKNKDEILAL | LKALFLDKRL | SIIHCPGHQK | GHSAEARGNR | MADQAARKAA | ITETPDTSTL | LIENSSPNSR | LIN | | | | |
| 3 | 601 | RRWLTSEGKE | IKNKDEILAL | LKALFLPKRL | SIIHCPGHQK | GHSAEARGNR | MADQAARKAA | ITETPDTSTL | LIENSSPNSR | LIN | | | | |
| 4 | 601 | RGLLTSEGKE | IKNKDEILAL | LKALFLPKRL | SIIHCPGHQR | GHSAEARGNR | MANQAARKAA | ITETPDTSTL | LIENSSPNSR | LIN | | | | |
| 5 | 601 | RGLLTSEGKE | IRNKDEILAL | LKALFLPKRL | SIIHCPGHQK | GHSAEARGNR | MADQAARKAA | ITETPDTSTL | LIENSSPNSR | LIN | | | | |
| 6 | 601 | RGLLTSKGKE | IKNKDEILAL | LKALFLPKRL | SIIHCPGHQK | GHSAEARGNR | MADQAARKAA | ITDTPDTSTL | LIENSSPNSR | LIN | | | | |
| 7 | 601 | RGLLTSEGKE | IKNKDEILAL | LKALFLPKRL | SIIHCPGHQK | GHSAEARGNR | MADQAARKAA | ITETPDTSTL | LIENSSPNSR | LIN | | | | |
| 8 | 601 | RGLLTSEGKE | IKNKDEILAL | LKALFLPKRL | SIIHCPGHQK | GHSAEARGNR | MADQAATRKAA | ITETPDTSTL | LIENSSPNSR | LIN | | | | |
| 9 | 601 | RGLLTSEGKE | IKNKDEILAL | LKALFLPKRL | SIIHCPGHQK | GHSAEARGNR | MADQAARKAA | ITETPDTSTL | LIENSSPNSR | LIN | | | | |
| 10 | 601 | RGLLTSEGKE | IKNKDEILAL | LKALFLPKRL | SIIHCPGHQK | GHSAEARGNR | MADQAARKAA | ITETPDTSTL | LIENSSPNSR | LIN | | | | |
| 11 | 601 | RGLLTSEGKE | IKNKDEILAL | LKALFLPKRL | SIIHCPGHQK | GHSAEARGNR | MAAQAARKAA | ITETPDTSTL | LIENSSPNSR | LIN | | | | |
| 12 | 601 | RGLLTSEGKE | IKNKDEILAL | LKVLFLPKRL | SIIHCPGHQK | GHSAEARGNR | MADQAARKAA | ITETPDTSTL | LIENSSPIPA | | | | | |
| 13 | 601 | RGWLTSEGKE | IKNKDEILAL | LKALFLPKRL | SIIHCPGHQK | GHSAEARGNR | MADQAARKAA | ITETPDTSTL | LIENSSPIPA | | | | | |
| 14 | 601 | RGLLTSEGKE | IKNKDEILAL | LKALFLPKRL | SIIHCPGHQK | GHSAEARGNR | MADQAARKAA | ITETPDTSTL | LIENSSPNSR | LIN | | | | |
| 15 | 601 | RGLLTSEGKE | IKNKDEILAL | LKALFLPKRL | SIIHCPGHQK | GHSAEARGNR | MADQAARKAA | ITETPDTSTL | LIENSSPNSR | LIN | | | | |
| 16 | 601 | RGLLTSEGKE | IKNKDEILAL | LKALFLPKRL | SIIHCPGHQK | GHSAEARGNR | MADQAARKAA | ITETPDTSTL | LIENSSPNSR | LIN | | | | |
| 17 | 601 | RGLLTSEGKE | IKNKDEILAL | LKALFLPKRL | SIIHCPGHQK | GHSAEARGNR | MADQAARKAA | ITETPDTSTL | LIENSSPNSR | LIN | | | | |
| 18 | 601 | RGLLTSEGKE | IKNKDEILAL | LKALFLPKRL | SIIHCPGHQK | GHSAEARGNR | MADQAARKAA | ITETPDTSTL | LIENSSPIPA | | | | | |
| 19 | 601 | RGLLTSEGKE | IKNKDEILAL | LKALFLPKRL | SIIHCPGHQK | GHSAEARGNR | MAGQAARKAA | ITETPDTSTL | LIENSSPNSR | LIN | | | | |
| 20 | 601 | RGLLTSEGKE | IKNKDEILAL | LKALFLPKRL | SIIHCPGHQK | GHSAEARGNR | MADQAARKAA | ITETPDTSTL | LIENSSPNSR | LIN | | | | |
| 21 | 601 | RGLLTSEGKE | IKNKDEILAL | LKALFLPKRL | SIIHCPGHQK | GHSAEARGNR | MADQAARKAA | ITETPDTSTL | LIENSSPIPA | | | | | |
| 22 | 601 | RGLLTSEGKE | IKNKDEILAL | LKALFLPKRL | SIIHCPGHQK | GHSVEARGNR | MADQAARKAA | VTETPDTSTL | LIENSSPNSR | LIN | | | | |
| 23 | 601 | RGLLTSEGKE | IKNKDEILAL | LKALFLPKRL | STIHCPGHQK | GHSVEARGNR | MADQAARKAA | ITETPDTSTL | LIENSSPNSR | LIN | | | | |
| 24 | 601 | RGLLTSEGKE | IKNKDEILAL | LKALFLPKRL | SIIHCPGHQK | GHSVEARGNR | MADQAARKAA | ITETPDTSTL | LIENSSPNSR | LIN | | | | |
| 25 | 601 | RGLLTSEGKE | IKNKDEILAL | LKALFLPKRL | SIIHCPGHQK | GHSAEARGNR | MADQAARKAA | ITETPDTSTL | LIENSSPNSA | | | | | |
| 26 | 601 | RGLLTSEGKE | IKNKDEILAL | LKALFLPKRL | SIIHCPGHQR | GHSAEARGNR | MADQAARKAA | ITETPDTSTL | LIENSSPNSP | | | | | |
| 27 | 601 | RGLLTSEGKE | IKNKDEILAL | LKALFLPKRL | SIIHCPGHQK | GHSAEARGNR | MADQAARKAA | ITETPDTSTL | LIENSSPNPA | | | | | |
| 28 | 601 | RGLLTSEGKE | IKNKDEILAL | LKALFLPKRL | SIIHCPGHQK | GHSAEARGNR | MADQAARKAA | ITETPDTSTL | LIENSSPNSR | LIN | | | | |
| 29 | 601 | RGLLTSEGKE | IKNKDEILAL | LKALFLPKRL | SIIHCPGHQK | GHSAEARGNR | MADQAARKAA | ITETPDTSTL | LIENSSPNSR | LIN | | | | |
| 30 | 601 | RGLLTSEGKE | IKNKDEILAL | LKALFLPKRL | SIIHCPGHQK | GHSAEARGNR | MADQAARKAA | ITETPDTSTL | LIENSSPNSR | LIN | | | | |
| L5_1 | 601 | RGLLTSEGKE | IKNKDEILAL | LKALFLPKRL | SIIHCPGHQK | GHSAEARGNR | MADQAARKAA | ITETPDTSTL | LIENSSPNSR | LIN | | | | |
| L5_2 | 601 | RGLLTSPGKE | IKNKDFILAL | IKALFLPKRL | SIIHCPGHQK | GHSAFARGNR | MADQAARKAA | ITFTPDTSTL | LIFNSSPNSR | LIN | | | | |
| L5_3 | 601 | RGWLTSEGKE | IRNKDEILAL | LRALFLPKRL | SIIHCPGHQK | GHSAEARGNR | MADQAARQAA | ITETPDTPTL | LIENSSENSR | LIN | | | | |
| L5_4 | 601 | RGLLTSEGKE | IKNKDEILAL | LKALFLDKRL | SIIHCPGHQK | GHSAEARGNR | MADQAARKAA | ITETPDTSTL | LIENSSPNSA | | | | | |
| L5_6 | 601 | RGLLTSEGKE | IKNKDEILAL | LKALFLPKRL | SIIHCPGHQR | GHSAEARGNR | MADQAARKAA | ITETPDTSTL | LIENSSPNSP | | | | | |
| L5_8 | 601 | RGLLTSEGKE | IKNKDEILAL | LKALFLPKRL | SIIHCPGHQK | GHSAEARGNR | MADQAARKAA | ITETPDTSTL | LIENSSPNPA | | | | | |
| L5_9 | 601 | RGLLTSEGKE | IKNKDEILAL | LKALFLPKRL | SIIHCPGHQK | GHSAEARGNR | MADQAARKAA | ITETPDTSTL | LIENSSPNSR | LIN | | | | |
| L5_11 | 601 | RGLLTSAGKE | IKNKDEILAL | LKALFLPKRL | SIIHCPGHQK | GHSAFARGNR | MADQAARKAA | ITETPDTSTL | LIENSSPNSR | LIN | | | | |
| L5_13 | 601 | RGLLTSEGKE | IKNKDEILAL | LKALFLPKRL | SIIHCPGHQK | GHSAEARGNR | MADQAARKAA | ITETPDTSTL | LIENSSPNSR | LIN | | | | |
| L5_14 | 601 | RGWLTSEGKE | IKNKDEILAL | LKALFLPKRL | SIIHCPGHQK | GHSAEARGNR | MADQAARKAA | ITETPDTSTL | LIENSSPNSR | LIN | | | | |
| L5_15 | 601 | RGLLTSPGKE | IKNKDEILAL | LKALFLPKRL | SIIHCPGHQK | GHSAFARGNR | MAGQAARKAA | ITETPDTSTL | LIENSSPNSR | LIN | | | | |
| L5_16 | 601 | RGMLTSEGKE | IKNKDEILAL | LKALFLPKRL | SIIHCPGHQK | GHSAEARGNR | MADQAARKAA | ITETPDTSTL | LIENSSPNSR | LIN | | | | |
| L5_18 | 601 | RGLLTSPGKE | IKNKDFILAL | IKALFLPKRL | SIIHCPGHQK | GHSAFARGNR | MADQVARKAA | ITFTPDTSTL | LIFNSSPNSR | LIN | | | | |
| L5_20 | 601 | RGWLTSEGKE | IKNKDEILAL | LKALFLPKRL | SIIHCPGHQK | GHSAEARGNR | MADQAARKAA | ITETPDTSTL | LIENSSPNSR | LIN | | | | |
| L5_21 | 601 | RGWLTSEGKE | IKNKDEILAL | LKALFLPKFL | SIIHCPGHQK | GHSAEARGNR | MANQAARKAA | ITETPDTSTL | LIENSSPNSR | LIN | | | | |
| L5_23 | 601 | RGLLTSEGKE | IKNKDEILAL | LKALFLPKRL | SIIHCPGHQK | GHSAEARGNR | MADQAARKAA | ITETPDTSTL | LIENSSPNSR | LIN | | | | |
| L5_24 | 601 | RGWLTSPGKE | IKNKDFILAL | LKALFLPKRL | SIIHCPGHQK | GHSAFARGSR | MADQAAPRAA | ITFTPDTSTL | LIFNSSPNSR | LIN | | | | |
| L5_25 | 601 | RGLLTSPGKE | IKNKDFILAL | LKALFLERRL | SIIHCPGHQK | GHSAFARGNR | MADQAARKAA | ITFTPDTSTL | LIFNSSPNSR | LIN | | | | |
| L5_28 | 601 | RGLLTSEGKE | IKNKDEILAL | LKALSLPKRL | SIIHCPGHQK | GHSAEARGNR | MADQAARKAA | ITETPDTSTL | LIENSSPNSR | LIN | | | | |
| L5_29 | 601 | RGLLTSPGKE | IKNKDFILAL | LKALFLPKRL | SIIHCPGHQK | GHSAFARGNR | MADQAARKAA | ITFTPDTSTL | LIFNSSPNSR | LIN | | | | |
| L5_33 | 601 | RGLLTSEGKE | IKNKDEILAL | LKALFLPKRL | SIIHCPGHQK | GHSAEARGNR | MADQAARKAA | ITETPDTSTL | PIENSSPNSR | LIN | | | | |
| L5_34 | 601 | RGLLTSEGKE | IKNKDEILAL | LRALFLPKRL | SIIHCPGHQK | GHSAEARGNR | MADQAARKAA | ITETPDTSTL | LIENSSPNFR | LIN | | | | |
| L5_35 | 601 | RGLLTSEGKE | IKNKDEILAL | LKALFLPKFL | SIIHCPGHQK | GHSAEARGNR | MADQAARKAA | ITETPDTSTL | LIENSSPNSP | LIN | | | | |
| L5_37 | 601 | RGLLTSEGKE | IKNKDEILAL | LKALFLPKRL | SIIHCPGHQK | GHSAEARGNR | MADQAARKAA | ITETPDTSTL | LIENSSPNSR | LIN | | | | |
| L5_39 | 601 | RGLLTSPGKE | IKNKDFILAL | LKALFLPKRL | SIIHCPGHQK | GHSAFVRGNR | MADQAARKAA | ITFTPDTSTL | LIFNSSPNSR | LIN | | | | |
| L5_40 | 601 | RGLLTSPGKE | IKNKDFILAL | LKALFLPKRL | SIIHCPGHQK | GHSAFARGND | MAGQAARKAA | ITFTPDTSTL | LIFNSSPNSR | | | | | |
| L5_41 | 601 | RGLLTSEGKE | IKNKDEILAL | LKALFLPKRL | SIIHCPGHQK | GHSAEARGNR | MADQAARKAA | ITETPDTSTL | LIENSSPNSR | LIN | | | | |
| L5_42 | 601 | RGWLTSKGKE | IKNKDEILAL | LKALFLPKRL | SIIHCPGHQK | GHSAEARGNR | MADQAARKAA | ITETPDTSTL | LIENSSPNSR | LIN | | | | |
| L5_43 | 601 | RGLLTSEGKE | IKNKDEILAL | LKALSLPKRL | SIIHCPGHQK | GHSAEARGNR | MANQAARKAA | ITETPDTSTL | LIENSSPNSR | LIN | | | | |
| L5_44 | 601 | RGLLTSEGKE | IKNKDEILAL | LKALFLPKRL | SIIHCPGHQK | GHSAEARGNR | MADQAARKAA | ITETPDTSTL | LIENSSPNSP | LIN | | | | |
| L5_46 | 601 | RGLLTSPGKE | IKNKDFILAL | LKALFLPKFL | SIIHCPGHQK | GHSTEARGNR | MAGQAARKAA | ITFTPDTSTL | LIFNSSPNSR | LIN | | | | |
| L5_47 | 601 | RGLLTSPGKE | IKNKDFILAL | IKALFLPKRL | SIIHCPGHQK | GYSAEARGNR | MANQAARKAA | ITFTPDTSTL | LIFNSSPNSR | LIN | | | | |
| L5_48 | 601 | RGLLTSPGKE | IKNKDFILAL | LKALFLPKRL | SIIHCPGHQK | GHSAEARGNP | MADQAARKAA | ITFTPDTSTL | LIFNSSPNSR | LIN | | | | |
| L5_49 | 601 | RGLLTSEGKE | IKNKDEILAL | LKALFLDKRL | SIIHCPGHQK | GHSAEARGNR | MADQAARKAA | ITETPDTSTL | LIENSSPNSR | LIN | | | | |
| L5_51 | 601 | RGLLTSPGKE | IKNKDEILAL | LRALFLPKRL | SIIHCPGHQK | GESAEARGNR | MANQAARKAA | ITETPDTSTL | LIENSSPNSR | LIN | | | | |

[0398] 附录2续

[0399]



[0400] 附录3

[0401]

| | | |
|---|---|---|
| L5_9 | 12 个突变, | 83D->N, 135L->P, 166P->S, 214H->R, 222Y->C, 293T->A, 344Y->H, 407P->L, 415A->V, 436A->T, 444V->A, 447P->L |
| 18 | 11 个突变, | 83D->N, 139L->P, 200D->N, 330T->P, 479N->D, 577K->Q, 583D->G, 618L->V, 678N->I, 679S->P, 680R->A |
| L5_79 | 10 个突变, | 74I->T, 104P->R, 325Y->H, 333L->Q, 430Q->R, 597I->T, 616E->K, 649N->S, 658K->R, 673E->G |
| L5_82 | 10 个突变, | 26L->P, 130P->S, 137S->G, 343A->T, 356A->G, 524D->G, 539A->T, 603L->W, 653D->N, 676S->P |
| L5_117 | 10 个突变, | 14S->P, 95N->S, 139L->P, 190Q->R, 333L->Q, 339D->G, 380V->A, 383Q->P, 524D->G, 532G->R |
| 21 | 9 个突变, | 131N->S, 179I->T, 204H->R, 323P->L, 353T->A, 583D->A, 678N->I, 679S->P, 680R->A |
| L5_3 | 9 个突变, | 125I->V, 138G->R, 143H->R, 380V->A, 552L->P, 603L->W, 622K->R, 658K->Q, 668S->P |
| L5_20 | 9 个突变, | 12E->A, 87L->P, 95N->S, 200D->A, 221Q->R, 242A->T, 331G->E, 428M->L, 603L->W |
| L5_75 | 9 个突变, | 11H->Y, 325Y->H, 333L->Q, 430Q->R, 597I->T, 616E->K, 649N->S, 658K->R, 673E->G |
| L5_76 | 9 个突变, | 66M->L, 105G->E, 200D->N, 289M->V, 314I->T, 436A->T, 491L->P, 573A->T, 583D->N |
| 8 | 8 个突变, | 89P->S, 139L->P, 287T->A, 330T->P, 514L->F, 607E->K, 635C->S, 663E->D |
| L5_13 | 8 个突变, | 130P->S, 159R->K, 200D->A, 221Q->R, 330T->P, 449D->G, 524D->G, 583D->E |
| L5_41 | 8 个突变, | 60S->A, 139L->P, 168F->S, 199F->Y, 346E->D, 450R->H, 524D->G, 583D->G |
| L5_68 | 8 个突变, | 39M->V, 49I->V, 91Q->R, 204H->R, 287T->A, 454N->K, 625F->L, 653D->H |
| L5_72 | 8 个突变, | 139L->P, 200D->N, 330T->P, 393A->T, 572M->L, 594H->Q, 671L->P, 672I->T |
| L5_114 | 8 个突变, | 17P->S, 108D->E, 249N->D, 307A->V, 344Y->H, 524D->G, 653D->G, 673E->K |
| L5_115 | 8 个突变, | 8H->P, 139L->P, 197T->A, 200D->G, 358G->W, 524D->G, 623A->V, 653D->A |
| 17 | 7 个突变, | 49I->V, 65P->S, 200D->N, 409C->R, 470V->A, 502A->V, 583D->N |
| L5_15 | 7 个突变, | 17P->S, 108D->E, 249N->D, 307A->V, 344Y->H, 524D->G, 653D->G |
| L5_21, L5_111 | 7 个突变, | 65P->S, 233E->K, 407P->S, 478L->P, 603L->W, 638H->R, 653D->N |
| L5_40 | 7 个突变, | 24T->A, 41L->R, 127P->S, 151L->F, 330T->P, 503H->R, 633D->G |
| L5_47 | 7 个突变, | 107N->S, 126H->R, 128T->A, 179I->V, 200D->N, 642H->Y, 653D->N |
| L5_71 | 7 个突变, | 97P->S, 184T->A, 250L->P, 289M->L, 497D->G, 551A->T, 562E->K |
| L5_78 | 7 个突变, | 11H->R, 148V->M, 330T->P, 459H->R, 502A->V, 653D->G, 667T->A |
| L5_81 | 7 个突变, | 166P->S, 199F->L, 446Q->R, 468D->A, 501E->K, 530Q->H, 558A->V |
| L5_118 | 7 个突变, | 221Q->R, 283A->T, 287T->A, 369F->I, 376Y->C, 434I->T, 603L->W |
| 13 | 6 个突变, | 51P->S, 136L->W, 207L->P, 428M->L, 560R->W, 603L->W |
| L5_30 | 6 个突变, | 200D->A, 330T->P, 449D->G, 479N->D, 583D->N, 671L->P |
| L5_32 | 6 个突变, | 69E->G, 135L->P, 139L->P, 431P->Q, 583D->G, 679S->F |
| L5_43, L5_112 | 6 个突变, | 200D->N, 237Q->R, 330T->P, 524D->G, 625F->S, 653D->N |
| L5_53 | 6 个突变, | 130P->S, 139L->P, 417A->V, 524D->G, 583D->N, 634H->Y |
| L5_56 | 6 个突变, | 8H->P, 211R->W, 292P->L, 486L->P, 524D->A, 594H->R |
| L5_57 | 6 个突变, | 49I->V, 173R->K, 302E->K, 392V->A, 603L->W, 669T->S |
| L5_60 | 6 个突变, | 112V->A, 200D->N, 280L->P, 322A->T, 379G->S, 653D->N |
| L5_62 | 6 个突变, | 118V->A, 204H->R, 282E->G, 346E->D, 524D->A, 528L->I |
| L5_84 | 6 个突变, | 126H->R, 139L->P, 417A->V, 491L->P, 524D->G, 653D->A |
| L5_95 | 6 个突变, | 124D->G, 187R->G, 263Q->R, 494N->K, 583D->G, 618L->V |
| L5_107 | 6 个突变, | 139L->P, 233E->K, 295K->E, 633I->T, 642H->R, 643S->G |

[0402] 附录3续

| | | |
|---|---|---|
| 16 | 5 个突变,, | 241R->Q, 259A->T, 311R->H, 544T->I, 656A->T |
| 20 | 5 个突变,, | 200D->N, 603L->W, 678N->I, 679S->P, 680R->A |
| 23 | 5 个突变,, | 221Q->R, 332T->I, 524D->A, 644A->V, 661I->V |
| L5_2 | 5 个突变,, | 5E->K, 43V->I, 184T->A, 391P->L, 543E->K |
| L5_14 | 5 个突变,, | 37G->W, 197T->A, 200D->G, 433V->A, 603L->W |
| L5_18 | 5 个突变,, | 457M->T, 462A->T, 504G->R, 559Q->R, 655A->V |
| L5_23 | 5 个突变,, | 30P->L, 223V->M, 388W->R, 390R->W, 435L->P |
| L5_24 | 5 个突变,, | 139L->P, 221Q->R, 603L->W, 649N->S, 658K->R |
| L5_28 | 5 个突变,, | 306T->A, 309F->S, 524D->A, 594H->R, 625F->S |
| L5_37 | 5 个突变,, | 126H->S, 149L->F, 200D->N, 454N->K, 583D->N |
| L5_49, L5_63 | 5 个突变,, | 50I->V, 194N->S, 204H->R, 287T->A, 524D->A |
| L5_55 | 5 个突变,, | 5E->K, 200D->N, 240T->A, 653D->H, 671L->P |
| L5_69 | 5 个突变,, | 14S->T, 49I->T, 538A->T, 603L->W, 653D->N |
| L5_88 | 5 个突变,, | 70A->V, 139L->P, 479N->D, 524D->G, 625F->L |
| L5_101 | 5 个突变,, | 65P->S, 204H->R, 283A->D, 391P->S, 583D->N |
| L5_104 | 5 个突变,, | 132P->S, 164S->G, 388W->R, 524D->G, 533Q->K |
| 3 | 4 个突变,, | 221Q->R, 428M->L, 602G->R, 603L->W |
| 5 | 4 个突变,, | 139L->P, 283A->D, 358G->V, 653D->N |
| 7 | 4 个突变,, | 204H->R, 433V->A, 524D->G, 572M->I |
| 11 | 4 个突变,, | 29F->L, 148V->M, 390R->W, 653D->A |
| 12 | 4 个突变,, | 23S->P, 88V->A, 237Q->R, 623A->V |
| 30 | 4 个突变,, | 200D->N, 306T->M, 524D->N, 583D->G |
| L5_1 | 4 个突变,, | 249N->D, 409C->R, 470V->A, 502A->V |
| L5_6 | 4 个突变,, | 287T->A, 524D->A, 594H->R, 680R->P |
| L5_8 | 4 个突变,, | 475V->G, 524D->G, 679S->P, 680R->A |
| L5_11 | 4 个突变,, | 15K->T, 200D->N, 576K->R, 607E->A |
| L5_16 | 4 个突变,, | 91Q->L, 583D->G, 600R->K, 603L->M |
| L5_29 | 4 个突变,, | 174D->G, 312L->P, 502A->V, 524D->G |
| L5_39 | 4 个突变,, | 12E->V, 86I->V, 200D->N, 646A->V |
| L5_42 | 4 个突变,, | 192F->L, 333L->P, 556T->A, 603L->W |
| L5_44 | 4 个突变,, | 92S->P, 430Q->R, 479N->D, 583D->N |
| L5_46 | 4 个突变,, | 200D->N, 330T->P, 583D->N, 644A->T |
| L5_52 | 4 个突变,, | 200D->N, 330T->P, 374Q->R, 583D->N |
| L5_61 | 4 个突变,, | 64Y->C, 351L->V, 449D->A, 530Q->H |
| L5_64, L5_93 | 4 个突变,, | 200D->N, 216D->G, 524D->A, 545E->G |
| L5_65 | 4 个突变,, | 200D->N, 238Q->H, 570L->I, 603L->W |
| L5_85, L5_96 | 4 个突变,, | 200D->N, 330T->P, 583D->A, 638H->R |
| L5_90 | 4 个突变,, | 200D->N, 298R->G, 330T->P, 374Q->R |
| L5_103 | 4 个突变,, | 96T->M, 200D->N, 559Q->P, 607E->G |
| L5_106 | 4 个突变,, | 524D->G, 583D->N, 635C->R, 670L->F |
| L5_120 | 4 个突变,, | 252Y->H, 308G->S, 441E->G, 603L->W |
| 1 | 3 个突变,, | 49I->V, 524D->A, 594H->R |
| 28 | 3 个突变,, | 8H->R, 632I->T, 644A->V |
| L5_4 | 3 个突变,, | 208A->V, 225D->G, 680R->A |
| L5_25 | 3 个突变,, | 39M->L, 302E->K, 628K->E |
| L5_35 | 3 个突变,, | 200D->N, 330T->P, 479N->D |
| L5_51 | 3 个突变,, | 110R->G, 431P->Q, 653D->N |
| L5_58 | 3 个突变,, | 66M->L, 90C->Y, 653D->V |
| L5_66 | 3 个突变,, | 93P->L, 457M->R, 603L->W |
| L5_73 | 3 个突变,, | 67S->P, 139L->P, 307A->V |
| L5_80 | 3 个突变,, | 326P->S, 583D->G, 676S->P |
| L5_92 | 3 个突变,, | 126H->S, 200D->N, 653D->G |
| L5_94 | 3 个突变,, | 494N->D, 524D->G, 607E->K |
| L5_97 | 3 个突变,, | 484L->P, 498I->V, 653D->A |
| 10 | 2 个突变,, | 481A->T, 524D->G |
| L5_59 | 2 个突变,, | 450R->H, 503H->R |
| L5_99 | 2 个突变,, | 36T->I, 524D->G |
| L5_116 | 2 个突变,, | 653D->H, 662T->A |
| L5_48 | 1 个突变,, | 77H->R |
| WT_MLV, L5_87 | 无突变, | |

[0403]

[0404] 附录4

[0405] 位置524,总突变31,3个不同的氨基酸

[0406] D->G,20个序列(7,10,L5_8,L5_13,L5_15,L5_29,L5_41,L5_43,L5_53,L5_82,L5_84,L5_88,L5_94,L5_99,L5_104,L5_106,L5_112,L5_114,L5_115,L5_117)

[0407] D->A,10个序列(1,23,L5_6,L5_28,L5_49,L5_56,L5_62,L5_63,L5_64,L5_93)

[0408] D->N,1个序列(30)

[0409] 位置200,总突变30,3个不同的氨基酸

[0410] D->G,2个序列(L5_14,L5_115)

[0411] D->A,3个序列(L5_13,L5_20,L5_30)

[0412] D->N,25个序列(17,18,20,30,L5_11,L5_35,L5_37,L5_39,L5_43,L5_46,L5_47,L5_52,

[0413] L5_55,L5_60,L5_64,L5_65,L5_72,L5_76,L5_85,L5_90,L5_92,L5_93,L5_96,L5_103,L5_112)

[0414] 位置653,总突变23,5个不同的氨基酸

[0415] D->V,1个序列(L5_58)

[0416] D->G,5个序列(L5_15,L5_40,L5_78,L5_92,L5_114)

[0417] D->A,4个序列(11,L5_84,L5_97,L5_115)

[0418] D->N,10个序列(5,L5_21,L5_43,L5_47,L5_51,L5_60,L5_69,L5_82,L5_111,L5_112)

[0419] D->H,3个序列(L5_55,L5_68,L5_116)

[0420] 位置583,总突变21,4个不同的氨基酸

[0421] D->E,1个序列(L5_13)

[0422] D->G,7个序列(18,30,L5_16,L5_32,L5_41,L5_80,L5_95)

[0423] D->A,3个序列(21,L5_85,L5_96)

[0424] D->N,10个序列(17,L5_30,L5_37,L5_44,L5_46,L5_52,L5_53,L5_76,L5_101,L5_106)

[0425] 位置603,总突变18,2个不同的氨基酸

[0426] L->W,17个序列(3,13,20,L5_3,L5_14,L5_20,L5_21,L5_24,L5_42,L5_57,L5_65,L5_66,

[0427] L5_69,L5_82,L5_111,L5_118,L5_120)

[0428] L->M,1个序列(L5_16)

[0429] 位置330,总突变15,1个不同的氨基酸

[0430] T->P,15个序列(8,18,L5_13,L5_30,L5_35,L5_40,L5_43,L5_46,L5_52,L5_72,L5_78,L5_85,L5_90,L5_96,L5_112)

[0431] 位置139,总突变14,1个不同的氨基酸

[0432] L->P,14个序列(5,8,18,L5_24,L5_32,L5_41,L5_53,L5_72,L5_73,L5_84,L5_88,L5_107,

[0433] L5_115,L5_117)

[0434] 位置204,总突变7,1个不同的氨基酸

[0435] H->R,7个序列(7,21,L5_49,L5_62,L5_63,L5_68,L5_101)

[0436] 位置221,总突变6,1个不同的氨基酸

[0437] Q->R,6个序列(3,23,L5_13,L5_20,L5_24,L5_118)

[0438] 位置287,总突变6,1个不同的氨基酸

[0439] T->A,6个序列(8,L5_6,L5_49,L5_63,L5_68,L5_118)

[0440] 位置680,总突变6,2个不同的氨基酸

[0441] R->P,1个序列(L5_6)

[0442] R->A,5个序列(18,20,21,L5_4,L5_8)

[0443] 位置49,总突变5,2个不同的氨基酸

[0444] I->T,1个序列(L5_69)

[0445] I->V,4个序列(1,17,L5_57,L5_68)

[0446] 位置479,总突变5,1个不同的氨基酸

[0447] N->D,5个序列(18,L5_30,L5_35,L5_44,L5_88)

[0448] 附录4续

[0449] 位置594,总突变5,2个不同的氨基酸

[0450] H->Q,1个序列(L5_72)

[0451] H->R,4个序列(1,L5_6,L5_28,L5_56)

[0452] 位置625,总突变5,2个不同的氨基酸

[0453] F->S,3个序列(L5_28,L5_43,L5_112)

[0454] F->L,2个序列(L5_68,L5_88)

[0455] 位置679,总突变5,2个不同的氨基酸

[0456] S->F,1个序列(L5_32)

[0457] S->P,4个序列(18,20,21,L5_8)

[0458] 位置65,总突变4,1个不同的氨基酸

[0459] P->S,4个序列(17,L5_21,L5_101,L5_111)

[0460] 位置126,总突变4,2个不同的氨基酸

[0461] H->S,2个序列(L5_37,L5_92)

[0462] H->R,2个序列(L5_47,L5_84)

[0463] 位置333,总突变4,2个不同的氨基酸

[0464] L->Q,3个序列(L5_75,L5_79,L5_117)

[0465] L->P,1个序列(L5_42)

[0466] 位置502,总突变4,1个不同的氨基酸

[0467] A->V,4个序列(17,L5_1,L5_29,L5_78)

[0468] 位置607,总突变4,3个不同的氨基酸

[0469] E->G,1个序列(L5_103)

[0470] E->A,1个序列(L5_11)

[0471] E->K,2个序列(8,L5_94)

[0472] 位置638,总突变4,1个不同的氨基酸

[0473] H->R,4个序列(L5_21,L5_85,L5_96,L5_111)

[0474] 位置658,总突变4,2个不同的氨基酸

[0475] K->Q,1个序列(L5_3)

[0476] K->R,3个序列(L5_24,L5_75,L5_79)

[0477] 位置8,总突变3,2个不同的氨基酸

[0478] H->P,2个序列(L5_56,L5_115)

[0479] H->R,1个序列(28)

[0480] 位置130,总突变3,1个不同的氨基酸

[0481] P->S,3个序列(L5_13,L5_53,L5_82)

[0482] 位置233,总突变3,1个不同的氨基酸

[0483] E->K,3个序列(L5_21,L5_107,L5_111)

[0484] 位置237,总突变3,1个不同的氨基酸

[0485] Q->R,3个序列(12,L5_43,L5_112)

[0486] 位置249,总突变3,1个不同的氨基酸

[0487] N->D,3个序列(L5_1,L5_15,L5_114)

[0488] 位置283,总突变3,2个不同的氨基酸

[0489] A->D,2个序列(5,L5_101)

[0490] A->T,1个序列(L5_118)

[0491] 附录4续

[0492] 位置307,总突变3,1个不同的氨基酸

[0493] A->V,3个序列(L5_15,L5_73,L5_114)

[0494] 位置344,总突变3,1个不同的氨基酸

[0495] Y->H,3个序列(L5_9,L5_15,L5_114)

[0496] 位置407,总突变3,2个不同的氨基酸

[0497] P->S,2个序列(L5_21,L5_111)

[0498] P->L,1个序列(L5_9)

[0499] 位置428,总突变3,1个不同的氨基酸

[0500] M->L,3个序列(3,13,L5_20)

[0501] 位置430,总突变3,1个不同的氨基酸

[0502] Q->R,3个序列(L5_44,L5_75,L5_79)

[0503] 位置449,总突变3,2个不同的氨基酸

[0504] D->G,2个序列(L5_13,L5_30)

[0505] D->A,1个序列(L5_61)

[0506] 位置644,总突变3,2个不同的氨基酸

[0507] A->T,1个序列(L5_46)

[0508] A->V,2个序列(23,28)

[0509] 位置649,总突变3,1个不同的氨基酸

[0510] N->S,3个序列(L5_24,L5_75,L5_79)

[0511] 位置671,总突变3,1个不同的氨基酸

[0512] L->P,3个序列(L5_30,L5_55,L5_72)

[0513] 位置673,总突变3,2个不同的氨基酸

[0514] E->G,2个序列(L5_75,L5_79)

[0515] E->K,1个序列(L5_114)

[0516] 位置678,总突变3,1个不同的氨基酸

[0517] N->I,3个序列(18,20,21)

[0518] 位置5,总突变2,1个不同的氨基酸

[0519] E->K,2个序列(L5_2,L5_55)

[0520] 位置11,总突变2,2个不同的氨基酸

[0521] H->R,1个序列(L5_78)

[0522] H->Y,1个序列(L5_75)

[0523] 位置12,总突变2,2个不同的氨基酸

[0524] E->V,1个序列(L5_39)

[0525] E->A,1个序列(L5_20)

[0526] 位置14,总突变2,2个不同的氨基酸

[0527] S->T,1个序列(L5_69)

[0528] S->P,1个序列(L5_117)

[0529] 位置17,总突变2,1个不同的氨基酸

[0530] P->S,2个序列(L5_15,L5_114)

[0531] 附录4续

[0532] 位置39,总突变2,2个不同的氨基酸

[0533] M->V,1个序列(L5_68)

[0534] M->L,1个序列(L5_25)

[0535] 位置50,总突变2,1个不同的氨基酸

[0536] I->V,2个序列(L5_49,L5_63)

[0537] 位置66,总突变2,1个不同的氨基酸

[0538] M->L,2个序列(L5_58,L5_76)

[0539] 位置83,总突变2,1个不同的氨基酸

[0540] D->N,2个序列(18,L5_9)

[0541] 位置91,总突变2,2个不同的氨基酸

[0542] Q->R,1个序列(L5_68)

[0543] Q->L,1个序列(L5_16)

[0544] 位置95,总突变2,1个不同的氨基酸

[0545] N->S,2个序列(L5_20,L5_117)

[0546] 位置108,总突变2,1个不同的氨基酸

[0547] D->E,2个序列(L5_15,L5_114)

[0548] 位置135,总突变2,1个不同的氨基酸

[0549] L->P,2个序列(L5_9,L5_32)

[0550] 位置148,总突变2,1个不同的氨基酸

[0551] V->M,2个序列(11,L5_78)

[0552] 位置166,总突变2,1个不同的氨基酸

[0553] P->S,2个序列(L5_9,L5_81)

[0554] 位置179,总突变2,2个不同的氨基酸

[0555] I->T,1个序列(21)

[0556] I->V,1个序列(L5_47)

[0557] 位置184,总突变2,1个不同的氨基酸

[0558] T->A,2个序列(L5_2,L5_71)

[0559] 位置194,总突变2,1个不同的氨基酸

[0560] N->S,2个序列(L5_49,L5_63)

[0561] 位置197,总突变2,1个不同的氨基酸

[0562] T->A,2个序列(L5_14,L5_115)

[0563] 位置199,总突变2,2个不同的氨基酸

[0564] F->L,1个序列(L5_81)

[0565] F->Y,1个序列(L5_41)

[0566] 位置216,总突变2,1个不同的氨基酸

[0567] D->G,2个序列(L5_64,L5_93)

[0568] 位置289,总突变2,2个不同的氨基酸

[0569] M->V,1个序列(L5_76)

[0570] M->L,1个序列(L5_71)

[0571] 附录4续

[0572] 位置302,总突变2,1个不同的氨基酸

[0573] E->K,2个序列(L5_25,L5_57)

[0574] 位置306,总突变2,2个不同的氨基酸

[0575] T->A,1个序列(L5_28)

[0576] T->M,1个序列(30)

[0577] 位置325,总突变2,1个不同的氨基酸

[0578] Y->H,2个序列(L5_75,L5_79)

[0579] 位置346,总突变2,1个不同的氨基酸

[0580] E->D,2个序列(L5_41,L5_62)

[0581] 位置358,总突变2,2个不同的氨基酸

[0582] G->W,1个序列(L5_115)

[0583] G->V,1个序列(5)

[0584] 位置374,总突变2,1个不同的氨基酸

[0585] Q->R,2个序列(L5_52,L5_90)

[0586] 位置380,总突变2,1个不同的氨基酸

[0587] V->A,2个序列(L5_3,L5_117)

[0588] 位置388,总突变2,1个不同的氨基酸

[0589] W->R,2个序列(L5_23,L5_104)

[0590] 位置390,总突变2,1个不同的氨基酸

[0591] R->W,2个序列(11,L5_23)

[0592] 位置391,总突变2,2个不同的氨基酸

[0593] P->S,1个序列(L5_101)

[0594] P->L,1个序列(L5_2)

[0595] 位置409,总突变2,1个不同的氨基酸

[0596] C->R,2个序列(17,L5_1)

[0597] 位置417,总突变2,1个不同的氨基酸

[0598] A->V,2个序列(L5_53,L5_84)

[0599] 位置431,总突变2,1个不同的氨基酸

[0600] P->Q,2个序列(L5_32,L5_51)

[0601] 位置433,总突变2,1个不同的氨基酸

[0602] V->A,2个序列(7,L5_14)

[0603] 位置436,总突变2,1个不同的氨基酸

[0604] A->T,2个序列(L5_9,L5_76)

[0605] 位置450,总突变2,1个不同的氨基酸

[0606] R->H,2个序列(L5_41,L5_59)

[0607] 位置454,总突变2,1个不同的氨基酸

[0608] N->K,2个序列(L5_37,L5_68)

[0609] 位置457,总突变2,2个不同的氨基酸

[0610] M->T,1个序列(L5_18)

[0611] M->R,1个序列(L5_66)

[0612] 附录4续

[0613] 位置470,总突变2,1个不同的氨基酸

[0614] V->A,2个序列(17,L5_1)

[0615] 位置478,总突变2,1个不同的氨基酸

[0616] L->P,2个序列(L5_21,L5_111)

[0617] 位置491,总突变2,1个不同的氨基酸

[0618] L->P,2个序列(L5_76,L5_84)

[0619] 位置494,总突变2,2个不同的氨基酸

[0620] N->D,1个序列(L5_94)

[0621] N->K,1个序列(L5_95)

[0622] 位置503,总突变2,1个不同的氨基酸

[0623] H->R,2个序列(L5_40,L5_59)

[0624] 位置530,总突变2,1个不同的氨基酸

[0625] Q->H,2个序列(L5_61,L5_81)

[0626] 位置545,总突变2,1个不同的氨基酸

[0627] E->G,2个序列(L5_64,L5_93)

[0628] 位置559,总突变2,2个不同的氨基酸

[0629] Q->P,1个序列(L5_103)

[0630] Q->R,1个序列(L5_18)

[0631] 位置572,总突变2,2个不同的氨基酸

[0632] M->L,1个序列(L5_72)

[0633] M->I,1个序列(7)

[0634] 位置597,总突变2,1个不同的氨基酸

[0635] I->T,2个序列(L5_75,L5_79)

[0636] 位置616,总突变2,1个不同的氨基酸

[0637] E->K,2个序列(L5_75,L5_79)

[0638] 位置618,总突变2,1个不同的氨基酸

[0639] L->V,2个序列(18,L5_95)

[0640] 位置623,总突变2,1个不同的氨基酸

[0641] A->V,2个序列(12,L5_115)

[0642] 位置635,总突变2,2个不同的氨基酸

[0643] C->S,1个序列(8)

[0644] C->R,1个序列(L5_106)

[0645] 位置642,总突变2,2个不同的氨基酸

[0646] H->R,1个序列(L5_107)

[0647] H->Y,1个序列(L5_47)

[0648] 位置676,总突变2,1个不同的氨基酸

[0649] S->P,2个序列(L5_80,L5_82)

[0650] 位置15,总突变1,1个不同的氨基酸

[0651] K->T,1个序列(L5_11)

[0652] 附录4续

[0653] 位置23,总突变1,1个不同的氨基酸

[0654] S->P,1个序列(12)

[0655] 位置24,总突变1,1个不同的氨基酸

[0656] T->A,1个序列(L5_40)

[0657] 位置26,总突变1,1个不同的氨基酸

[0658] L->P,1个序列(L5_82)

[0659] 位置29,总突变1,1个不同的氨基酸

[0660] F->L,1个序列(11)

[0661] 位置30,总突变1,1个不同的氨基酸

[0662] P->L,1个序列(L5_23)

[0663] 位置36,总突变1,1个不同的氨基酸

[0664] T->I,1个序列(L5_99)

[0665] 位置37,总突变1,1个不同的氨基酸

[0666] G->W,1个序列(L5_14)

[0667] 位置41,总突变1,1个不同的氨基酸

[0668] L->R,1个序列(L5_40)

[0669] 位置43,总突变1,1个不同的氨基酸

[0670] V->I,1个序列(L5_2)

[0671] 位置51,总突变1,1个不同的氨基酸

[0672] P->S,1个序列(13)

[0673] 位置60,总突变1,1个不同的氨基酸

[0674] S->A,1个序列(L5_41)

[0675] 位置64,总突变1,1个不同的氨基酸

[0676] Y->C,1个序列(L5_61)

[0677] 位置67,总突变1,1个不同的氨基酸

[0678] S->P,1个序列(L5_73)

[0679] 位置69,总突变1,1个不同的氨基酸

[0680] E->G,1个序列(L5_32)

[0681] 位置70,总突变1,1个不同的氨基酸

[0682] A->V,1个序列(L5_88)

[0683] 位置74,总突变1,1个不同的氨基酸

[0684] I->T,1个序列(L5_79)

[0685] 位置77,总突变1,1个不同的氨基酸

[0686] H->R,1个序列(L5_48)

[0687] 位置86,总突变1,1个不同的氨基酸

[0688] I->V,1个序列(L5_39)

[0689] 位置87,总突变1,1个不同的氨基酸

[0690] L->P,1个序列(L5_20)

[0691] 附录4续

[0692] 位置88,总突变1,1个不同的氨基酸

[0693] V->A,1个序列(12)

[0694] 位置89,总突变1,1个不同的氨基酸

[0695] P->S,1个序列(8)

[0696] 位置90,总突变1,1个不同的氨基酸

[0697] C->Y,1个序列(L5_58)

[0698] 位置92,总突变1,1个不同的氨基酸

[0699] S->P,1个序列(L5_44)

[0700] 位置93,总突变1,1个不同的氨基酸

[0701] P->L,1个序列(L5_66)

[0702] 位置96,总突变1,1个不同的氨基酸

[0703] T->M,1个序列(L5_103)

[0704] 位置97,总突变1,1个不同的氨基酸

[0705] P->S,1个序列(L5_71)

[0706] 位置104,总突变1,1个不同的氨基酸

[0707] P->R,1个序列(L5_79)

[0708] 位置105,总突变1,1个不同的氨基酸

[0709] G->E,1个序列(L5_76)

[0710] 位置107,总突变1,1个不同的氨基酸

[0711] N->S,1个序列(L5_47)

[0712] 位置110,总突变1,1个不同的氨基酸

[0713] R->G,1个序列(L5_51)

[0714] 位置112,总突变1,1个不同的氨基酸

[0715] V->A,1个序列(L5_60)

[0716] 位置118,总突变1,1个不同的氨基酸

[0717] V->A,1个序列(L5_62)

[0718] 位置124,总突变1,1个不同的氨基酸

[0719] D->G,1个序列(L5_95)

[0720] 位置125,总突变1,1个不同的氨基酸

[0721] I->V,1个序列(L5_3)

[0722] 位置127,总突变1,1个不同的氨基酸

[0723] P->S,1个序列(L5_40)

[0724] 位置128,总突变1,1个不同的氨基酸

[0725] T->A,1个序列(L5_47)

[0726] 位置131,总突变1,1个不同的氨基酸

[0727] N->S,1个序列(21)

[0728] 位置132,总突变1,1个不同的氨基酸

[0729] P->S,1个序列(L5_104)

[0730] 附录4续

[0731] 位置136,总突变1,1个不同的氨基酸

[0732] L->W,1个序列(13)

[0733] 位置137,总突变1,1个不同的氨基酸

[0734] S->G,1个序列(L5_82)

[0735] 位置138,总突变1,1个不同的氨基酸

[0736] G->R,1个序列(L5_3)

[0737] 位置143,总突变1,1个不同的氨基酸

[0738] H->R,1个序列(L5_3)

[0739] 位置149,总突变1,1个不同的氨基酸

[0740] L->F,1个序列(L5_37)

[0741] 位置151,总突变1,1个不同的氨基酸

[0742] L->F,1个序列(L5_40)

[0743] 位置159,总突变1,1个不同的氨基酸

[0744] R->K,1个序列(L5_13)

[0745] 位置164,总突变1,1个不同的氨基酸

[0746] S->G,1个序列(L5_104)

[0747] 位置168,总突变1,1个不同的氨基酸

[0748] F->S,1个序列(L5_41)

[0749] 位置173,总突变1,1个不同的氨基酸

[0750] R->K,1个序列(L5_57)

[0751] 位置174,总突变1,1个不同的氨基酸

[0752] D->G,1个序列(L5_29)

[0753] 位置187,总突变1,1个不同的氨基酸

[0754] R->G,1个序列(L5_95)

[0755] 位置190,总突变1,1个不同的氨基酸

[0756] Q->R,1个序列(L5_117)

[0757] 位置192,总突变1,1个不同的氨基酸

[0758] F->L,1个序列(L5_42)

[0759] 位置207,总突变1,1个不同的氨基酸

[0760] L->P,1个序列(13)

[0761] 位置208,总突变1,1个不同的氨基酸

[0762] A->V,1个序列(L5_4)

[0763] 位置211,总突变1,1个不同的氨基酸

[0764] R->W,1个序列(L5_56)

[0765] 位置214,总突变1,1个不同的氨基酸

[0766] H->R,1个序列(L5_9)

[0767] 位置222,总突变1,1个不同的氨基酸

[0768] Y->C,1个序列(L5_9)

[0769] 附录4续

[0770] 位置223,总突变1,1个不同的氨基酸

[0771] V->M,1个序列(L5_23)

[0772] 位置225,总突变1,1个不同的氨基酸

[0773] D->G,1个序列(L5_4)

[0774] 位置238,总突变1,1个不同的氨基酸

[0775] Q->H,1个序列(L5_65)

[0776] 位置240,总突变1,1个不同的氨基酸

[0777] T->A,1个序列(L5_55)

[0778] 位置241,总突变1,1个不同的氨基酸

[0779] R->Q,1个序列(16)

[0780] 位置242,总突变1,1个不同的氨基酸

[0781] A->T,1个序列(L5_20)

[0782] 位置250,总突变1,1个不同的氨基酸

[0783] L->P,1个序列(L5_71)

[0784] 位置252,总突变1,1个不同的氨基酸

[0785] Y->H,1个序列(L5_120)

[0786] 位置259,总突变1,1个不同的氨基酸

[0787] A->T,1个序列(16)

[0788] 位置263,总突变1,1个不同的氨基酸

[0789] Q->R,1个序列(L5_95)

[0790] 位置280,总突变1,1个不同的氨基酸

[0791] L->P,1个序列(L5_60)

[0792] 位置282,总突变1,1个不同的氨基酸

[0793] E->G,1个序列(L5_62)

[0794] 位置292,总突变1,1个不同的氨基酸

[0795] P->L,1个序列(L5_56)

[0796] 位置293,总突变1,1个不同的氨基酸

[0797] T->A,1个序列(L5_9)

[0798] 位置295,总突变1,1个不同的氨基酸

[0799] K->E,1个序列(L5_107)

[0800] 位置298,总突变1,1个不同的氨基酸

[0801] R->G,1个序列(L5_90)

[0802] 位置308,总突变1,1个不同的氨基酸

[0803] G->S,1个序列(L5_120)

[0804] 位置309,总突变1,1个不同的氨基酸

[0805] F->S,1个序列(L5_28)

[0806] 位置311,总突变1,1个不同的氨基酸

[0807] R->H,1个序列(16)

[0808] 附录4续

[0809] 位置312,总突变1,1个不同的氨基酸

[0810] L->P,1个序列(L5_29)

[0811] 位置314,总突变1,1个不同的氨基酸

[0812] I->T,1个序列(L5_76)

[0813] 位置322,总突变1,1个不同的氨基酸

[0814] A->T,1个序列(L5_60)

[0815] 位置323,总突变1,1个不同的氨基酸

[0816] P->L,1个序列(21)

[0817] 位置326,总突变1,1个不同的氨基酸

[0818] P->S,1个序列(L5_80)

[0819] 位置331,总突变1,1个不同的氨基酸

[0820] G->E,1个序列(L5_20)

[0821] 位置332,总突变1,1个不同的氨基酸

[0822] T->I,1个序列(23)

[0823] 位置339,总突变1,1个不同的氨基酸

[0824] D->G,1个序列(L5_117)

[0825] 位置343,总突变1,1个不同的氨基酸

[0826] A->T,1个序列(L5_82)

[0827] 位置351,总突变1,1个不同的氨基酸

[0828] L->V,1个序列(L5_61)

[0829] 位置353,总突变1,1个不同的氨基酸

[0830] T->A,1个序列(21)

[0831] 位置356,总突变1,1个不同的氨基酸

[0832] A->G,1个序列(L5_82)

[0833] 位置369,总突变1,1个不同的氨基酸

[0834] F->I,1个序列(L5_118)

[0835] 位置376,总突变1,1个不同的氨基酸

[0836] Y->C,1个序列(L5_118)

[0837] 位置379,总突变1,1个不同的氨基酸

[0838] G->S,1个序列(L5_60)

[0839] 位置383,总突变1,1个不同的氨基酸

[0840] Q->P,1个序列(L5_117)

[0841] 位置392,总突变1,1个不同的氨基酸

[0842] V->A,1个序列(L5_57)

[0843] 位置393,总突变1,1个不同的氨基酸

[0844] A->T,1个序列(L5_72)

[0845] 位置415,总突变1,1个不同的氨基酸

[0846] A->V,1个序列(L5_9)

[0847] 附录4续

[0848] 位置434,总突变1,1个不同的氨基酸

[0849] I->T,1个序列(L5_118)

[0850] 位置435,总突变1,1个不同的氨基酸

[0851] L->P,1个序列(L5_23)

[0852] 位置441,总突变1,1个不同的氨基酸

[0853] E->G,1个序列(L5_120)

[0854] 位置444,总突变1,1个不同的氨基酸

[0855] V->A,1个序列(L5_9)

[0856] 位置446,总突变1,1个不同的氨基酸

[0857] Q->R,1个序列(L5_81)

[0858] 位置447,总突变1,1个不同的氨基酸

[0859] P->L,1个序列(L5_9)

[0860] 位置459,总突变1,1个不同的氨基酸

[0861] H->R,1个序列(L5_78)

[0862] 位置462,总突变1,1个不同的氨基酸

[0863] A->T,1个序列(L5_18)

[0864] 位置468,总突变1,1个不同的氨基酸

[0865] D->A,1个序列(L5_81)

[0866] 位置475,总突变1,1个不同的氨基酸

[0867] V->G,1个序列(L5_8)

[0868] 位置481,总突变1,1个不同的氨基酸

[0869] A->T,1个序列(10)

[0870] 位置484,总突变1,1个不同的氨基酸

[0871] L->P,1个序列(L5_97)

[0872] 位置486,总突变1,1个不同的氨基酸

[0873] L->P,1个序列(L5_56)

[0874] 位置497,总突变1,1个不同的氨基酸

[0875] D->G,1个序列(L5_71)

[0876] 位置498,总突变1,1个不同的氨基酸

[0877] I->V,1个序列(L5_97)

[0878] 位置501,总突变1,1个不同的氨基酸

[0879] E->K,1个序列(L5_81)

[0880] 位置504,总突变1,1个不同的氨基酸

[0881] G->R,1个序列(L5_18)

[0882] 位置514,总突变1,1个不同的氨基酸

[0883] L->F,1个序列(8)

[0884] 位置528,总突变1,1个不同的氨基酸

[0885] L->I,1个序列(L5_62)

[0886] 附录4续

[0887] 位置532,总突变1,1个不同的氨基酸

[0888] G->R,1个序列(L5_117)

[0889] 位置533,总突变1,1个不同的氨基酸

[0890] Q->K,1个序列(L5_104)

[0891] 位置538,总突变1,1个不同的氨基酸

[0892] A->T,1个序列(L5_69)

[0893] 位置539,总突变1,1个不同的氨基酸

[0894] A->T,1个序列(L5_82)

[0895] 位置543,总突变1,1个不同的氨基酸

[0896] E->K,1个序列(L5_2)

[0897] 位置544,总突变1,1个不同的氨基酸

[0898] T->I,1个序列(16)

[0899] 位置551,总突变1,1个不同的氨基酸

[0900] A->T,1个序列(L5_71)

[0901] 位置552,总突变1,1个不同的氨基酸

[0902] L->P,1个序列(L5_3)

[0903] 位置556,总突变1,1个不同的氨基酸

[0904] T->A,1个序列(L5_42)

[0905] 位置558,总突变1,1个不同的氨基酸

[0906] A->V,1个序列(L5_81)

[0907] 位置560,总突变1,1个不同的氨基酸

[0908] R->W,1个序列(13)

[0909] 位置562,总突变1,1个不同的氨基酸

[0910] E->K,1个序列(L5_71)

[0911] 位置570,总突变1,1个不同的氨基酸

[0912] L->I,1个序列(L5_65)

[0913] 位置573,总突变1,1个不同的氨基酸

[0914] A->T,1个序列(L5_76)

[0915] 位置576,总突变1,1个不同的氨基酸

[0916] K->R,1个序列(L5_11)

[0917] 位置577,总突变1,1个不同的氨基酸

[0918] K->Q,1个序列(18)

[0919] 位置600,总突变1,1个不同的氨基酸

[0920] R->K,1个序列(L5_16)

[0921] 位置602,总突变1,1个不同的氨基酸

[0922] G->R,1个序列(3)

[0923] 位置622,总突变1,1个不同的氨基酸

[0924] K->R,1个序列(L5_3)

[0925] 附录4续

[0926] 位置628,总突变1,1个不同的氨基酸

[0927] K->E,1个序列(L5_25)

[0928] 位置632,总突变1,1个不同的氨基酸

[0929] I->T,1个序列(28)

[0930] 位置633,总突变1,1个不同的氨基酸

[0931] I->T,1个序列(L5_107)

[0932] 位置634,总突变1,1个不同的氨基酸

[0933] H->Y,1个序列(L5_53)

[0934] 位置643,总突变1,1个不同的氨基酸

[0935] S->G,1个序列(L5_107)

[0936] 位置646,总突变1,1个不同的氨基酸

[0937] A->V,1个序列(L5_39)

[0938] 位置655,总突变1,1个不同的氨基酸

[0939] A->V,1个序列(L5_18)

[0940] 位置656,总突变1,1个不同的氨基酸

[0941] A->T,1个序列(16)

[0942] 位置661,总突变1,1个不同的氨基酸

[0943] I->V,1个序列(23)

[0944] 位置662,总突变1,1个不同的氨基酸

[0945] T->A,1个序列(L5_116)

[0946] 位置663,总突变1,1个不同的氨基酸

[0947] E->D,1个序列(8)

[0948] 位置667,总突变1,1个不同的氨基酸

[0949] T->A,1个序列(L5_78)

[0950] 位置668,总突变1,1个不同的氨基酸

[0951] S->P,1个序列(L5_3)

[0952] 位置669,总突变1,1个不同的氨基酸

[0953] T->S,1个序列(L5_57)

[0954] 位置670,总突变1,1个不同的氨基酸

[0955] L->F,1个序列(L5_106)

[0956] 位置672,总突变1,1个不同的氨基酸

[0957] I->T,1个序列(L5_72)

[0958] 附录5

[0959]

| | 选择频率* | 浓度 | 比活性 | | | RNA 酶 H 活性 | | CDNA 合成的最高温度 |
|---|---|---|---|---|---|---|---|---|
| | | mg/ml | 37°C (u/mg) | 50°C | 5 min 50°C => 37°C | u/mol | % | 1kb |
| **M-MuLV (wt)** | | **3.4** | **~200 000** | **45 - 50 %** | **11 %** | **~160-200** | **100 %** | **47.8°C** |
| D200N | 25 (30) | 2.91 | 254 000 | 84 % | 15 % | 174 | 97 % | 50.4°C |
| D200A | 3 (30) | 2.73 | 187 000 | 87 % | 18 % | 179 | 99 % | 47.8°C |
| D200Q | | 3.82 | 135 000 | 103 % | 23 % | 161 | 89 % | 47.8°C |
| D200E | | 2.33 | 71 000 | 79 % | 7 % | | | 47.8°C |
| D200V | | 2.04 | 87 000 | 131 % | 2 % | 102 | 57 % | 47.8°C |
| D200W | | 2.84 | 114 000 | 103 % | 4.3 % | | | 47.8°C |
| D200G | 2 (30) | 2.68 | 276 000 | 88 % | 10 % | | | 47.8°C |
| D200P | | 1.15 | 0 | 0 % | 0 % | | | |
| D200K | | 3.14 | 175 000 | 102 % | 2.5 % | | | 47.8°C |
| D200R | | 3.69 | 52 000 | 68 % | 27 % | 171 | 95 % | 45.5°C |
| D200H | | 6.23 | 234 000 | 54 % | 27 % | | | 50.4°C |
| L603W | 17 (18) | 2.41 | 110 000 | 105 % | 23 % | 202 | 112 % | 53.1°C |
| L603F | | 3.18 | 109 000 | 104 % | 9 % | 112 | 62 % | 50.4°C |
| L603Y | | 3.12 | 139 000 | 95 % | 13 % | | | 47.8 - 50.4°C |
| L603P | | 3.24 | 138 000 | 20 % | 15 % | | | 45.5°C |
| L603I | | 2.85 | 178 000 | 33 % | 7 % | 142 | 79 % | |
| L603M | 1 (18) | 3.7 | 149 000 | 77 % | 9 % | 161 | 89 % | 47.8°C |
| L603V | | 2.69 | 190 000 | 39 % | 7 % | | | 45.5°C |
| L603G | | 3.42 | 149 000 | 36 % | 10 % | | | |
| D653N | 10 (23) | 2.39 | 179 000 | 93 % | 21 % | | | 50.4-53.1°C |
| D653K | | 2.38 | 211 000 | 106 % | 15 % | | | 50.4-53.1°C |
| D653A | 4 (23) | 2.98 | 198 000 | 99 % | 18 % | | | 50.4°C |
| D653V | 1 (23) | 2.82 | 172 000 | 98 % | 16 % | | | 50.4°C |
| D653Q | | 2.79 | 182 000 | 93 % | 18 % | | | 50.4°C |
| D653L | | 2.60 | 212 000 | 83 % | 11 % | | | 50.4°C |
| D653H | 3 (23) | 2.41 | 180 000 | 116 % | 13 % | | | 50.4-53.1°C |
| D653G | 5 (23) | 3.05 | 135 000 | 90 % | 13 % | | | 50.4°C |
| D653W | | 2.68 | 178 000 | 93 % | 13 % | | | 50.4°C |
| D653E | | 2.46 | 200 000 | 80 % | 19 % | | | 47.8°C |
| T330P | 15 (15) | 1.76 | 154 000 | 80 % | 21 % | 185 | 103 % | 47.8°C |
| T330N | | 2.62 | 223 000 | 69 % | 13 % | | | 47.8°C |

[0960] 附录5续

[0961]

| | 选择频率* | 浓度 | 比活性 | | | RNA 酶 H 活性 | | CDNA 合成的最高温度 |
|---|---|---|---|---|---|---|---|---|
| | | mg/ml | 37°C (u/mg) | 50°C | 5 min 50°C => 37°C | u/mol | % | 1kb |
| **M-MuLV (wt)** | | **3.4** | **~200 000** | **45 - 50 %** | **11 %** | **~160-200** | **100 %** | **47.8°C** |
| T330L | | 3.26 | 128 000 | 49 % | 9 % | | | 45.5°C |
| T330D | | 3.32 | 240 000 | 55 % | 16 % | | | 47.8°C |
| T330V | | 2.56 | 197 000 | 65 % | 12 % | | | 47.8°C |
| T330S | | 2.33 | 201 000 | 67 % | 15 % | | | 47.8°C |
| Q221R | 6 (6) | 2.38 | 268 000 | 94 % | 0 % | 117 | 65 % | 50.4°C |
| Q221N | | 3.09 | 114 000 | 53 % | 10 % | | | 45.5°C |
| Q221H | | 2.47 | 55 000 | 44 % | 8 % | | | 43.6-45.5°C |
| Q221K | | 1.85 | 179 000 | 77 % | 0 % | | | 45.5 |
| Q221E | | 1.85 | 83 000 | 64 % | 0 % | | | 43.6-45.5°C |
| Q221M | | 1.3 | 107 000 | 58 % | 2 % | | | 45.5 |
| Q221Y | | 1.76 | 60 000 | 77 % | 0 % | | | 43.6-45.5°C |
| E607K | 2 (4) | 3.94 | 182 000 | 84 % | 9 % | 260 | 144 % | 47.8-50.4°C |
| E607A | 1 (4) | 4.28 | 204 000 | 98 % | | | | 47.8°C |
| E607G | 1 (4) | 4.02 | 184 000 | 72 % | | | | 47.8°C |
| E607D | | 2.24 | 181 000 | 69 % | | | | 45.5°C |
| E607Q | | 1.02 | 0 | 0 % | | | | |
| L139P | 14 (14) | 1.34 | 148 000 | 59 % | 7 % | 176 | 98 % | 47.8°C |
| T287A | 6 (6) | 2.34 | 204 000 | 52 % | 13 % | 174 | 96 % | 45.5-47.8°C |
| T287S | | 2.75 | 213 000 | 68 % | 11 % | | | 45.5-47.8°C |
| T287L | | 2.28 | 179 000 | 38 % | 11 % | | | 45.5°C |
| T287N | | 2.46 | 197 000 | 50 % | 12 % | | | 45.5°C |
| T287F | | 2.18 | 190 000 | 48 % | 13 % | | | 45.5°C |
| T287P | | 1.64 | 134 000 | 7 % | 11 % | | | 41.9°C |
| N479D | 5 (5) | 3.15 | 190 000 | 81 % | 9 % | 234 | 130 % | 47.8°C |
| N479Q | | 2.82 | 176 000 | 33 % | | | | 43.6°C |
| N479E | | 2.58 | 179 000 | 48 % | | | | 45.5°C |
| H594R | 4 (5) | 2.67 | 184 000 | 69 % | 14 % | 232 | 129 % | 47.8°C |
| H594K | | 3.66 | 270 000 | 80 % | 9 % | 170 | 94 % | 47.8-50.4°C |
| H594Q | 1 (5) | 3.85 | 231 000 | 75 % | 13 % | | | 47.8-50.4°C |
| H594N | | 4.15 | 216 000 | 61 % | 11 % | | | 47.8°C |
| D449G | 2 (3) | 3.53 | 123 000 | 79 % | 13 % | 162 | 90 % | 45.5-47.8°C |

[0962] 附录5续

[0963]

| | 选择频率* | 浓度 | 比活性 | | | RNA 酶 H 活性 | | CDNA 合成的最高温度 |
|---|---|---|---|---|---|---|---|---|
| | | mg/ml | 37°C (u/mg) | 50°C | 5 min 50°C => 37°C | u/mol | % | 1kb |
| **M-MuLV (wt)** | | **3.4** | **~200 000** | **45 - 50 %** | **11 %** | **~160-200** | **100 %** | **47.8°C** |
| D449E | | 3.25 | 224 000 | 77 % | | | | 45.5°C |
| D449N | | 3.25 | 221 000 | 75 % | | | | 45.5°C |
| D449A | 1 (3) | 4.15 | 212 000 | 99 % | | | | 45.5-47.8°C |
| D449V | | 4.45 | 216 000 | 83 % | | | | 45.5°C |
| M39V | 1 (2) | 3.09 | 146 000 | 54 % | 13 % | 150 | 83 % | 45.5-47.8°C |
| M39L | 1 (2) | 2.24 | 212 000 | 47 % | 9 % | | | 45.5°C |
| M39F | | 2.51 | 141 000 | 43 % | 8 % | | | 45.5°C |
| M39N | | 2.38 | 349 000 | 71 % | 0 % | | | 45.5-47.8°C |
| M66L | 1 (2) | 3.35 | 237 000 | 79 % | 13 % | 215 | 119 % | 47.8°C |
| M66V | 1 (2) | 2.75 | 227 000 | 73 % | | | | 45.5-47.8°C |
| M66I | | 3.27 | 240 000 | 80 % | | | | 45.5-47.8°C |
| L333Q | 3 (4) | 3.34 | 177 000 | 54 % | 9 % | 233 | 129 % | 47.8°C |
| H126R | 2 (4) | 3.58 | 227 000 | 58 % | 4 % | 218 | 121 % | 47.8°C |
| Y344H | 3 (3) | 3.29 | 85 000 | 30 % | 13 % | 117 | 65 % | 45.5°C |
| P130S | 3 (3) | 2.59 | 120 000 | 70 % | 10 % | 156 | 87 % | 47.8°C |
| Q91R | 1 (2) | 2.44 | 160 000 | 56 % | 13 % | 162 | 90 % | 47,8°C |
| N649S | 3 (3) | 3.47 | 157 000 | 45 % | 16 % | | | 47.8°C |
| W388R | 2 (2) | 1.75 | 266 000 | 72 % | 14 % | | | 45.5-47.8°C |
| R390W | 2 (2) | 1.98 | 161 000 | 64 % | 10 % | | | 45.5-47.8°C |
| I179V | 1 (2) | 5.44 | 251 000 | 52 % | | | | 48°C |
| Q374R | 2 (2) | 3.17 | 135 000 | 56 % | 9 % | | | |
| E5K | 2 (2) | 3.2 | 160 000 | 67 % | 5 % | | | |

[0964] 附录6

[0965]

| | 浓度 | 比活性 | | CDNA 合成的最高温度 | |
|---|---|---|---|---|---|
| | mg/ml | 37°C (u/mg) | 50°C | 1kb | 4.5kb |
| **M-MuLV (wt)** | **3.4** | **~200 000** | **45 - 50 %** | **47.8°C** | |
| D200N, L603W (20 mut.) | 0.46 | 260 000 | 131 % | 56°C | |
| D200N, L603W, T330P (M0_1) | 4,17 | 321 194 | 175 % | 56-58°C | 53°C |
| D200N, L603W, T330P, E607K (M1) | 4,27 | 311 576 | 174 % | 60-62°C | 56°C |
| D200N, L603W, T330P, E607K, L139P (M2) | 2,91 | 348 694 | 176 % | 62°C | 61°C |
| D200N, L603W, N479D, H594R (M3) | 5,37 | 371 510 | 182 % | 56-58°C | 56°C |
| D200N, L603W, D653N, D524G (M4) | 4,84 | 358 112 | 155 % | 58-60°C | 58°C |
| D200N, L603W, D653N, D524G, T330P (M6) | 5,23 | 272 031 | 180 % | 60-62°C | 58°C |

[0966] 附录7

[0967] SEQ ID No:1

[0968]

```
基因座       pET-his-MLV-pD          7873 bp    DNA     环状      5-JUN-2007
来源
  生物体
注释      该文件由 Vector NTI 创建
            http://www.invitrogen.com/
注释      ORIGDB|GenBank
注释      VNTDATE|448383004|
注释      VNTDBDATE|448383004|
注释      LSOWNER|
注释      VNTNAME|pET-his-MLV-pD|
注释      VNTAUTHORNAME|Remigijus Skirgaila|
注释      VNTAUTHORTEL|+370-5-2394224|
注释      VNTAUTHOREML|skirgaila@fermentas.lt|
注释      VNTAUTHORWWW|www.fermentas.com|
特征              位置/限定符
     CDS             3534..4391
                     /vntifkey="4"
                     /标记=Ap
                     /note="ORF: Frame #2   Start: atg   Stop:  taa"
     CDS             互补(6586..7674)
                     /vntifkey="4"
                     /标记=lacI
                     /note="ORF: Frame #3   Start: gtg   Stop:  tga"
     终止子        2864..2910
                     /vntifkey="43"
                     /标记=T7\终止子
     复制原点       2947..3402
                     /vntifkey="33"
                     /标记=f1\origin
     复制原点       互补(2936..2936)
                     /vntifkey="33"
                     /标记=ori
     启动子         174..190
                     /vntifkey="30"
                     /标记=PT7
     misc_特征     2394..2669
                     /vntifkey="21"
                     /标记=pD
     misc_特征     2364..2393
                     /vntifkey="21"
                     /标记=gs\linker
     misc_特征     2670..2759
                     /vntifkey="21"
                     /标记=gs\linker
     misc_特征     306..2363
                     /vntifkey="21"
                     /标记=MLV\H+
     misc_特征     258..305
                     /vntifkey="21"
                     /标记=his
碱基计数     1873 a     2155 c     2084 g     1761 t
原点
       1 cacggggcct gccaccatac ccacgccgaa acaagcgctc atgagcccga agtggcgagc
```

```
     61 ccgatcttcc ccatcggtga tgtcggcgat ataggcgcca gcaaccgcac ctgtggcgcc
    121 ggtgatgccg gccacgatgc gtccggcgta gaggatcgag atctatacga aattaatacg
    181 actcactata gggagaccac aacggtttcc ctctagaaat aattttgttt aactttaaga
    241 aagaggagaa attacatatg agaggatcgc atcaccatca ccatcacgga tctggttcca
    301 tgggcatgac cctaaatata gaagatgagc atcggctaca tgagacctca aaagagccag
    361 atgtttctct agggtccaca tggctgtctg attttcctca ggcctgggcg gaaaccgggg
    421 gcatgggact ggcagttcgc caagctcctc tgatcatacc tctgaaagca acctctaccc
    481 ccgtgtccat aaaacaatac cccatgtcac aagaagccag actggggatc aagccccaca
    541 tacagagact gttggaccag ggaatactgg taccctgcca gtccccctgg aacacgcccc
    601 tgctacccgt taagaaacca gggactaatg attataggcc tgtccaggat ctgagagaag
    661 tcaacaagcg ggtggaagac atccacccca ccgtgcccaa cccttacaac ctcttgagcg
    721 ggctcccacc gtcccaccag tggtacactg tgcttgattt aaaggatgcc ttttttctgcc
    781 tgagactcca ccccaccagt cagcctctct tcgcctttga gtggagagat ccagagatgg
    841 gaatctcagg acaattgacc tggaccagac tcccacaggg tttcaaaaac agtcccaccc
    901 tgtttgatga ggcactgcac agagacctag cagacttccg gatccagcac ccagacttga
    961 tcctgctaca gtacgtggat gacttactgc tggccgccac ttctgagcta gactgccaac
   1021 aaggtactcg ggccctgtta caaaccctag ggaacctcgg gtatcgggcc tcggccaaga
   1081 aagcccaaat ttgccagaaa caggtcaagt atctggggta tcttctaaaa gagggtcaga
   1141 gatggctgac tgaggccaga aaagagactg tgatggggca gcctactccg aagacccctc
   1201 gacaactaag ggagttccta gggacggcag gcttctgtcg cctctggatc cctgggtttg
   1261 cagaaatggc agcccccttg taccctctca ccaaaaacggg gactctgttt aattggggcc
   1321 cagaccaaca aaaggcctat caagaaatca agcaagctct tctaactgcc ccagccctgg
   1381 ggttgccaga tttgactaag ccctttgaac tctttgtcga cgagaagcag ggctacgcca
   1441 aaggtgtcct aacgcaaaaa ctgggacctt ggcgtcggcc ggtggcctac ctgtccaaaa
   1501 agctagaccc agtagcagct gggtggcccc cttgcctacg gatggtagca gccattgccg
   1561 tactgacaaa ggatgcaggc aagctaacca tgggacagcc actagtcatt ctggcccccc
   1621 atgcagtaga ggcactagtc aaacaacccc ccgaccgctg gctttccaac gcccggatga
   1681 ctcactatca ggccttgctt ttggacacgg accgggtcca gttcggaccg gtggtagccc
   1741 tgaacccggc tacgctgctc ccactgcctg aggaagggct gcaacacaac tgccttgata
[0969] 1801 tcctggccga agcccacgga acccgacccg acctaacgga ccagccgctc ccagacgccg
   1861 accacacctg gtacacggat ggaagcagtc tcttacaaga gggacagcgt aaggcgggag
   1921 ctgcggtgac caccgagacc gaggtaatct gggctaaagc cctgccagcc gggacatccg
   1981 ctcagcgggc tgaactgata gcactcaccc aggccctaaa gatggcagaa ggtaagaagc
   2041 taaatgttta tactgatagc cgttatgctt ttgctactgc ccatatccat ggagaaatat
   2101 acagaaggcg tgggttgctc acatcagaag gcaaagagat caaaaataaa gacgagatct
   2161 tggccctact aaaagccctc tttctgccca aaagacttag cataatccat tgtccaggac
   2221 atcaaaaggg acacagcgcc gaggctagag gcaaccggat ggctgaccaa gcggcccgaa
   2281 aggcagccat cacagagact ccagacacct ctaccctcct catagaaaat tcatcaccca
   2341 attcccgctt aattaatgaa ttcggatccg gtggcggttc cggcggtgga tctatgggta
   2401 ccgcaaccgc gcccggcgga ttgagtgcga aagcgcctgc aatgaccccg ctgatgctgg
   2461 acacctccag ccgtaagctg gttgcgtggg atggcaccac cgacggtgct gccgttggca
   2521 ttcttgcggt tgctgctgac cagaccagca ccacgctgac gttctacaag tccggcacgt
   2581 tccgttatga ggatgtgctc tggccggagg ctgccagcga cgagacgaaa aaacggaccg
   2641 cgtttgccgg aacggcaatc agcatcgttg gatctggtgg cggttccggc ggtggatctg
   2701 gtggcggttc cggcggtgga tctggtggcg gttccggcgg tggatcgtgt cttctttaag
   2761 cttgcggccg cactcgagca ccaccaccac caccactgag atccggctgc taacaaagcc
   2821 cgaaaggaag ctgagttggc tgctgccacc gctgagcaat aactagcata acccccttggg
   2881 gcctctaaac gggtcttgag gggttttttg ctgaaaggag gaactatatc cggattggcg
   2941 aatgggacgc gccctgtagc ggcgcattaa gcgcggcggg tgtggtggtt acgcgcagcg
   3001 tgaccgctac acttgccagc gccctagcgc ccgctccttt cgctttcttc ccttcctttc
   3061 tcgccacgtt cgccggcttt ccccgtcaag ctctaaatcg ggggctccct ttagggttcc
   3121 gatttagtgc tttacggcac ctcgacccca aaaaacttga ttagggtgat ggttcacgta
   3181 gtgggccatc gccctgatag acggtttttc gccctttgac gttggagtcc acgttcttta
   3241 atagtggact cttgttccaa actggaacaa cactcaaccc tatctcggtc tattcttttg
   3301 atttataagg gattttgccg atttcggcct attggttaaa aaatgagctg atttaacaaa
   3361 aatttaacgc gaattttaac aaaatattaa cgtttacaat ttcaggtggc acttttcggg
   3421 gaaatgtgcg cggaacccct atttgtttat ttttctaaat acattcaaat atgtatccgc
   3481 tcatgagaca ataaccctga taaatgcttc aataatattg aaaaaggaag agtatgagta
   3541 ttcaacattt ccgtgtcgcc cttattccct tttttgcggc attttgcctt cctgtttttg
```

3601 ctcacccaga aacgctggtg aaagtaaaag atgctgaaga tcagttgggt gcacgagtgg
3661 gttacatcga actggatctc aacagcggta agatccttga gagttttcgc cccgaagaac
3721 gttttccaat gatgagcact tttaaagttc tgctatgtgg cgcggtatta tcccgtattg
3781 acgccgggca agagcaactc ggtcgccgca tacactattc tcagaatgac ttggttgagt
3841 actcaccagt cacagaaaag catcttacgg atggcatgac agtaagagaa ttatgcagtg
3901 ctgccataac catgagtgat aacactgcgg ccaacttact tctgacaacg atcggaggac
3961 cgaaggagct aaccgctttt ttgcacaaca tgggggatca tgtaactcgc cttgatcgtt
4021 gggaaccgga gctgaatgaa gccataccaa acgacgagcg tgacaccacg atgcctgcag
4081 caatggcaac aacgttgcgc aaactattaa ctggcgaact acttactcta gcttcccggc
4141 aacaattaat agactggatg gaggcggata aagttgcagg accacttctg cgctcggccc
4201 ttccggctgg ctggtttatt gctgataaat ctggagccgg tgagcgtggg tctcgcggta
4261 tcattgcagc actggggcca gatggtaagc cctcccgtat cgtagttatc tacacgacgg
4321 ggagtcaggc aactatggat gaacgaaata gacagatcgc tgagataggt gcctcactga
4381 ttaagcattg gtaactgtca gaccaagttt actcatatat actttagatt gatttaaaac
4441 ttcatttta atttaaaagg atctaggtga agatcctttt tgataatctc atgaccaaaa
4501 tcccttaacg tgagttttcg ttccactgag cgtcagaccc cgtagaaaag atcaaaggat
4561 cttcttgaga tcctttttt ctgcgcgtaa tctgctgctt gcaaacaaaa aaaccaccgc
4621 taccagcggt ggtttgtttg ccggatcaag agctaccaac tctttttccg aaggtaactg
4681 gcttcagcag agcgcagata ccaaatactg tccttctagt gtagccgtag ttaggccacc
4741 acttcaagaa ctctgtagca ccgcctacat acctcgctct gctaatcctg ttaccagtgg
4801 ctgctgccag tggcgataag tcgtgtctta ccgggttgga ctcaagacga tagttaccgg
4861 ataaggcgca gcggtcgggc tgaacggggg gttcgtgcac acagcccagc ttggagcgaa
4921 cgacctacac cgaactgaga tacctacagc gtgagctatg agaaagcgcc acgcttcccg
4981 aagggagaaa ggcggacagg tatccggtaa gcggcagggt cggaacagga gagcgcacga
5041 gggagcttcc agggggaaac gcctggtatc tttatagtcc tgtcgggttt cgccacctct
5101 gacttgagcg tcgatttttg tgatgctcgt caggggggcg gagcctatgg aaaaacgcca
5161 gcaacgcggc ctttttacgg ttcctggcct tttgctggcc ttttgctcac atgttctttc
5221 ctgcgttatc ccctgattct gtggataacc gtattaccgc ctttgagtga gctgataccg
5281 ctcgccgcag ccgaacgacc gagcgcagcg agtcagtgag cgaggaagcg gaagagcgcc
[0970] 5341 tgatgcggta ttttctcctt acgcatctgt gcggtatttc acaccgcata tatggtgcac
5401 tctcagtaca atctgctctg atgccgcata gttaagccag tatacactcc gctatcgcta
5461 cgtgactggg tcatggctgc gccccgacac ccgccaacac ccgctgacgc gccctgacgg
5521 gcttgtctgc tcccggcatc cgcttacaga caagctgtga ccgtctccgg gagctgcatg
5581 tgtcagaggt tttcaccgtc atcaccgaaa cgcgcgaggc agctgcggta aagctcatca
5641 gcgtggtcgt gaagcgattc acagatgtct gcctgttcat ccgcgtccag ctcgttgagt
5701 ttctccagaa gcgttaatgt ctggcttctg ataaagcggg ccatgttaag ggcggttttt
5761 tcctgtttgg tcactgatgc ctccgtgtaa ggggggatttc tgttcatggg ggtaatgata
5821 ccgatgaaac gagagaggat gctcacgata cgggttactg atgatgaaca tgcccggtta
5881 ctggaacgtt gtgagggtaa acaactggcg gtatggatgc ggcgggacca gagaaaaatc
5941 actcagggtc aatgccagcg cttcgttaat acagatgtag gtgttccaca gggtagccag
6001 cagcatcctg cgatgcagat ccggaacata atggtgcagg gcgctgactt ccgcgtttcc
6061 agactttacg aaacacggaa accgaagacc attcatgttg ttgctcaggt cgcagacgtt
6121 ttgcagcagc agtcgcttca cgttcgctcg cgtatcggtg attcattctg ctaaccagta
6181 aggcaacccc gccagcctag ccgggtcctc aacgacagga gcacgatcat gcgcacccgt
6241 ggggccgcca tgccggcgat aatggcctgc ttctcgccga aacgtttggt ggcgggacca
6301 gtgacgaagg cttgagcgag ggcgtgcaag attccgaata ccgcaagcga caggccgatc
6361 atcgtcgcgc tccagcgaaa gcggtcctcg ccgaaaatga cccagagcgc tgccggcacc
6421 tgtcctacga gttgcatgat aaagaagaca gtcataagtg cggcgacgat agtcatgccc
6481 cgcgcccacc ggaaggagct gactgggttg aaggctctca agggcatcgg tcgagatccc
6541 ggtgcctaat gagtgagcta acttacatta attgcgttgc gctcactgcc cgctttccag
6601 tcgggaaacc tgtcgtgcca gctgcattaa tgaatcggcc aacgcgcggg gagaggcggt
6661 ttgcgtattg ggcgccaggg tggtttttct tttcaccagt gagacgggca acagctgatt
6721 gcccttcacc gcctggccct gagagagttg cagcaagcgg tccacgctgg tttgccccag
6781 caggcgaaaa tcctgtttga tggtggttaa cggcgggata taacatgagc tgtcttcggt
6841 atcgtcgtat cccactaccg agatatccgc accaacgcgc agcccggact cggtaatggc
6901 gcgcattgcg cccagcgcca tctgatcgtt ggcaaccagc atcgcagtgg gaacgatgcc
6961 ctcattcagc atttgcatgg tttgttgaaa accggacatg gcactccagt cgccttcccg
7021 ttccgctatc ggctgaattt gattgcgagt gagatattta tgccagccag ccagacgcag
7081 acgcgccgag acagaactta atgggcccgc taacagcgcg atttgctggt gacccaatgc

```
7141 gaccagatgc tccacgccca gtcgcgtacc gtcttcatgg gagaaaataa tactgttgat
7201 gggtgtctgg tcagagacat caagaaataa cgccggaaca ttagtgcagg cagcttccac
7261 agcaatggca tcctggtcat ccagcggata gttaatgatc agcccactga cgcgttgcgc
7321 gagaagattg tgcaccgccg ctttacaggc ttcgacgccg cttcgttcta ccatcgacac
7381 caccacgctg gcacccagtt gatcggcgcg agatttaatc gccgcgacaa tttgcgacgg
7441 cgcgtgcagg gccagactgg aggtggcaac gccaatcagc aacgactgtt tgcccgccag
7501 ttgttgtgcc acgcggttgg gaatgtaatt cagctccgcc atcgccgctt ccactttttc
7561 ccgcgttttc gcagaaacgt ggctggcctg gttcaccacg cgggaaacgg tctgataaga
7621 gacaccggca tactctgcga catcgtataa cgttactggt ttcacattca ccaccctgaa
7681 ttgactctct tccgggcgct atcatgccat accgcgaaag gttttgcgcc attcgatggt
7741 gtccgggatc tcgacgctct cccttatgcg actcctgcat taggaagcag cccagtagta
7801 ggttgaggcc gttgagcacc gccgccgcaa ggaatggtgc atgcaaggag atggcgccca
7861 acagtccccc ggc
```

[0971]

[0972] SEQ ID No:2

基因座 pET_his_del_pD 7702 bp DNA 环状 6-JUN-2007
来源
生物体
注释 该文件由Vector NTI创建
http://www.invitrogen.com/
注释 VNTDATE|448455242|
注释 VNTDBDATE|448455855|
注释 LSOWNER|
注释 VNTNAME|pET_his_del_pD|
注释 VNTAUTHORNAME|Remigijus Skirgaila|
注释 VNTAUTHORTEL|+370-5-2394224|
注释 VNTAUTHOREML|skirgaila@fermentas.lt|
注释 VNTAUTHORWWW|www.fermentas.com|
特征 位置/限定符
misc_特征 2499..2588
/vntifkey="21"
/标记=gs\linker
misc_特征 2193..2222
/vntifkey="21"

[0973]

/标记=gs\linker
misc_特征 2223..2498
/vntifkey="21"
/标记=pD
启动子 174..190
/vntifkey="30"
/标记=PT7
复制原点 互补(2765..2765)
/vntifkey="33"
/标记=ori
复制原点 2776..3231
/vntifkey="33"
/标记=f1\原点
终止子 2693..2739
/vntifkey="43"
/标记=T7\终止子
CDS 互补(6415..7503)
/vntifkey="4"
/标记=lacI
/note="ORF: Frame #3 Start: gtg Stop: tga"
CDS 3363..4220

[0974]

```
                  /vntifkey="4"
                  /标记=Ap
                  /note="ORF: Frame #2   Start: atg   Stop:  taa"
     misc_特征    258..305
                  /vntifkey="21"
                  /标记=his
     misc_特征    306..2192
                  /vntifkey="21"
                  /标记=del
碱基计数     1823 a     2115 c     2033 g     1731 t
原点
        1 cacggggcct gccaccatac ccacgccgaa acaagcgctc atgagcccga agtggcgagc
       61 ccgatcttcc ccatcggtga tgtcggcgat ataggcgcca gcaaccgcac ctgtggcgcc
      121 ggtgatgccg gccacgatgc gtccggcgta gaggatcgag atctatacga aattaatacg
      181 actcactata gggagaccac aacggtttcc ctctagaaat aattttgttt aactttaaga
      241 aagaggagaa attacatatg agaggatcgc atcaccatca ccatcacgga tctggttcca
      301 tgggcatgac cctaaatata gaagatgagc atcggctaca tgagacctca aaagagccag
      361 atgtttctct agggtccaca tggctgtctg attttcctca ggcctgggcg gaaaccgggg
      421 gcatgggact ggcagttcgc caagctcctc tgatcatacc tctgaaagca acctctaccc
      481 ccgtgtccat aaaacaatac cccatgtcac aagaagccag actggggatc aagccccaca
      541 tacagagact gttggaccag ggaatactgg taccctgcca gtccccctgg aacacgcccc
      601 tgctacccgt taagaaacca gggactaatg attataggcc tgtccaggat ctgagagaag
      661 tcaacaagcg ggtggaagac atccacccca ccgtgcccaa cccttacaac ctcttgagcg
      721 ggctcccacc gtcccaccag tggtacactg tgcttgattt aaaggatgcc ttttttctgcc
      781 tgagactcca ccccaccagt cagcctctct tcgcctttga gtggagagat ccagagatgg
      841 gaatctcagg acaattgacc tggaccagac tcccacaggg tttcaaaaac agtcccaccc
      901 tgtttgatga ggcactgcac agagacctag cagacttccg gatccagcac ccagacttga
      961 tcctgctaca gtacgtggat gacttactgc tggccgccac ttctgagcta gactgccaac
     1021 aaggtactcg ggccctgtta caaaccctag ggacggcagg cttctgtcgc ctctggatcc
     1081 ctgggtttgc agaaatggca gccccccttgt acccctctcac caaaacgggg actctgttta
     1141 attggggccc agaccaacaa aaggcctatc aagaaatcaa gcaagctctt ctaactgccc
     1201 cagccctggg gttgccagat ttgactaagc cctttgaact ctttgtcgac gagaagcagg
     1261 gctacgccaa aggtgtccta acgcaaaaac tgggaccttg gcgtcggccg gtggcctacc
     1321 tgtccaaaaa gctagaccca gtagcagctg ggtggccccc ttgcctacgg atggtagcag
     1381 ccattgccgt actgacaaag gatgcaggca agctaaccat gggacagcca ctagtcattc
     1441 tggcccccca tgcagtagag gcactagtca aacaaccccc cgaccgctgg ctttccaacg
     1501 cccggatgac tcactatcag gccttgcttt tggacacgga ccgggtccag ttcggaccgg
     1561 tggtagccct gaacccggct acgctgctcc cactgcctga ggaagggctg caacacaact
     1621 gccttgatat cctggccgaa gcccacggaa cccgacccga cctaacggac cagccgctcc
     1681 cagacgccga ccacacctgg tacacggatg gaagcagtct cttacaagag ggacagcgta
     1741 aggcgggagc tgcggtgacc accgagaccg aggtaatctg ggctaaagcc ctgccagccg
     1801 ggacatccgc tcagcgggct gaactgatag cactcaccca ggccctaaag atggcagaag
     1861 gtaagaagct aaatgtttat actaatagcc gttatgcttt tgctactgcc catatccatg
     1921 gagaaatata cagaaggcgt gggttgctca catcagaagg caaagagatc aaaaataaag
     1981 acgagatctt ggccctacta aaagccctct ttctgcccaa aagacttagc ataatccatt
     2041 gtccaggaca tcaaaaggga cacagcgccg aggctagagg caaccggatg gctgaccaag
     2101 cggcccgaaa ggcagccatc acagagactc cagacacctc taccctcctc atagaaaatt
     2161 catcacccaa ttcccgctta attaatgaat tcggatccgg tggcggttcc ggcggtggat
     2221 ctatgggtac cgcaaccgcg cccggcggat tgagtgcgaa agcgcctgca atgaccccgc
     2281 tgatgctgga cacctccagc cgtaagctgg ttgcgtggga tggcaccacc gacggtgctg
     2341 ccgttggcat tcttgcggtt gctgctgacc agaccagcac cacgctgacg ttctacaagt
     2401 ccggcacgtt ccgttatgag gatgtgctct ggccggaggc tgccagcgac gagacgaaaa
     2461 aacggaccgc gtttgccgga acggcaatca gcatcgttgg atctggtggc ggttccggcg
     2521 gtggatctgg tggcggttcc ggcggtggat ctggtggcgg ttccggcggt ggatcgtgtc
     2581 ttctttaagc ttgcggccgc actcgagcac caccaccacc accactgaga tccggctgct
     2641 aacaaagccc gaaaggaagc tgagttggct gctgccaccg ctgagcaata actagcataa
     2701 ccccttgggg cctctaaacg ggtcttgagg ggtttttttgc tgaaaggagg aactatatcc
     2761 ggattggcga atgggacgcg ccctgtagcg gcgcattaag cgcggcgggt gtggtggtta
     2821 cgcgcagcgt gaccgctaca cttgccagcg ccctagcgcc cgctcctttc gctttcttcc
```

2881 cttcctttct cgccacgttc gccggctttc cccgtcaagc tctaaatcgg gggctcccct
2941 tagggttccg atttagtgct ttacggcacc tcgaccccaa aaaacttgat tagggtgatg
3001 gttcacgtag tgggccatcg ccctgataga cggtttttcg ccctttgacg ttggagtcca
3061 cgttctttaa tagtggactc ttgttccaaa ctggaacaac actcaaccct atctcggtct
3121 attcttttga tttataaggg attttgccga tttcggccta ttggttaaaa aatgagctga
3181 tttaacaaaa atttaacgcg aattttaaca aaatattaac gtttacaatt tcaggtggca
3241 cttttcgggg aaatgtgcgc ggaacccctа tttgtttatt tttctaaata cattcaaata
3301 tgtatccgct catgagacaa taaccctgat aaatgcttca ataatattga aaaaggaaga
3361 gtatgagtat tcaacatttc cgtgtcgccc ttattcccct ttttgcggca ttttgccttc
3421 ctgtttttgc tcacccagaa acgctggtga aagtaaaaga tgctgaagat cagttgggtg
3481 cacgagtggg ttacatcgaa ctggatctca acagcggtaa gatccttgag agttttcgcc
3541 ccgaagaacg ttttccaatg atgagcactt ttaaagttct gctatgtggc gcggtattat
3601 cccgtattga cgccgggcaa gagcaactcg gtcgccgcat acactattct cagaatgact
3661 tggttgagta ctcaccagtc acagaaaagc atcttacgga tggcatgaca gtaagagaat
3721 tatgcagtgc tgccataacc atgagtgata acactgcggc caacttactt ctgacaacga
3781 tcggaggacc gaaggagcta accgcttttt tgcacaacat gggggatcat gtaactcgcc
3841 ttgatcgttg ggaaccggag ctgaatgaag ccataccaaa cgacgagcgt gacaccacga
3901 tgcctgcagc aatggcaaca acgttgcgca aactattaac tggcgaacta cttactctag
3961 cttcccggca acaattaata gactggatgg aggcggataa agttgcagga ccacttctgc
4021 gctcggccct tccggctggc tggtttattg ctgataaatc tggagccggt gagcgtgggt
4081 ctcgcggtat cattgcagca ctggggccag atggtaagcc ctcccgtatc gtagttatct
4141 acacgacggg gagtcaggca actatggatg aacgaaatag acagatcgct gagataggtg
4201 cctcactgat taagcattgg taactgtcag accaagttta ctcatatata ctttagattg
4261 atttaaaact tcatttttaa tttaaaagga tctaggtgaa gatccttttt gataatctca
4321 tgaccaaaat cccttaacgt gagttttcgt tccactgagc gtcagacccc gtagaaaaga
4381 tcaaaggatc ttcttgagat cctttttttc tgcgcgtaat ctgctgcttg caaacaaaaa
4441 aaccaccgct accagcggtg gtttgtttgc cggatcaaga gctaccaact ctttttccga
4501 aggtaactgg cttcagcaga gcgcagatac caaatactgt ccttctagtg tagccgtagt
4561 taggccacca cttcaagaac tctgtagcac cgcctacata cctcgctctg ctaatcctgt
[0975] 4621 taccagtggc tgctgccagt ggcgataagt cgtgtcttac cgggttggac tcaagacgat
4681 agttaccgga taaggcgcag cggtcgggct gaacgggggg ttcgtgcaca cagcccagct
4741 tggagcgaac gacctacacc gaactgagat acctacagcg tgagctatga gaaagcgcca
4801 cgcttcccga agggagaaag gcggacaggt atccggtaag cggcagggtc ggaacaggag
4861 agcgcacgag ggagcttcca gggggaaacg cctggtatct ttatagtcct gtcgggtttc
4921 gccacctctg acttgagcgt cgatttttgt gatgctcgtc aggggggcgg agcctatgga
4981 aaaacgccag caacgcggcc tttttacggt tcctggcctt ttgctggcct tttgctcaca
5041 tgttctttcc tgcgttatcc cctgattctg tggataaccg tattaccgcc tttgagtgag
5101 ctgataccgc tcgccgcagc cgaacgaccg agcgcagcga gtcagtgagc gaggaagcgg
5161 aagagcgcct gatgcggtat tttctcctta cgcatctgtg cggtatttca caccgcatat
5221 atggtgcact ctcagtacaa tctgctctga tgccgcatag ttaagccagt atacactccg
5281 ctatcgctac gtgactgggt catggctgcg ccccgacacc cgccaacacc cgctgacgcg
5341 ccctgacggg cttgtctgct cccggcatcc gcttacagac aagctgtgac cgtctccggg
5401 agctgcatgt gtcagaggtt ttcaccgtca tcaccgaaac gcgcgaggca gctgcggtaa
5461 agctcatcag cgtggtcgtg aagcgattca cagatgtctg cctgttcatc cgcgtccagc
5521 tcgttgagtt tctccagaag cgttaatgtc tggcttctga taaagcgggc catgttaagg
5581 gcggtttttt cctgtttggt cactgatgcc tccgtgtaag ggggatttct gttcatgggg
5641 gtaatgatac cgatgaaacg agagaggatg ctcacgatac gggttactga tgatgaacat
5701 gcccggttac tggaacgttg tgagggtaaa caactggcgg tatggatgcg gcgggaccag
5761 agaaaaatca ctcagggtca atgccagcgc ttcgttaata cagatgtagg tgttccacag
5821 ggtagccagc agcatcctgc gatgcagatc cggaacataa tggtgcaggg cgctgacttc
5881 cgcgtttcca gactttacga aacacggaaa ccgaagacca ttcatgttgt tgctcaggtc
5941 gcagacgttt tgcagcagca gtcgcttcac gttcgctcgc gtatcggtga ttcattctgc
6001 taaccagtaa ggcaaccccg ccagcctagc cgggtcctca acgacaggag cacgatcatg
6061 cgcacccgtg gggccgccat gccggcgata atggcctgct tctcgccgaa acgtttggtg
6121 gcgggaccag tgacgaaggc ttgagcgagg gcgtgcaaga ttccgaatac cgcaagcgac
6181 aggccgatca tcgtcgcgct ccagcgaaag cggtcctcgc cgaaaatgac ccagagcgct
6241 gccggcacct gtcctacgag ttgcatgata aagaagacag tcataagtgc ggcgacgata
6301 gtcatgcccc gcgcccaccg gaaggagctg actgggttga aggctctcaa gggcatcggt
6361 cgagatcccg gtgcctaatg agtgagctaa cttacattaa ttgcgttgcg ctcactgccc

[0976]

```
6421 gctttccagt cgggaaacct gtcgtgccag ctgcattaat gaatcggcca acgcgcgggg
6481 agaggcggtt tgcgtattgg gcgccagggt ggtttttctt ttcaccagtg agacgggcaa
6541 cagctgattg cccttcaccg cctggccctg agagagttgc agcaagcggt ccacgctggt
6601 ttgccccagc aggcgaaaat cctgtttgat ggtggttaac ggcgggatat aacatgagct
6661 gtcttcggta tcgtcgtatc ccactaccga gatatccgca ccaacgcgca gcccggactc
6721 ggtaatggcg cgcattgcgc ccagcgccat ctgatcgttg gcaaccagca tcgcagtggg
6781 aacgatgccc tcattcagca tttgcatggt ttgttgaaaa ccggacatgg cactccagtc
6841 gccttcccgt tccgctatcg gctgaatttg attgcgagtg agatatttat gccagccagc
6901 cagacgcaga cgcgccgaga cagaacttaa tgggcccgct aacagcgcga tttgctggtg
6961 acccaatgcg accagatgct ccacgcccag tcgcgtaccg tcttcatggg agaaaataat
7021 actgttgatg ggtgtctggt cagagacatc aagaaataac gccggaacat tagtgcaggc
7081 agcttccaca gcaatggcat cctggtcatc cagcggatag ttaatgatca gcccactgac
7141 gcgttgcgcg agaagattgt gcaccgccgc tttacaggct tcgacgccgc ttcgttctac
7201 catcgacacc accacgctgg cacccagttg atcggcgcga gatttaatcg ccgcgacaat
7261 ttgcgacggc gcgtgcaggg ccagactgga ggtggcaacg ccaatcagca acgactgttt
7321 gcccgccagt tgttgtgcca cgcggttggg aatgtaattc agctccgcca tcgccgcttc
7381 cacttttcc cgcgttttcg cagaaacgtg gctggcctgg ttcaccacgc gggaaacggt
7441 ctgataagag acaccggcat actctgcgac atcgtataac gttactggtt tcacattcac
7501 caccctgaat tgactctctt ccgggcgcta tcatgccata ccgcgaaagg ttttgcgcca
7561 ttcgatggtg tccgggatct cgacgctctc ccttatgcga ctcctgcatt aggaagcagc
7621 ccagtagtag gttgaggccg ttgagcaccg ccgccgcaag gaatggtgca tgcaaggaga
7681 tggcgcccaa cagtcccccg gc
```

[0977] SEQ ID No:3

[0978] 寡核苷酸(22b)

[0979] pro-pIVEX

[0980] 5'-GCGAGCCCGATCTTCCCCATCG-3'

[0981] SEQ ID No:4

[0982] 寡核苷酸(21b)

[0983] pD-ter

[0984] 5'-AAGAAGACACGATCCACCGCC-3'

[0985] SEQ ID No:5

[0986] 寡核苷酸(38b)

[0987] assrA

[0988] 5'-TTAAGCTGCTAAAGCGTAGTTTTCGTCGTTTGCGACTA-3'

[0989] SEQ ID No:6

[0990] 蛋白质(833aa)

[0991] MLV_pD

[0992]

mrgshhhhhhgsgsmgmtlniedehrlhetskepdvslgstwlsdfpqawaetggmglavrqapliiplkatstpvs
ikqypmsqearlgikphiqrlldqgilvpcqspwntpllpvkkpgtndyrpvqdlrevnkrvedihptvpnpynlls
glppshqwytvldlkdaffclrlhptsqplfafewrdpemgisgqltwtrlpqgfknsptlfdealhrdladfriqh
pdlillqyvddlllaatseldcqqgtrallqtlgnlgyrasakkaqicqkqvkylgyllkegqrwltearketvmgq
ptpktprqlreflgtagfcrlwipgfaemaaplypltktgtlfnwgpdqqkayqeikqalltapalglpdltkpfel
fvdekqgyakgvltqklgpwrrpvaylskkldpvaagwppclrmvaaiavltkdagkltmgqplvilaphavealvk
qppdrwlsnarmthyqallldtdrvqfgpvvalnpatllplpeeglqhncldilaeahgtrpdltdqplpdadhtwy
tdgssllqegqrkagaavtteteviwakalpagtsaqraelialtqalkmaegkklnvytdsryafatahihgeiyr
rrglltsegkeiknkdeilallkalflpkrlsiihcpghqkghsaeargnrmadqaarkaaitetpdtstllienss

pnsrlinefgsgggsgggsmgtatapgglsakapamtplmldtssrklvawdgttdgaavgilavaadqtsttltfy
ksgtfryedvlwpeaasdetkkrtafagtaisivgsgggsgggsgggsgggsgggsgggscll

[0993] SEQ ID No:7

[0994] 蛋白质(766aa)

[0995] del_pD

[0996]

mrgshhhhhhgsgsmgmtlniedehrlhetskepdvslgstwlsdfpqawaetggmglavrqapliiplkatstpvs
ikqypmsqearlgikphiqrlldqgilvpcqspwntpllpvkkpgtndyrpvqdlrevnkrvedihptvpnpynlls
glppshqwytvldlkdaffclrlhptsqplfafewrdpemgisgqltwtrlpqgfknsptlfdealhrdladfriqh
pdlillqyvddlllaatseldcqqgtrallqtlgtagfcrlwipgfaemaaplypltktgtlfnwgpdqqkayqeik
qalltapalglpdltkpfelfvdekqgyakgvltqklgpwrrpvaylskkldpvaagwppclrmvaaiavltkdagk
ltmgqplvilaphavealvkqppdrwlsnarmthyqalldtdrvqfgpvvalnpatllplpeeglqhncldilaea
hgtrpdltdqplpdadhtwytdgssllqegqrkagaavtteteviwakalpagtsaqraelialtqalkmaegkkln
vytdsryafatahihgeiyrrrglltsegkeiknkdeilallkalflpkrlsiihcpghqkghsaeargnrmadqaa
rkaaitetpdtstlliensspnsrlinefgsgggsgggsmgtatapgglsakapamtplmldtssrklvawdgttdg
aavgilavaadqtsttltfyksgtfryedvlwpeaasdetkkrtafagtaisivgsgggsgggsgggsgggsgggsg
ggscll

[0997] SEQ ID No:8

[0998] 寡核苷酸(15b)

[0999] pD_42

[1000] 5’-TTACGGCTGGAGGTG-3’

[1001] SEQ ID No:9

[1002] 寡核苷酸(28b)

[1003] RD_Nde

[1004] 5’-CTTTAAGAAAGAGGAGAAATTACATATG-3’

[1005] SEQ ID No:10

[1006] 寡核苷酸(14b)

[1007] pD_55

[1008] 5’-GCCGGGCGCGGTTG-3’

[1009] SEQ ID No:11

[1010] 寡核苷酸(23b)

[1011] M_F

[1012] 5’-GATCAAGCCCCACATACAGAGAC-3’

[1013] SEQ ID No:12

[1014] 寡核苷酸(19b)

[1015] M_2R

[1016] 5’-GCCCTGCTTCTCGTCGACA-3’

[1017] SEQ ID No:13

[1018] 寡核苷酸(40b)

[1019] M_Esp

[1020] 5’-ATCGTCTCCCATGGGCATGACCCTAAATATAGAAGATGAG-3’

[1021] SEQ ID No:14

[1022] 寡核苷酸(32b)

[1023] M_Eri

[1024] 5’-AATGAATTCATTAATTAAGCGGGAATTGGGTG-3’

[1025] SEQ ID No:15

[1026] 寡核苷酸(18b)

[1027] M_1R

[1028] 5’-CAGGGCCCGAGTACCTTG-3’

[1029] SEQ ID No:16

[1030] 寡核苷酸(17b)

[1031] M_3F

[1032] 5’-CCAGTTCGGACCGGTGG-3’

[1033] SEQ ID No:17

[1034] 寡核苷酸(19b)

[1035] pD_ter-

[1036] 5’-AAGAAGACACGATCCACCG-3’

[1037] SEQ ID No:18

[1038] 寡核苷酸(36b)

[1039] M_Hind3+

[1040] 5’-CGGATCAAGCTTAATTAATTAAGCGGGAATTGGGTG-3’

[1041] SEQ ID No:19

[1042]
```
基因座       pET_his_MLV          7474 bp    DNA     环状      4-JUN-2007
来源
  生物体
注释      该文件由Vector NTI创建
            http://www.invitrogen.com/
注释      ORIGDB|GenBank
注释      VNTDATE|446486240|
注释      VNTDBDATE|448293778|
注释      LSOWNER|
注释      VNTNAME|pET_his_MLV|
注释      VNTAUTHORNAME|Remigijus Skirgaila|
注释      VNTAUTHORTEL|+370-5-2394224|
```

注释 VNTAUTHOREML|skirgaila@fermentas.lt|
注释 VNTAUTHORWWW|www.fermentas.com|
特征 位置/限定符
CDS 3135..3992
/vntifkey="4"
/标记=Ap
/note="ORF: Frame #2 Start: atg Stop: taa"
CDS 互补(6187..7275)
/vntifkey="4"
/标记=lacI
/note="ORF: Frame #3 Start: gtg Stop: tga"
终止子 2465..2511
/vntifkey="43"
/标记=T7\终止子
复制原点 2548..3003
/vntifkey="33"
/标记=f1\原点
复制原点 互补(2537..2537)
/vntifkey="33"
/标记=ori
启动子 174..190
/vntifkey="30"
/标记=PT7
CDS 306..2360
/vntifkey="4"
/标记=MLV
CDS 258..305
/vntifkey="4"

[1043] /标记=his

碱基计数 1810 a 2046 c 1942 g 1676 t
原点

```
   1 cacggggcct gccaccatac ccacgccgaa acaagcgctc atgagcccga agtggcgagc
  61 ccgatcttcc ccatcggtga tgtcggcgat ataggcgcca gcaaccgcac ctgtggcgcc
 121 ggtgatgccg gccacgatgc gtccggcgta gaggatcgag atctatacga attaatacg
 181 actcactata gggagaccac aacggtttcc ctctagaaat aattttgttt aactttaaga
 241 aagaggagaa attacatatg agaggatcgc atcaccatca ccatcacgga tctggttcca
 301 tgggcatgac cctaaatata gaagatgagc atcggctaca tgagacctca aaagagccag
 361 atgtttctct agggtccaca tggctgtctg attttcctca ggcctgggcg gaaaccgggg
 421 gcatgggact ggcagttcgc caagctcctc tgatcatacc tctgaaagca acctctaccc
 481 ccgtgtccat aaaacaatac cccatgtcac aagaagccag actggggatc aagccccaca
 541 tacagagact gttggaccag ggaatactgg taccctgcca gtccccctgg aacacgcccc
 601 tgctacccgt taagaaacca gggactaatg attataggcc tgtccaggat ctgagagaag
 661 tcaacaagcg ggtggaagac atccacccca ccgtgcccaa cccttacaac ctcttgagcg
 721 ggctcccacc gtcccaccag tggtacactg tgcttgattt aaaggatgcc tttttctgcc
 781 tgagactcca ccccaccagt cagcctctct tcgcctttga gtggagagat ccagagatgg
 841 gaatctcagg acaattgacc tggaccagac tcccacaggg tttcaaaaac agtcccaccc
 901 tgtttgatga ggcactgcac agagacctag cagacttccg gatccagcac ccagacttga
 961 tcctgctaca gtacgtggat gacttactgc tggccgccac ttctgagcta gactgccaac
1021 aaggtactcg ggccctgtta caaaccctag ggaacctcgg gtatcgggcc tcggccaaga
1081 aagcccaaat ttgccagaaa caggtcaagt atctggggta tcttctaaaa gagggtcaga
1141 gatggctgac tgaggccaga aaagagactg tgatggggca gcctactccg aagacccctc
1201 gacaactaag ggagttccta gggacggcag gcttctgtcg cctctggatc cctgggtttg
1261 cagaaatggc agcccccttg taccctctca ccaaaacggg gactctgttt aattggggcc
1321 cagaccaaca aaaggcctat caagaaatca agcaagctct tctaactgcc ccagccctgg
1381 ggttgccaga tttgactaag ccctttgaac tctttgtcga cgagaagcag ggctacgcca
1441 aaggtgtcct aacgcaaaaa ctgggacctt ggcgtcggcc ggtggcctac ctgtccaaaa
1501 agctagaccc agtagcagct gggtggcccc cttgcctacg gatggtagca gccattgccg
1561 tactgacaaa ggatgcaggc aagctaacca tgggacagcc actagtcatt ctggcccccc
1621 atgcagtaga ggcactagtc aaacaacccc ccgaccgctg gctttccaac gcccggatga
```

1681 ctcactatca ggccttgctt ttggacacgg accgggtcca gttcggaccg gtggtagccc
1741 tgaacccggc tacgctgctc ccactgcctg aggaagggct gcaacacaac tgccttgata
1801 tcctggccga agcccacgga acccgacccg acctaacgga ccagccgctc ccagacgccg
1861 accacacctg gtacacggat ggaagcagtc tcttacaaga gggacagcgt aaggcgggag
1921 ctgcggtgac caccgagacc gaggtaatct gggctaaagc cctgccagcc gggacatccg
1981 ctcagcgggc tgaactgata gcactcaccc aggccctaaa gatggcagaa ggtaagaagc
2041 taaatgttta tactgatagc cgttatgctt ttgctactgc ccatatccat ggagaaatat
2101 acagaaggcg tgggttgctc acatcagaag gcaaagagat caaaaataaa gacgagatct
2161 tggccctact aaaagccctc tttctgccca aaagacttag cataatccat tgtccaggac
2221 atcaaaaggg acacagcgcc gaggctagag gcaaccggat ggctgaccaa gcggcccgaa
2281 aggcagccat cacagagact ccagacacct ctaccctcct catagaaaat tcatcaccca
2341 attcccgctt aattaattaa gcttgcggcc gcactcgagc accaccacca ccaccactga
2401 gatccggctg ctaacaaagc ccgaaaggaa gctgagttgg ctgctgccac cgctgagcaa
2461 taactagcat aacccccttgg ggcctctaaa cgggtcttga ggggtttttt gctgaaagga
2521 ggaactatat ccggattggc gaatgggacg cgccctgtag cggcgcatta agcgcggcgg
2581 gtgtggtggt tacgcgcagc gtgaccgcta cacttgccag cgccctagcg cccgctcctt
2641 tcgctttctt cccttcctttt ctcgccacgt tcgccggctt tccccgtcaa gctctaaatc
2701 gggggctccc tttagggttc cgatttagtg ctttacggca cctcgacccc aaaaaacttg
2761 attagggtga tggttcacgt agtgggccat cgccctgata gacggttttt cgccctttga
2821 cgttggagtc cacgttcttt aatagtggac tcttgttcca aactggaaca acactcaacc
2881 ctatctcggt ctattctttt gatttataag ggattttgcc gatttcggcc tattggttaa
2941 aaaatgagct gatttaacaa aaatttaacg cgaattttaa caaaatatta acgtttacaa
3001 tttcaggtgg cacttttcgg ggaaatgtgc gcggaacccc tatttgttta ttttctaaa
3061 tacattcaaa tatgtatccg ctcatgagac aataaccctg ataaatgctt caataatatt
3121 gaaaaaggaa gagtatgagt attcaacatt tccgtgtcgc ccttattccc ttttttgcgg
3181 cattttgcct tcctgttttt gctcacccag aaacgctggt gaaagtaaaa gatgctgaag
3241 atcagttggg tgcacgagtg ggttacatcg aactggatct caacagcggt aagatccttg
3301 agagttttcg ccccgaagaa cgttttccaa tgatgagcac ttttaaagtt ctgctatgtg
3361 gcgcggtatt atcccgtatt gacgccgggc aagagcaact cggtcgccgc atacactatt
[1044] 3421 ctcagaatga cttggttgag tactcaccag tcacagaaaa gcatcttacg gatggcatga
3481 cagtaagaga attatgcagt gctgccataa ccatgagtga taacactgcg gccaacttac
3541 ttctgacaac gatcggagga ccgaaggagc taaccgcttt tttgcacaac atgggggatc
3601 atgtaactcg ccttgatcgt tgggaaccgg agctgaatga agccatacca aacgacgagc
3661 gtgacaccac gatgcctgca gcaatggcaa caacgttgcg caaactatta actggcgaac
3721 tacttactct agcttcccgg caacaattaa tagactggat ggaggcggat aaagttgcag
3781 gaccacttct gcgctcggcc cttccggctg gctggtttat tgctgataaa tctggagccg
3841 gtgagcgtgg gtctcgcggt atcattgcag cactggggcc agatggtaag ccctcccgta
3901 tcgtagttat ctacacgacg gggagtcagg caactatgga tgaacgaaat agacagatcg
3961 ctgagatagg tgcctcactg attaagcatt ggtaactgtc agaccaagtt tactcatata
4021 tactttagat tgatttaaaa cttcattttt aatttaaaag gatctaggtg aagatccttt
4081 ttgataatct catgaccaaa atcccttaac gtgagttttc gttccactga gcgtcagacc
4141 ccgtagaaaa gatcaaagga tcttcttgag atccttttttt tctgcgcgta atctgctgct
4201 tgcaaacaaa aaaaccaccg ctaccagcgg tggtttgttt gccggatcaa gagctaccaa
4261 ctcttttttcc gaaggtaact ggcttcagca gagcgcagat accaaatact gtccttctag
4321 tgtagccgta gttaggccac cacttcaaga actctgtagc accgcctaca tacctcgctc
4381 tgctaatcct gttaccagtg gctgctgcca gtggcgataa gtcgtgtctt accgggttgg
4441 actcaagacg atagttaccg gataaggcgc agcggtcggg ctgaacgggg ggttcgtgca
4501 cacagcccag cttggagcga acgacctaca ccgaactgag atacctacag cgtgagctat
4561 gagaaagcgc cacgcttccc gaagggagaa aggcggacag gtatccggta agcggcaggg
4621 tcggaacagg agagcgcacg agggagcttc cagggggaaa cgcctggtat ctttatagtc
4681 ctgtcgggtt tcgccacctc tgacttgagc gtcgattttt gtgatgctcg tcaggggggc
4741 ggagcctatg gaaaaacgcc agcaacgcgg cctttttacg gttcctggcc ttttgctggc
4801 cttttgctca catgttcttt cctgcgttat cccctgattc tgtggataac cgtattaccg
4861 cctttgagtg agctgatacc gctcgccgca gccgaacgac cgagcgcagc gagtcagtga
4921 gcgaggaagc ggaagagcgc ctgatgcggt attttctcct tacgcatctg tgcggtattt
4981 cacaccgcat atatggtgca ctctcagtac aatctgctct gatgccgcat agttaagcca
5041 gtatacactc cgctatcgct acgtgactgg gtcatggctg cgccccgaca cccgccaaca
5101 cccgctgacg cgccctgacg ggcttgtctg ctcccggcat ccgcttacag acaagctgtg
5161 accgtctccg ggagctgcat gtgtcagagg ttttcaccgt catcaccgaa acgcgcgagg

[1045]
```
5221 cagctgcggt aaagctcatc agcgtggtcg tgaagcgatt cacagatgtc tgcctgttca
5281 tccgcgtcca gctcgttgag tttctccaga agcgttaatg tctggcttct gataaagcgg
5341 gccatgttaa gggcggtttt ttcctgtttg gtcactgatg cctccgtgta agggggattt
5401 ctgttcatgg gggtaatgat accgatgaaa cgagagagga tgctcacgat acgggttact
5461 gatgatgaac atgcccggtt actggaacgt tgtgagggta aacaactggc ggtatggatg
5521 cggcgggacc agagaaaaat cactcagggt caatgccagc gcttcgttaa tacagatgta
5581 ggtgttccac agggtagcca gcagcatcct gcgatgcaga tccggaacat aatggtgcag
5641 ggcgctgact tccgcgtttc cagactttac gaaacacgga aaccgaagac cattcatgtt
5701 gttgctcagg tcgcagacgt tttgcagcag cagtcgcttc acgttcgctc gcgtatcggt
5761 gattcattct gctaaccagt aaggcaaccc cgccagccta gccgggtcct caacgacagg
5821 agcacgatca tgcgcacccg tggggccgcc atgccggcga taatggcctg cttctcgccg
5881 aaacgtttgg tggcgggacc agtgacgaag gcttgagcga gggcgtgcaa gattccgaat
5941 accgcaagcg acaggccgat catcgtcgcg ctccagcgaa agcggtcctc gccgaaaatg
6001 acccagagcg ctgccggcac ctgtcctacg agttgcatga taaagaagac agtcataagt
6061 gcggcgacga tagtcatgcc ccgcgcccac cggaaggagc tgactgggtt gaaggctctc
6121 aagggcatcg gtcgagatcc cggtgcctaa tgagtgagct aacttacatt aattgcgttg
6181 cgctcactgc ccgctttcca gtcgggaaac ctgtcgtgcc agctgcatta atgaatcggc
6241 caacgcgcgg ggagaggcgg tttgcgtatt gggcgccagg gtggtttttc ttttcaccag
6301 tgagacgggc aacagctgat tgcccttcac cgcctggccc tgagagagtt gcagcaagcg
6361 gtccacgctg gtttgcccca gcaggcgaaa atcctgtttg atggtggtta acggcgggat
6421 ataacatgag ctgtcttcgg tatcgtcgta tcccactacc gagatatccg caccaacgcg
6481 cagcccggac tcggtaatgg cgcgcattgc gcccagcgcc atctgatcgt tggcaaccag
6541 catcgcagtg ggaacgatgc cctcattcag catttgcatg gtttgttgaa aaccggacat
6601 ggcactccag tcgccttccc gttccgctat cggctgaatt tgattgcgag tgagatattt
6661 atgccagcca gccagacgca gacgcgccga gacagaactt aatgggcccg ctaacagcgc
6721 gatttgctgg tgacccaatg cgaccagatg ctccacgccc agtcgcgtac cgtcttcatg
6781 ggagaaaata atactgttga tgggtgtctg gtcagagaca tcaagaaata acgccggaac
6841 attagtgcag gcagcttcca cagcaatggc atcctggtca tccagcggat agttaatgat
6901 cagcccactg acgcgttgcg cgagaagatt gtgcaccgcc gctttacagg cttcgacgcc
6961 gcttcgttct accatcgaca ccaccacgct ggcacccagt tgatcggcgc gagatttaat
7021 cgccgcgaca atttgcgacg gcgcgtgcag ggccagactg gaggtggcaa cgccaatcag
7081 caacgactgt ttgcccgcca gttgttgtgc cacgcggttg ggaatgtaat tcagctccgc
7141 catcgccgct tccacttttt cccgcgtttt cgcagaaacg tggctggcct ggttcaccac
7201 gcgggaaacg gtctgataag agacaccggc atactctgcg acatcgtata acgttactgg
7261 tttcacattc accaccctga attgactctc ttccgggcgc tatcatgcca taccgcgaaa
7321 ggttttgcgc cattcgatgg tgtccgggat ctcgacgctc tcccttatgc gactcctgca
7381 ttaggaagca gcccagtagt aggttgaggc cgttgagcac cgccgccgca aggaatggtg
7441 catgcaagga gatggcgccc aacagtcccc cggc
```

[1046] SEQ ID No:20

[1047] 寡核苷酸(19b)

[1048] pD-ter-

[1049] 5’-AAGAAGACACGATCCACCG-3’

[1050] SEQ ID No:21

[1051] 寡核苷酸(35b)

[1052] long+

[1053] 5’-CGAACGTGGCGAGAAAGGAAGGGAAGAAAGAAGTC-3’

[1054] SEQ ID No:22

[1055] 寡核苷酸(71b+3’修饰ddC)

[1056] ddC-Long2

[1057] 5’-TTTTTTTAGACTTCTTTCTTCCCTTCCTTTCTCGCCACGTTCGAAGAAGACACGATCCACCGCCGGTTCCG-ddC-3’

[1058] SEQ ID No:23

[1059] 寡核苷酸(35b+3’生物素(TEG)

[1060] Long+Tb

[1061] 5'-CGAACGTGGCGAGAAAGGAAGGGAAGAAAGAAGTC-Bio-3'

[1062] 参考文献

[1063] Mattheakis LC,Bhatt RR,Dower WJ.An in vitro polysome display system for identifying ligands from very large peptide libraries.Proc Natl Acad Sci USA.1994Sep13;91(19):9022-6.

[1064] Matsuura T,Pluckthun A.Selection based on the folding properties of proteins with ribosome display.FEES Lett.2003Mar 27;539(1-3):24-8.

[1065] Hanes J,Pluckthun A.In vitro selection and evolution of functional proteins by using ribosome display.Proc Natl Acad Sci USA.1997May 13;94(10):4937-42.

[1066] He M,Taussig MJ.Antibody-ribosome-mRNA(ARM)complexes as efficient selection particles for in vitro display and evolution of antibody combining sites.Nucleic Acids Res.1997Dec 15;25(24):5132-4.

[1067] Irving RA,Coia G,Roberts A,Nuttall SD,Hudson PJ.Ribosome display and affinity maturation:from antibodies to single V-domains and steps towards cancer therapeutics.J Immunol Methods.2001Feb 1;248(1-2):31-45.Review.

[1068] Amstutz P,Pelletier JN,Guggisberg A,Jermutus L,Cesaro-Tadic S,Zahnd C,Pluckthun A.In vitro selection for catalytic activity with ribosome display.J Am Chem Soc.2002Aug 14;124(32):9396-403.

[1069] Takahashi F,Ebihara T,Mie M,Yanagida Y,Endo Y,Kobatake E,Aizawa M.Ribosome display for selection of active dihydrofolate reductase mutants using immobilized methotrexate on agarose beads.FEBS Lett.2002Mar 6;514(1):106-10.

[1070] Tawfik DS,Griffiths AD.Man-made cell-like compartments for molecular evolution.Nat Biotechnol.1998Jul;16(7):652-6.

[1071] Lee YF,Tawfik DS,Griffiths AD.Investigating the target recognition of DNA cytosine-5methyltransferase Hhal by library selection using in vitro compartmentalisation.Nucleic Acids Res.2002Nov 15;30(22):4937-44.

[1072] Cohen HM,Tawfik DS,Griffiths AD.Altering the sequence specificity of HaeIII methyltransferase by directed evolution using in vitro compartmentalization.

[1073] Protein Eng Des Sel.2004Jan;17(1):3-11.

[1074] Ghadessy FJ,Ong JL,Holliger P.Directed evolution of polymerase function by compartmentalized self-replication.Proc Natl Acad Sci USA.2001Apr 10;98(8):4552-7.Epub 2001Mar 27.

[1075] Ghadessy FJ,Ramsay N,Boudsocq F,Loakes D,Brown A,Iwai S,Vaisman A,Woodgate R,Holliger P.Generic expansion of the substrate spectrum of a DNA polymerase by directed evolution.Nat Biotechnol.2004Jun;22(6):755-9.Epub

2004May 23.

[1076] Ong JL,Loakes D,Jaroslawski S,Too K,Holliger P.Directed evolution of DNA polymerase,RNA polymerase and reverse transcriptase activity in a single polypeptide.J Mol Biol.2006Aug 18;361(3):537-50.Epub 2006Jμl 5.

[1077] Bernath K,Hai M,Mastrobattista E,Griffiths AD,Magdassi S,Tawfik DS.In vitro compartmentalization by double emulsions:sorting and gene enrichment by fluorescence activated cell sorting.Anal Biochem.2004年2月1日;325(1):151-7.

[1078] Mastrobattista E,Taly V,Chanudet E,Treacy P,Kelly BT,Griffiths AD.High-throughput screening of enzyme libraries:in vitro evolution of a beta-galactosidase by fluorescence-activated sorting of double emulsions.Chem Biol.2005年12月;12(12):1291-300.

[1079] Bertschinger J,Neri D.Covalent DNA display as a novel tool for directed evolution of proteins in vitro.Protein Eng Des Sel.2004年9月;17(9):699-707.Epub 2004年11月2日。

[1080] Doi N,Yanagawa H.STABLE:protein-DNA fusion system for screening of combinatorial protein libraries in vitro.FEBS Lett.1999年8月27日;457(2):227-30.

[1081] Yonezawa M,Doi N,Kawahashi Y,Higashinakagawa T,Yanagawa H.DNA display for in vitro selection of diverse peptide libraries.Nucleic acids Res.2003年11月1日;31(19):e118.

[1082] Sepp A,Choo Y.Cell-free selection of zinc finger DNA-binding proteins using in vitro compartmentalization.J Mol Biol.2005年11月25日;354(2):212-9.Epub 2005年10月3日.

[1083] Sepp A,Tawfik DS,Griffiths AD.Microbead display by in vitro compartmentalisation:selection for binding using flow cytometry.FEBS Lett.2002年12月18日;532(3):455-8.Griffiths AD,Tawfik DS.Directed evolution of an extremely fast phosphotriesterase by in vitro compartmentalization.EMBO J.2003年1月2日;22(1):24-35.

[1084] Bernath K,Magdassi S,Tawfik DS.Directed evolution of protein inhibitors of DNA-nucleases by in vitro compartmentalization(IVC)and nano-droplet delivery.J Mol Biol.2005年2月4日;345(5):1015-26.Epub 2004年12月7日。

[1085] Thorsen T,Roberts RW,Arnold FH,Quake SR.Dynamic pattern formation in a vesicle-generating microfiuidic device.Phys Rev Lett.2001年4月30日;86(18):4163-6.Okushima S,Nisisako T,Torii T,Higuchi T.Controlled production of monodisperse double emulsions by two-step droplet breakup in microfiuidic devices.Langmuir.2004年11月9日;20(23):9905-8.

[1086] Song H,Tice JD,Ismagilov RF.A microfluidic system for controlling reaction networks in time.Angew Chem Int Ed Engl.2003年2月17日;42(7):768-72.

[1087] Link DR,Grasland-Mongrain E,Duri A,Sarrazin F,Cheng Z,Cristobal G,

Marquez M,Weitz DA.Electric control of droplets in microfluidic devices.Angew Chem Int Ed Engl.2006年4月10日;45(16):2556-60.

[1088] Matsuura T,Yanagida H,Ushioda J,Urabe I,Yomo T.Nascent chain,mRNA,and ribosome complexes generated by a pure translation system.Biochem Biophys Res Commun.2007年1月12日;352(2):372-7.Epub 2006年11月17日。

[1089] Gerard GF,D'Alessio JM,Kotewicz ML,Noon MC.Influence on stability in Escherichia coli of the carboxy-terminal structure of cloned Moloney murine leukemia virus reverse transcriptase.DNA.1986年8月;5(4):271-9.

[1090] Forrer P,Jaussi R.High-level expression of soluble heterologous proteins in the cytoplasm of Escherichia coli by fusion to the bacteriophage lambda head protein D.Gene.1998年12月11日;224(1-2):45-52.

[1091] Gerard GF,Potter RJ,Smith MD,Rosenthal K,Dhariwal G,Lee J,Chatterjee DK.The role of template-primer in protection of reverse transcriptase from thermal inactivation.Nucleic Acids Res.2002年7月15日;30(14):3118-29.

[1092] Vichier-Guerre S,Ferris S,Auberger N,Mahiddine K,Jestin JL.A population of thermostable reverse transcriptases evolved from Thermus aquaticus DNA polymerase I by phage display.Angew Chem Int Ed Engl.2006年9月18日;45(37):6133-7.

[1093] Ghadessy FJ,Holliger P.A novel emulsion mixture for in vitro compartmentalization of transcription and translation in the rabbit reticulocyte system.Protein Eng Des Sel.2004年3月;17(3):201-4.Epub 2004年2月27日。

[1094] Roberts RW,Szostak JW.RNA-peptide fusions for the in vitro selection of peptides and proteins.Proc Natl Acad Sci USA.1997年11月11日;94(23):12297-302.

[1095] Odegrip R,Coomber D,Eldridge B,Hederer R,Kuhlman PA,Ullman C,FitzGerald K,McGregor D.CIS display:In vitro selection of peptides from libraries of protein-DNA complexes.Proc Natl Acad Sci USA.2004年3月2日;101(9):2806-10.Epub 2004年2月23日.

[1096] Reiersen H,Løbersli I,LøsetGA,Hvattum E,Simonsen B,Stacy JE,McGregor D,Fitzgerald K,Welschof M,Brekke OH,Marvik OJ.Covalent antibody display—an in vitro antibody-DNA library selection system.Nucleic Acids Res.2005年1月14日;33(1):e10.Bertschinger J,Neri D.Covalent DNA display as a novel tool for directed evolution of proteins in vitro.Protein Eng Des Sel.2004年9月;17(9):699-707.Epub 2004年11月2日。

[1097] Stein V,Sielaff I,Johnsson K,Hollfelder F.A covalent chemical genotype-phenotype linkage for in vitro protein evolution.Chembiochem.2007年12月17日;8(18):2191-4.

[1098] Tabuchi I,Soramoto S,Nemoto N,Husimi Y.An in vitro DNA virus for in

vitro protein evolution.FEBS Lett.2001年1月23日;508(3):309-12.

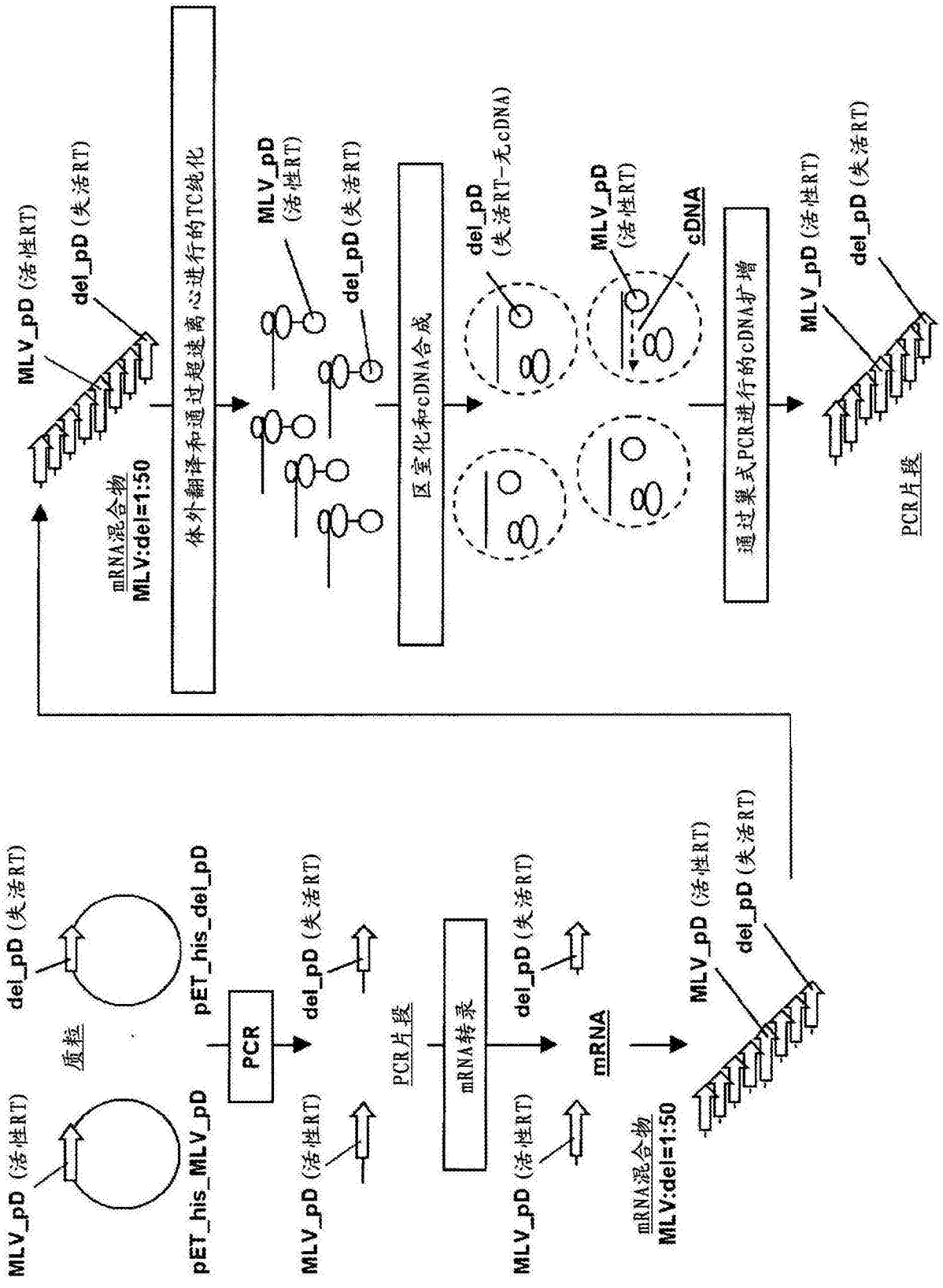

MLV_pD (活性RT)
del_pD (失活RT)
质粒
pET_his_MLV_pD
pET_his_del_pD
PCR
PCR片段
mRNA转录
mRNA
mRNA混合物
MLV:del=1:50
体外翻译和通过超速离心进行的TC纯化
区室化和cDNA合成
del_pD (失活RT-无cDNA)
cDNA
通过巢式PCR进行的cDNA扩增

图1

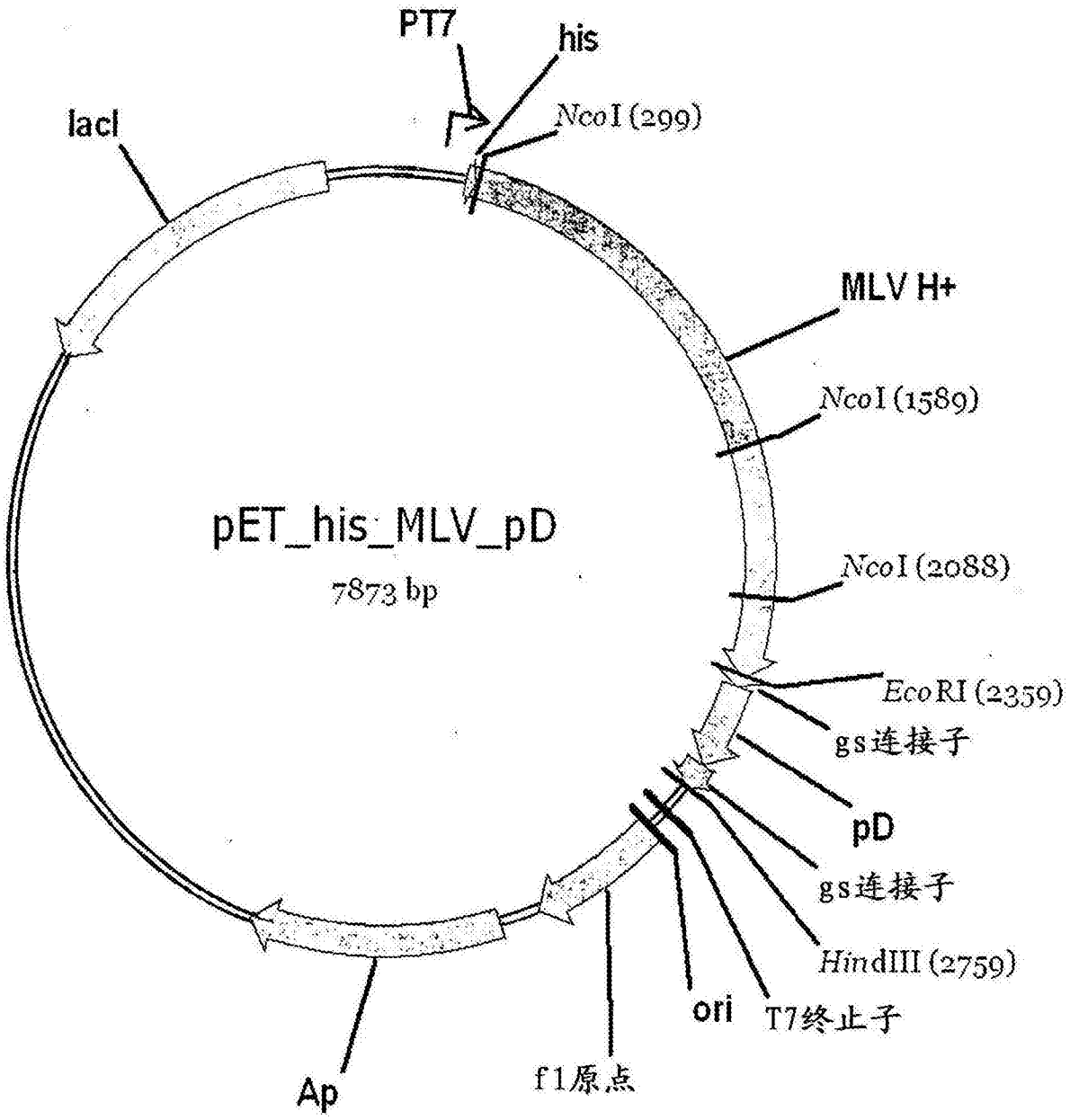

PT7
his
NcoI (299)
lacI
MLV H+
NcoI (1589)
pET_his_MLV_pD
7873 bp
NcoI (2088)
EcoRI (2359)
gs连接子
pD
gs连接子
HindIII (2759)
ori
T7终止子
Ap
f1原点

图2

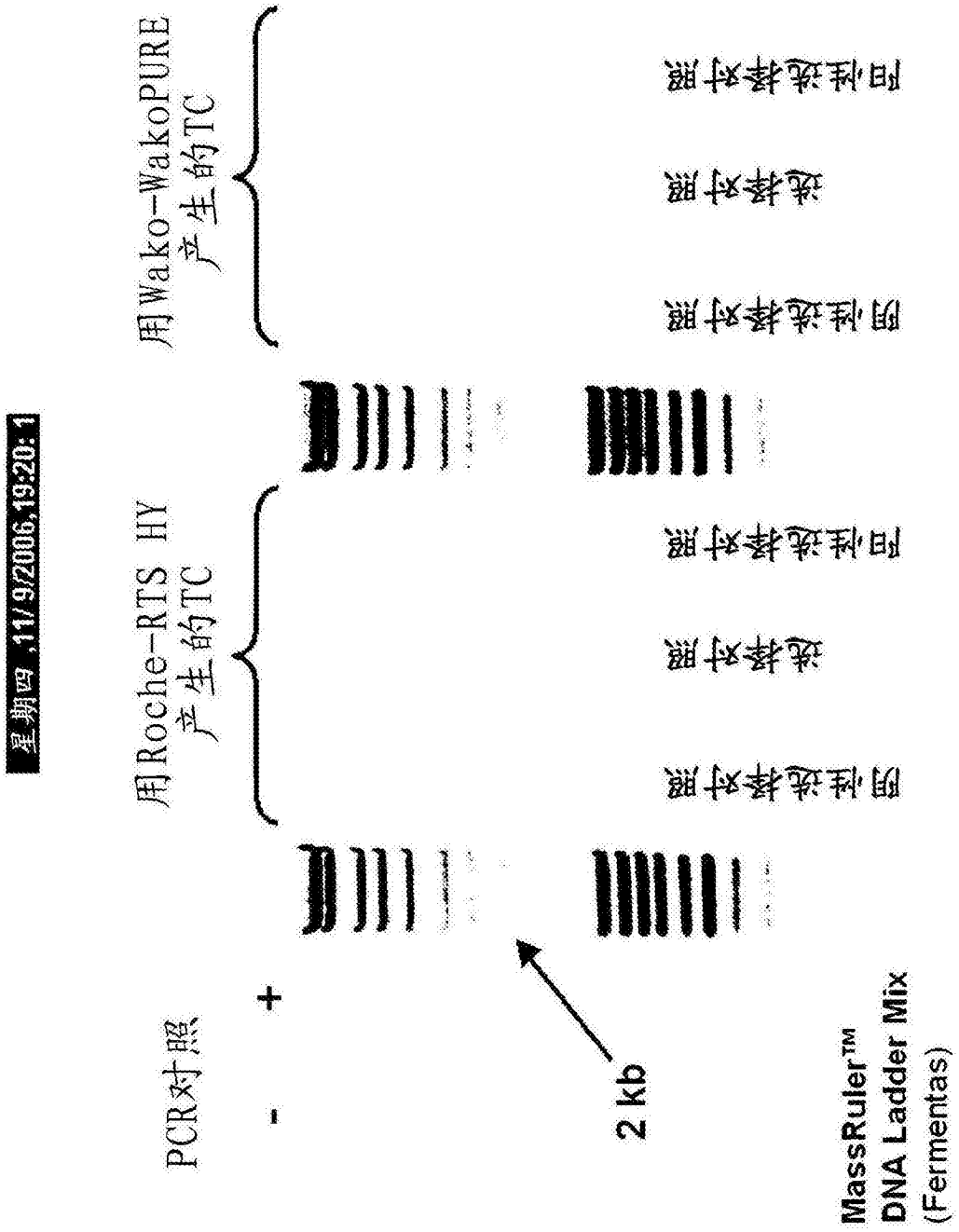

星期四 ,11/9/2006,19:20:1
用Wako-WakoPURE
产生的TC
用Roche-RTS HY
产生的TC
阴性选择对照
选择对照
阳性选择对照
阴性选择对照
选择对照
阳性选择对照
PCR对照
+
-
2 kb
MassRuler™
DNA Ladder Mix
(Fermentas)

图3

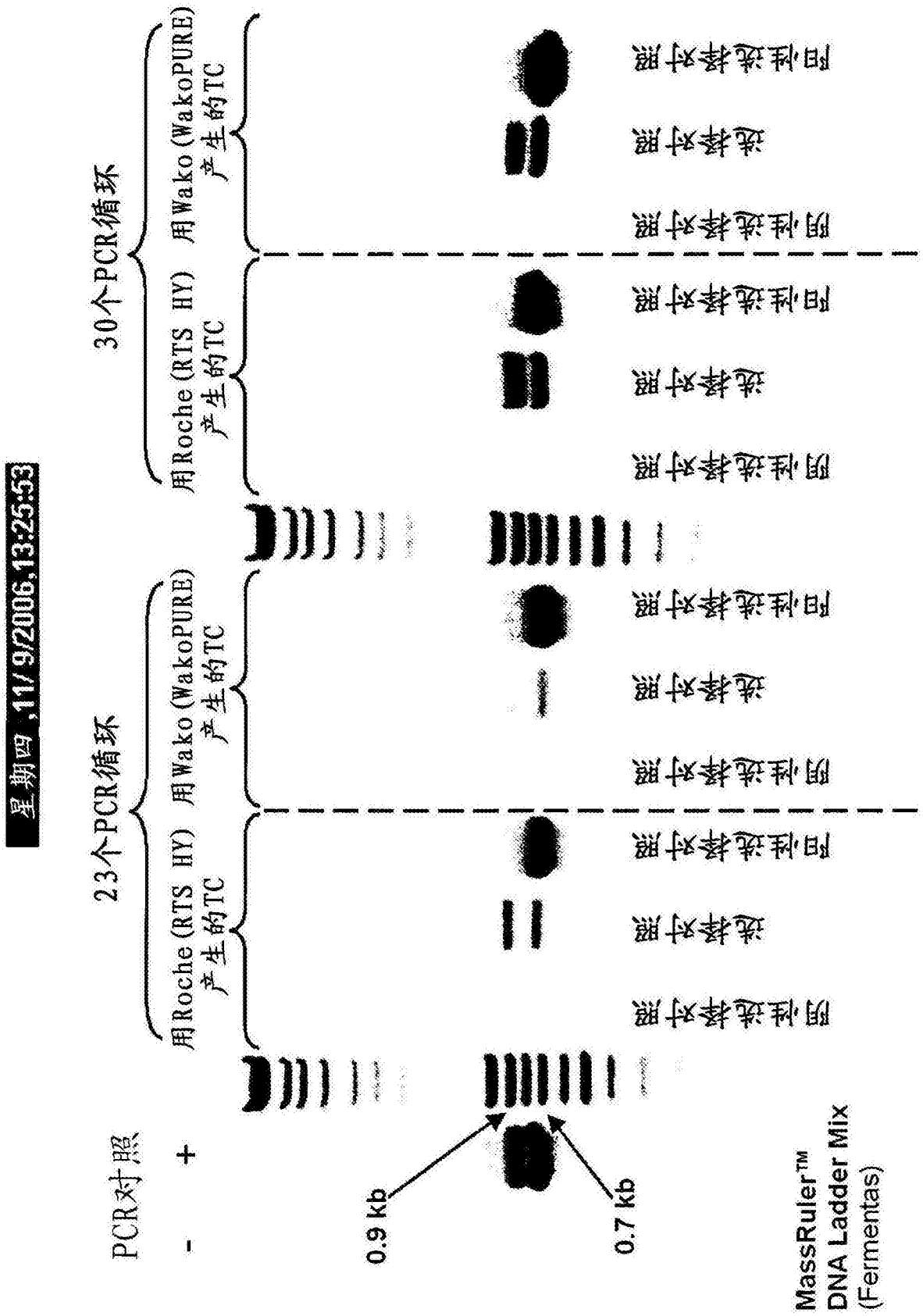

星期四 ,11/ 9/2006,13:25:53
PCR对照
-
+
0.9 kb
0.7 kb
MassRuler™
DNA Ladder Mix
(Fermentas)
23个PCR循环
用Roche (RTS HY) 产生的TC
用Wako (WakoPURE) 产生的TC
30个PCR循环
用Roche (RTS HY) 产生的TC
用Wako (WakoPURE) 产生的TC
阴性选择对照
选择对照
阳性选择对照

图4

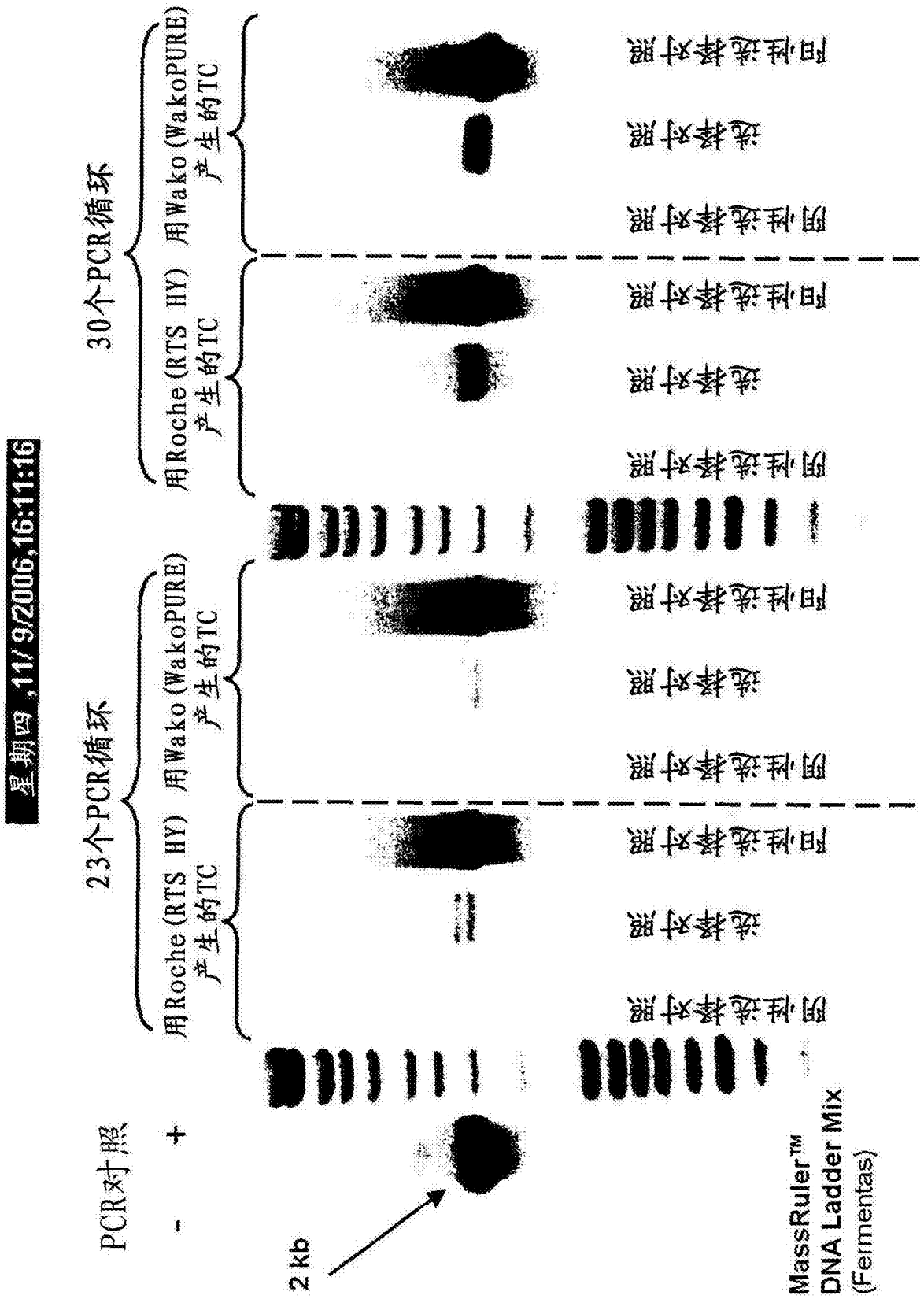

星期四 ,11/ 9/2006,16:11:16
PCR对照
- +
23个PCR循环
用Roche (RTS HY) 产生的TC
用Wako (WakoPURE) 产生的TC
30个PCR循环
用Roche (RTS HY) 产生的TC
用Wako (WakoPURE) 产生的TC
阴性选择对照
选择对照
阳性选择对照
2 kb
MassRuler™
DNA Ladder Mix
(Fermentas)

图5

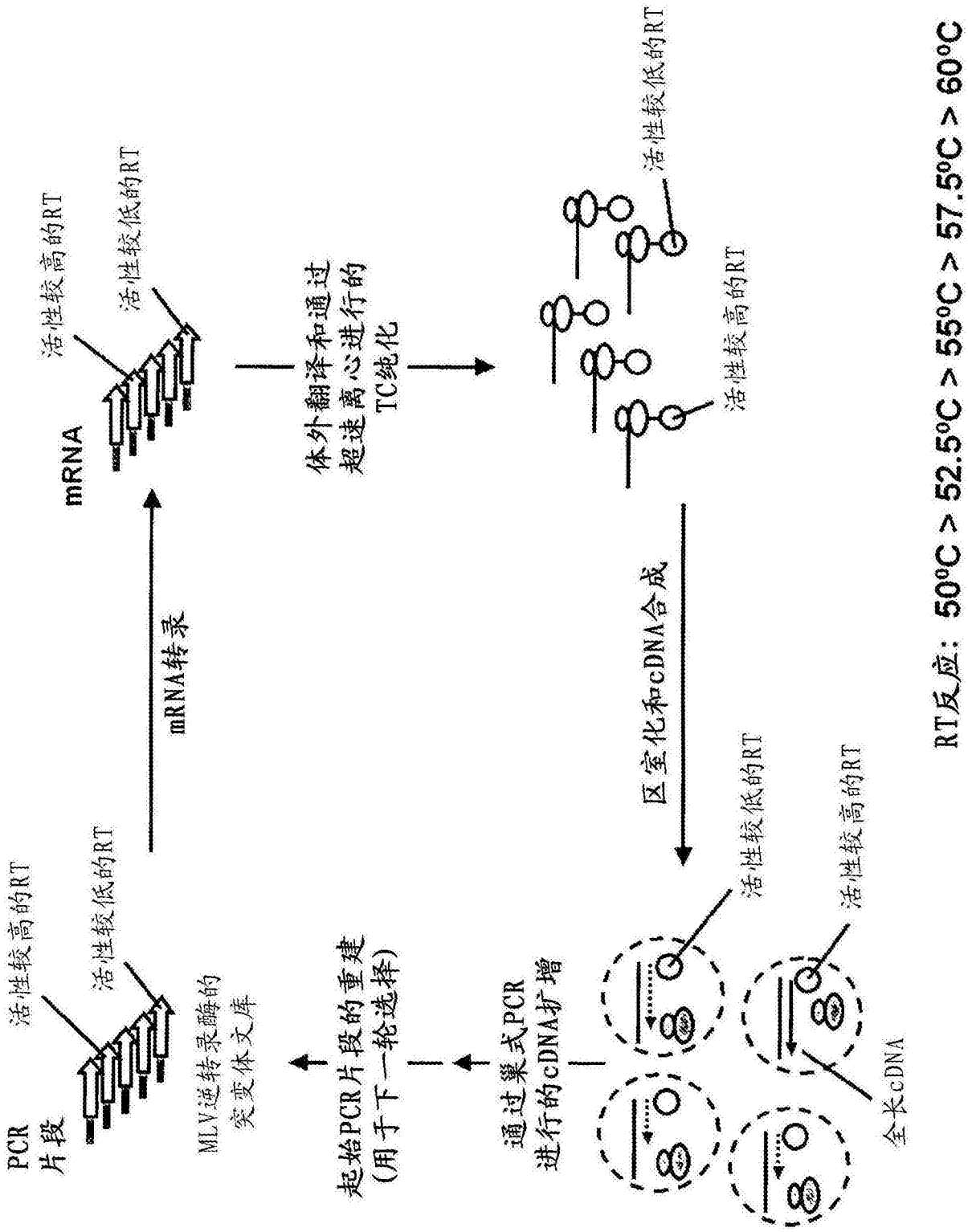

PCR
片段
活性较高的RT
活性较低的RT
MLV逆转录酶的
突变体文库
起始PCR片段的重建
(用于下一轮选择)
通过巢式PCR
进行的cDNA扩增
区室化和cDNA合成
全长cDNA
活性较低的RT
活性较高的RT
mRNA转录
mRNA
活性较高的RT
活性较低的RT
体外翻译和通过
超速离心进行的
TC纯化
活性较低的RT
活性较高的RT
RT反应: 50°C > 52.5°C > 55°C > 57.5°C > 60°C

图6

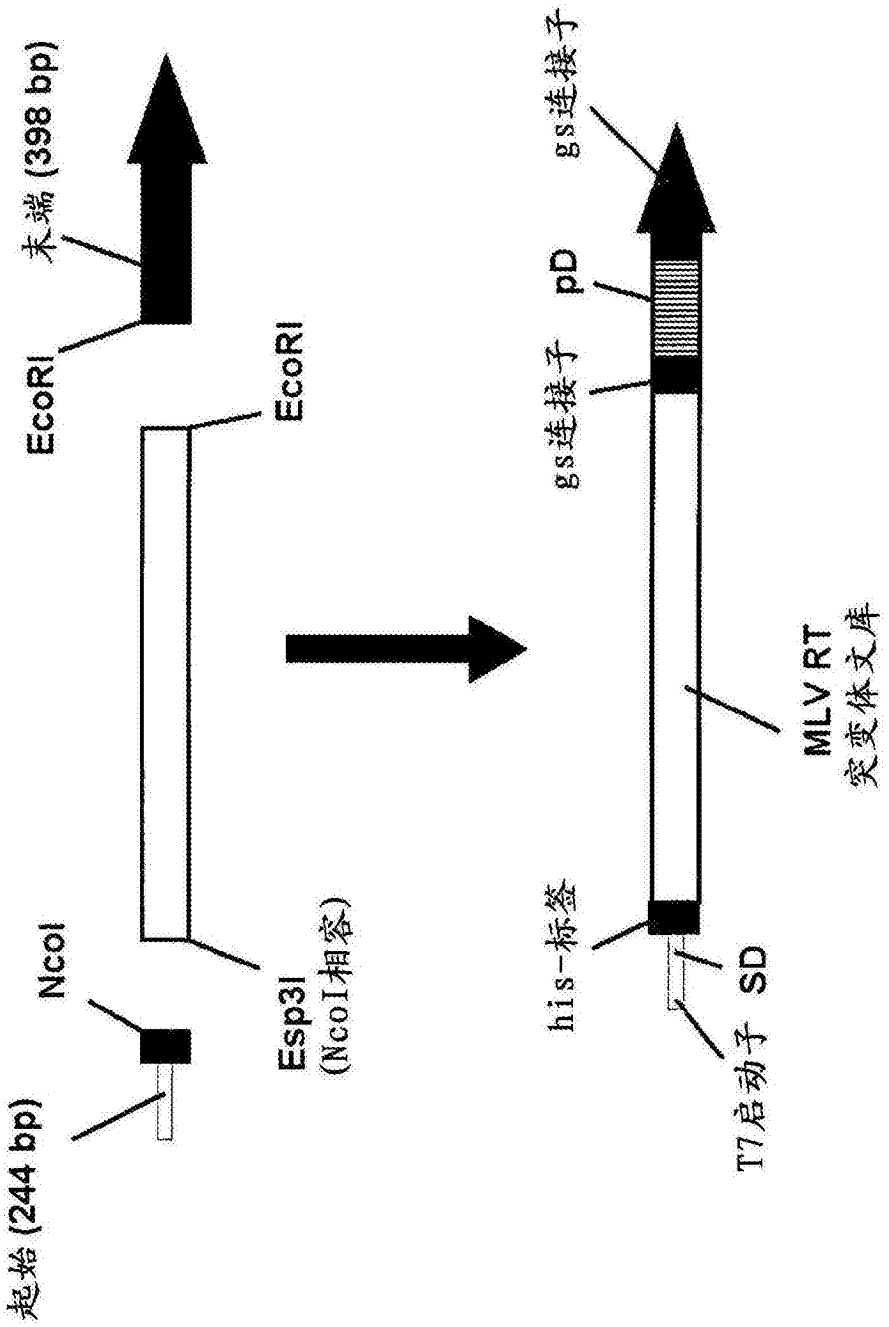

末端 (398 bp)
EcoRI
EcoRI
NcoI
起始 (244 bp)
Esp3I
(NcoI相容)
gs连接子
pD
gs连接子
MLV RT
突变体文库
his-标签
T7启动子
SD

图7

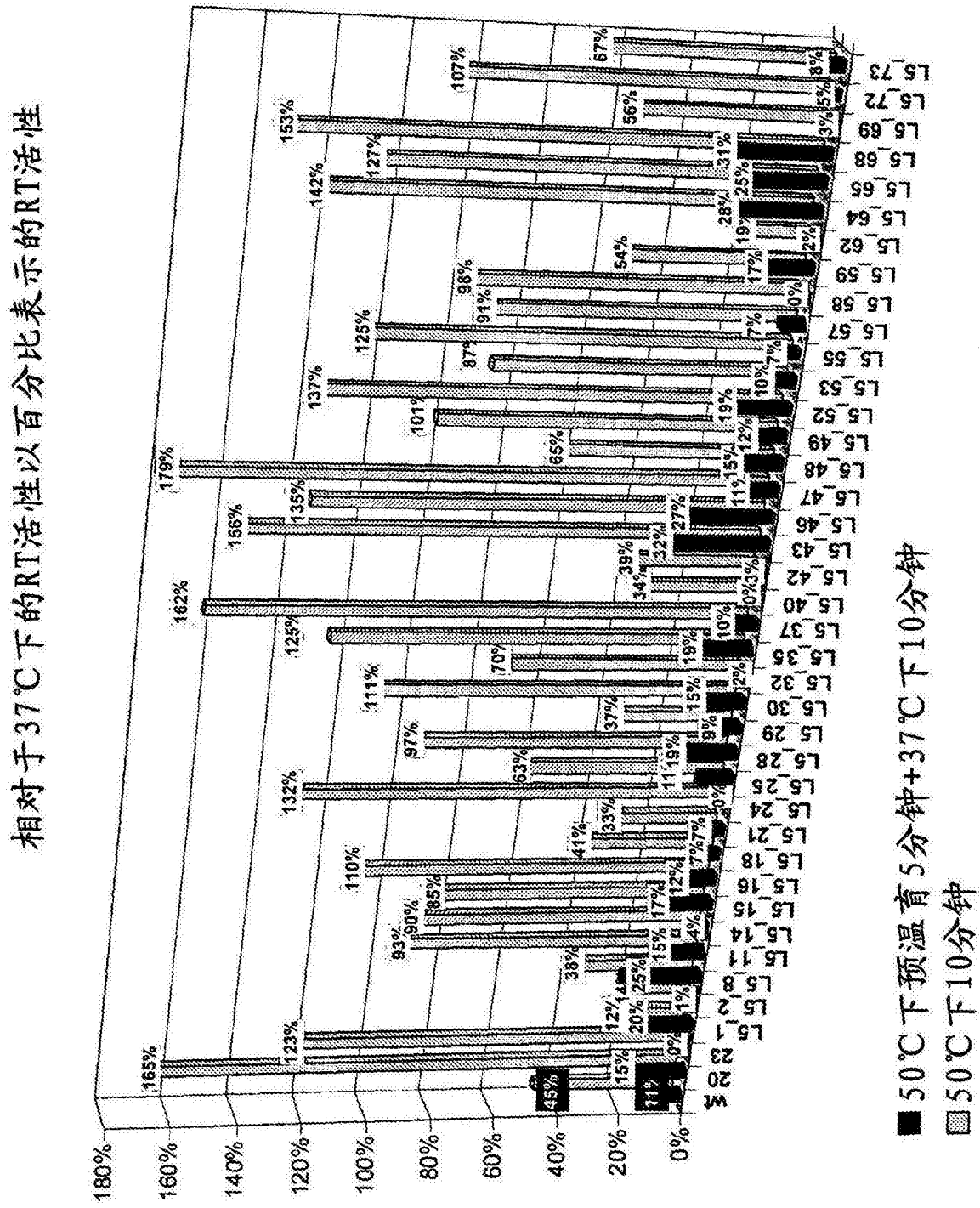

相对于37℃下的RT活性以百分比表示的RT活性
180%
160%
140%
120%
100%
80%
60%
40%
20%
0%
■ 50℃下预温育5分钟+37℃下10分钟
□ 50℃下10分钟

图8

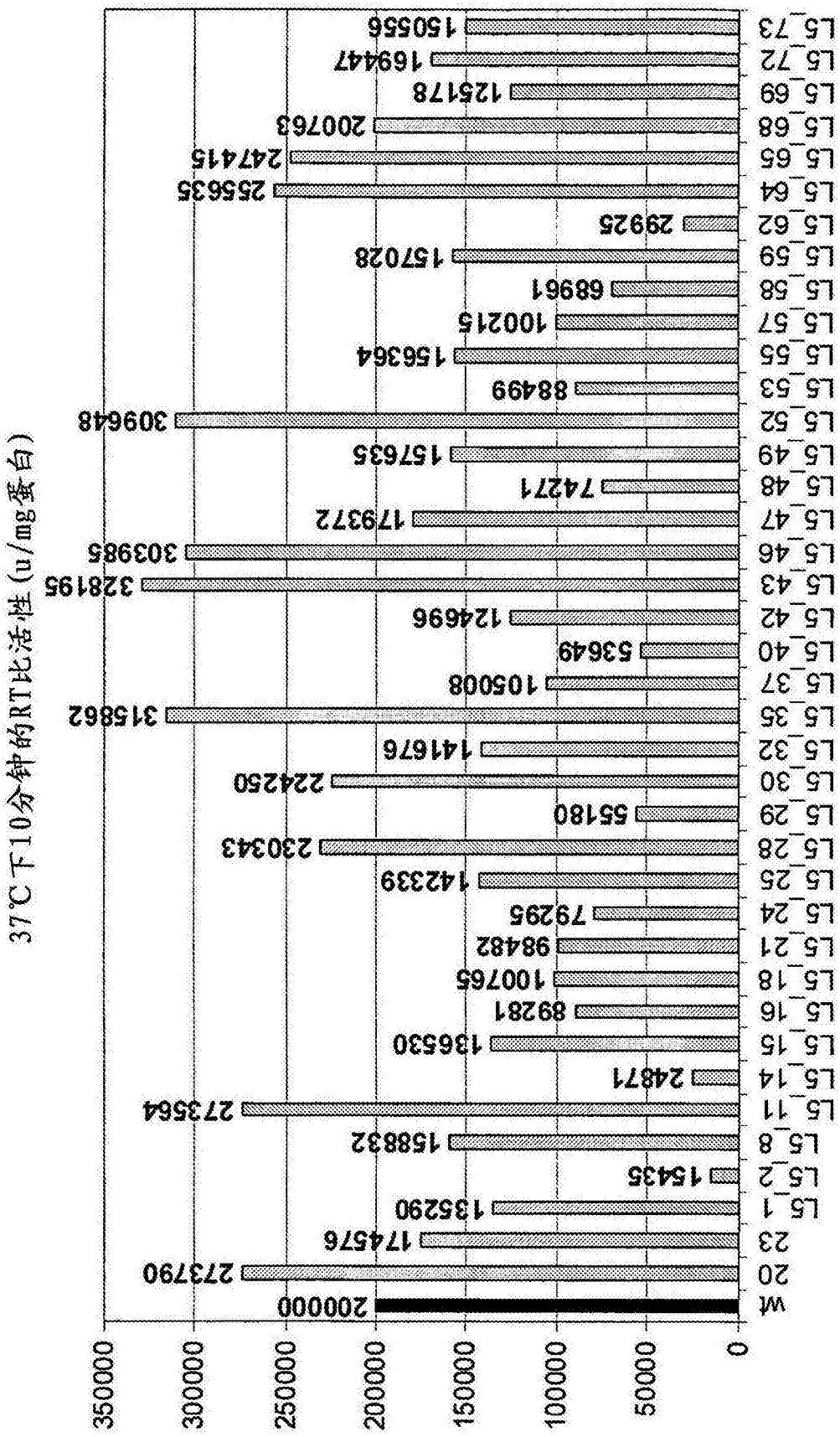

37℃下10分钟的RT比活性(u/mg蛋白)
350000
300000
250000
200000
150000
100000
50000
0
wt
20
23
L5_1
L5_2
L5_8
L5_11
L5_14
L5_15
L5_16
L5_18
L5_21
L5_24
L5_25
L5_28
L5_29
L5_30
L5_32
L5_35
L5_37
L5_40
L5_42
L5_43
L5_46
L5_47
L5_48
L5_49
L5_52
L5_53
L5_55
L5_57
L5_58
L5_59
L5_62
L5_64
L5_65
L5_68
L5_69
L5_72
L5_73
200000
273790
174576
135290
15435
158832
273564
24871
136530
89281
100765
98482
79295
142339
230343
55180
224250
141676
315862
105008
53649
124696
328195
303985
179372
74271
157635
309648
88499
156364
100215
68961
157028
299 25
255635
247415
200763
125178
169447
150556

图9

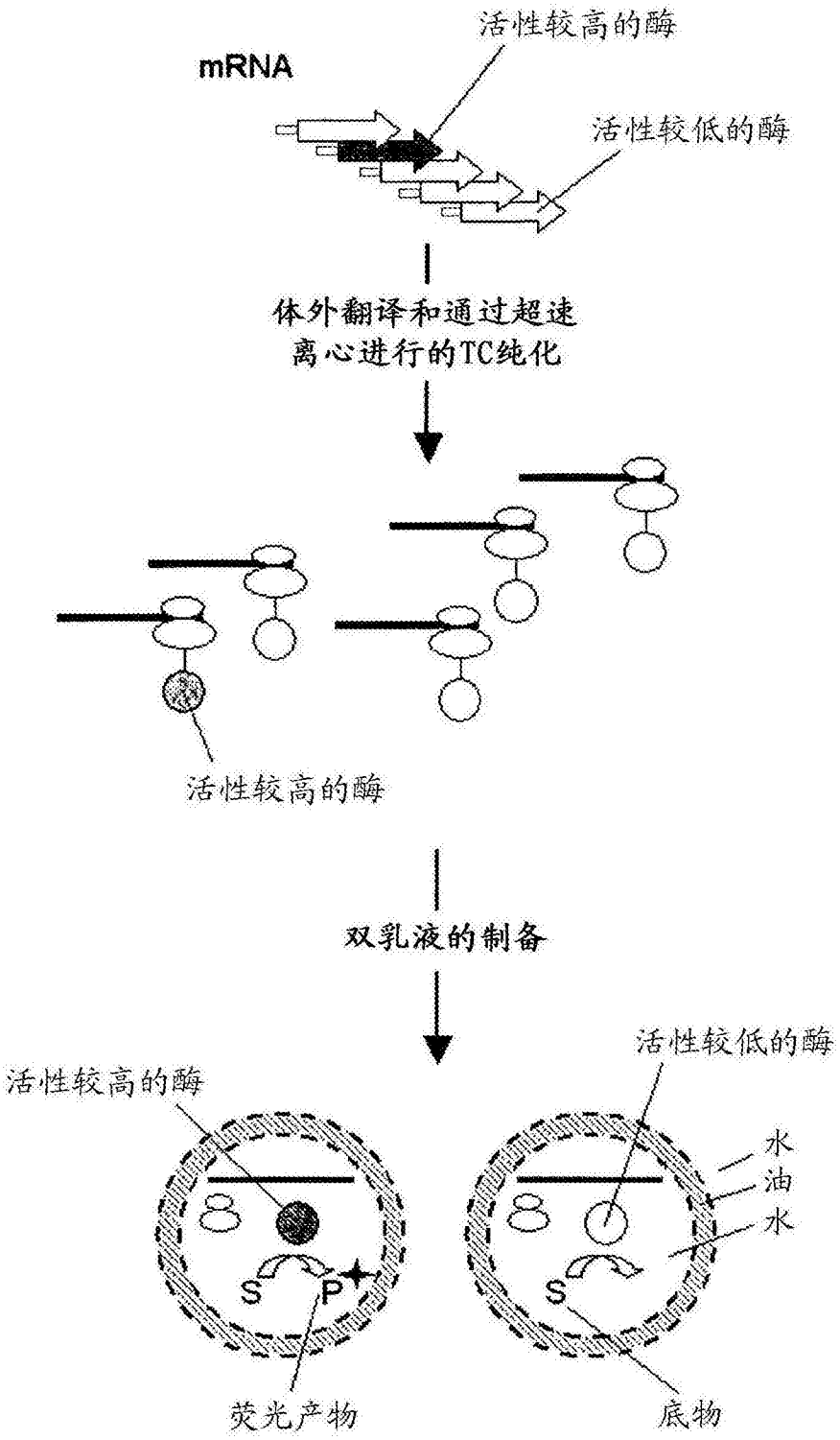

活性较高的酶
mRNA
活性较低的酶
体外翻译和通过超速
离心进行的TC纯化
活性较高的酶
双乳液的制备
活性较低的酶
活性较高的酶
水
油
水
S
P
S
荧光产物
底物
以FACS进行的分选

图10

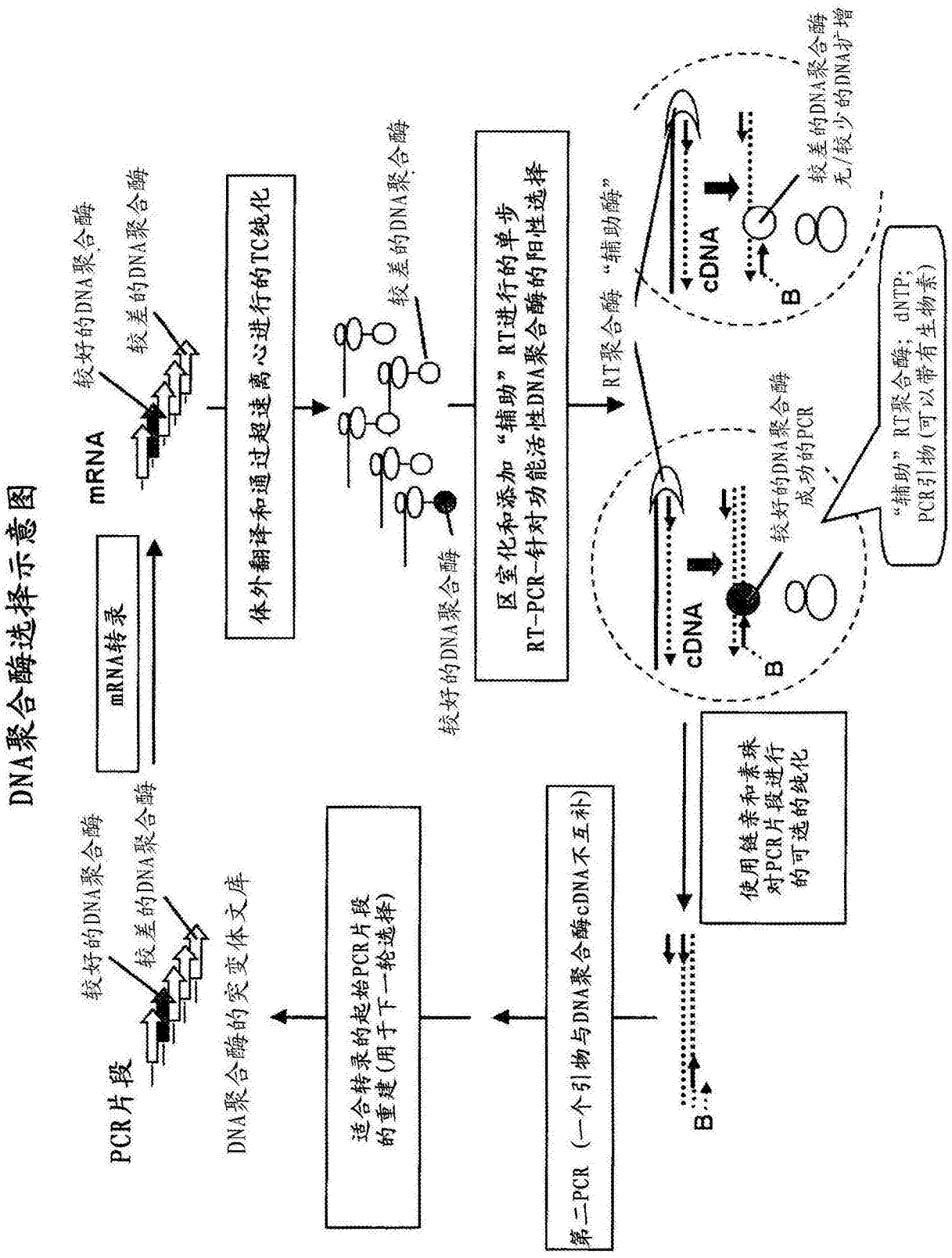

DNA聚合酶选择示意图
较好的DNA聚合酶
较差的DNA聚合酶
mRNA
mRNA转录
PCR片段
较好的DNA聚合酶
较差的DNA聚合酶
DNA聚合酶的突变体文库
体外翻译和通过超速离心进行的TC纯化
较差的DNA聚合酶
较好的DNA聚合酶
区室化和添加“辅助”RT进行的单步
RT-PCR-针对功能活性DNA聚合酶的阳性选择
RT聚合酶 “辅助酶”
cDNA
B
较差的DNA聚合酶
无/较少的DNA扩增
cDNA
B
较好的DNA聚合酶
成功的PCR
“辅助”RT聚合酶；dNTP；
PCR引物（可以带有生物素）
适合转录的起始PCR片段
的重建（用于下一轮选择）
第二PCR（一个引物与DNA聚合酶cDNA不互补）
使用链亲和素珠
对PCR片段进行
的可选的纯化
B

图11

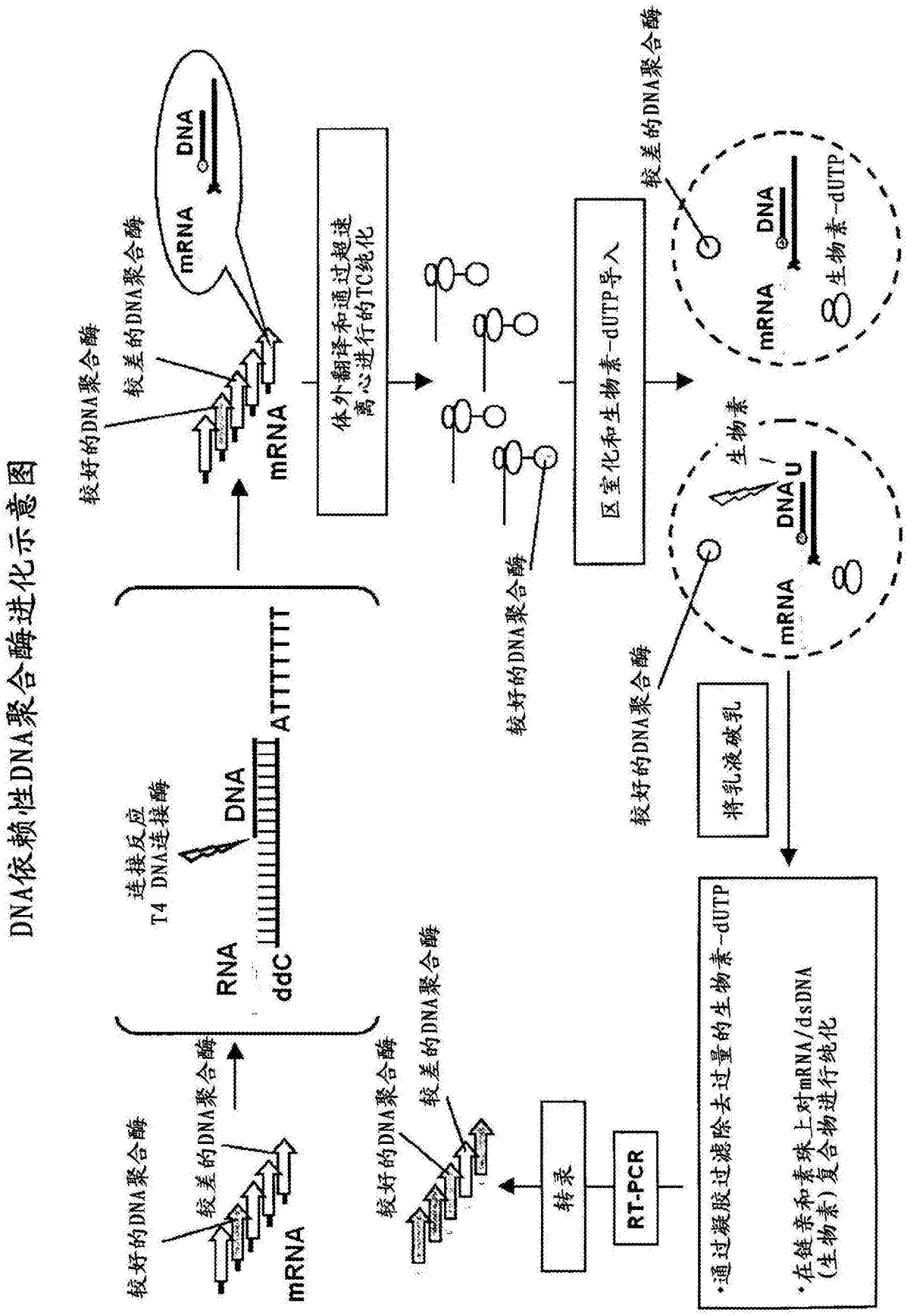

DNA依赖性DNA聚合酶进化示意图
较好的DNA聚合酶
较差的DNA聚合酶
mRNA
RNA
ddC
连接反应
T4 DNA连接酶
DNA
ATTTTTTT
体外翻译和通过超速离心进行的TC纯化
区室化和生物素-dUTP导入
生物素
U
将乳液破乳
转录
RT-PCR
·通过凝胶过滤除去过量的生物素-dUTP
·在链亲和素珠上对mRNA/dsDNA(生物素)复合物进行纯化

图12

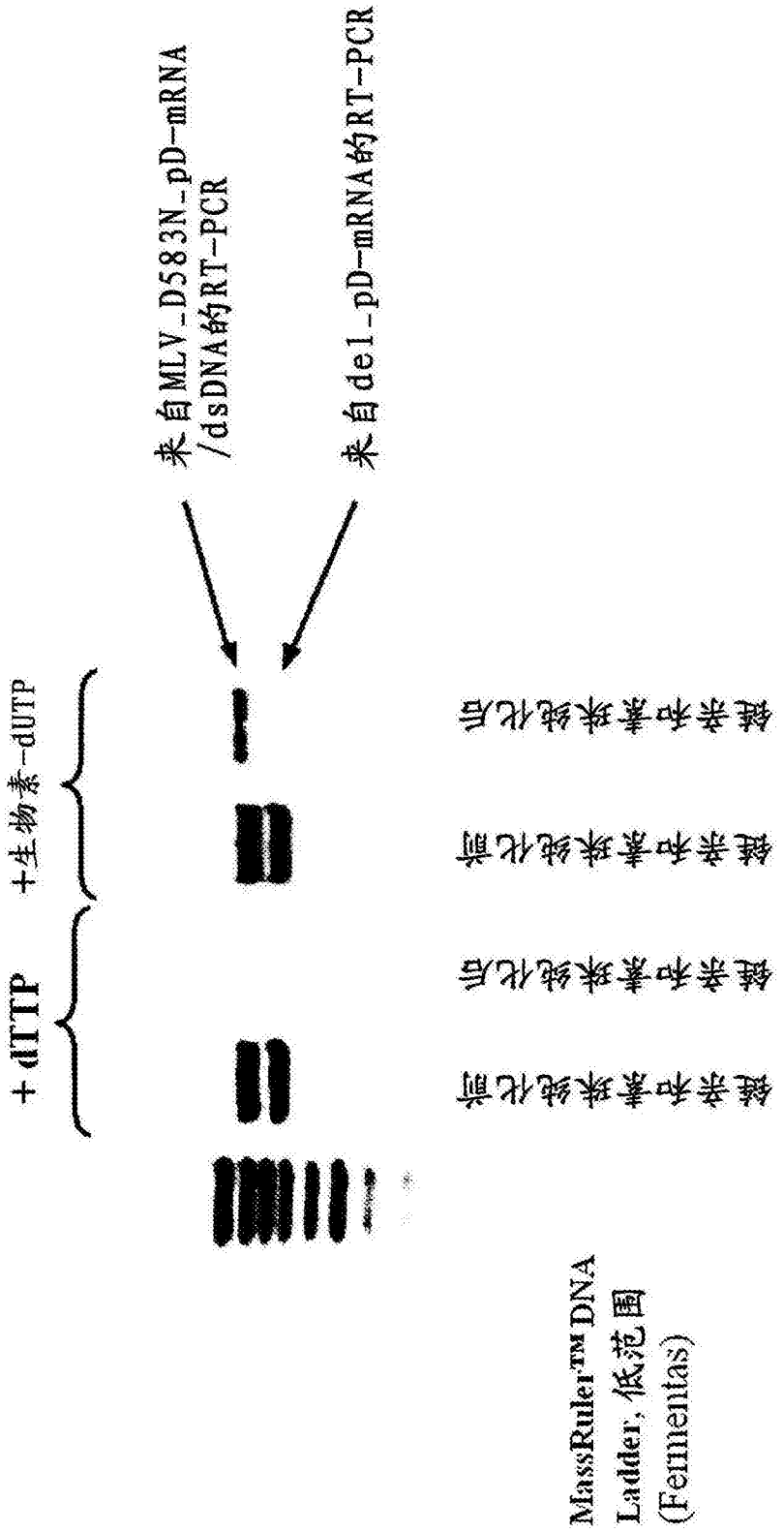

来自MLV_D583N_pD-mRNA /dsDNA的RT-PCR
来自del-pD-mRNA的RT-PCR
+生物素-dUTP
+ dTTP
链亲和素珠纯化后
链亲和素珠纯化前
链亲和素珠纯化后
链亲和素珠纯化前
MassRuler™ DNA Ladder, 低范围 (Fermentas)

图13


A
+dTTP
+生物素-dUTP
无连接酶
有连接酶
无连接酶
有连接酶
RiboRuler™
RNA Ladder,
高范围
(Fermentas)
B
+dTTP
+生物素-dUTP
无连接酶
有连接酶
无连接酶
有连接酶
C
long+
5′-.....mRNA.....CGAACGTGGCGAGAAAGGAAGGGAAGAAAGAAGTC +dTTP (生物素-dUTP) +[α-P33]dATP
ddC-GCCTTGGCCGCCACCTAGCACAGAAGAAGCTTGCACCGCTCTTTCCTTCCCTTCTTTCTTCAGATTTTTTT-5′
ddC-Long2

图14

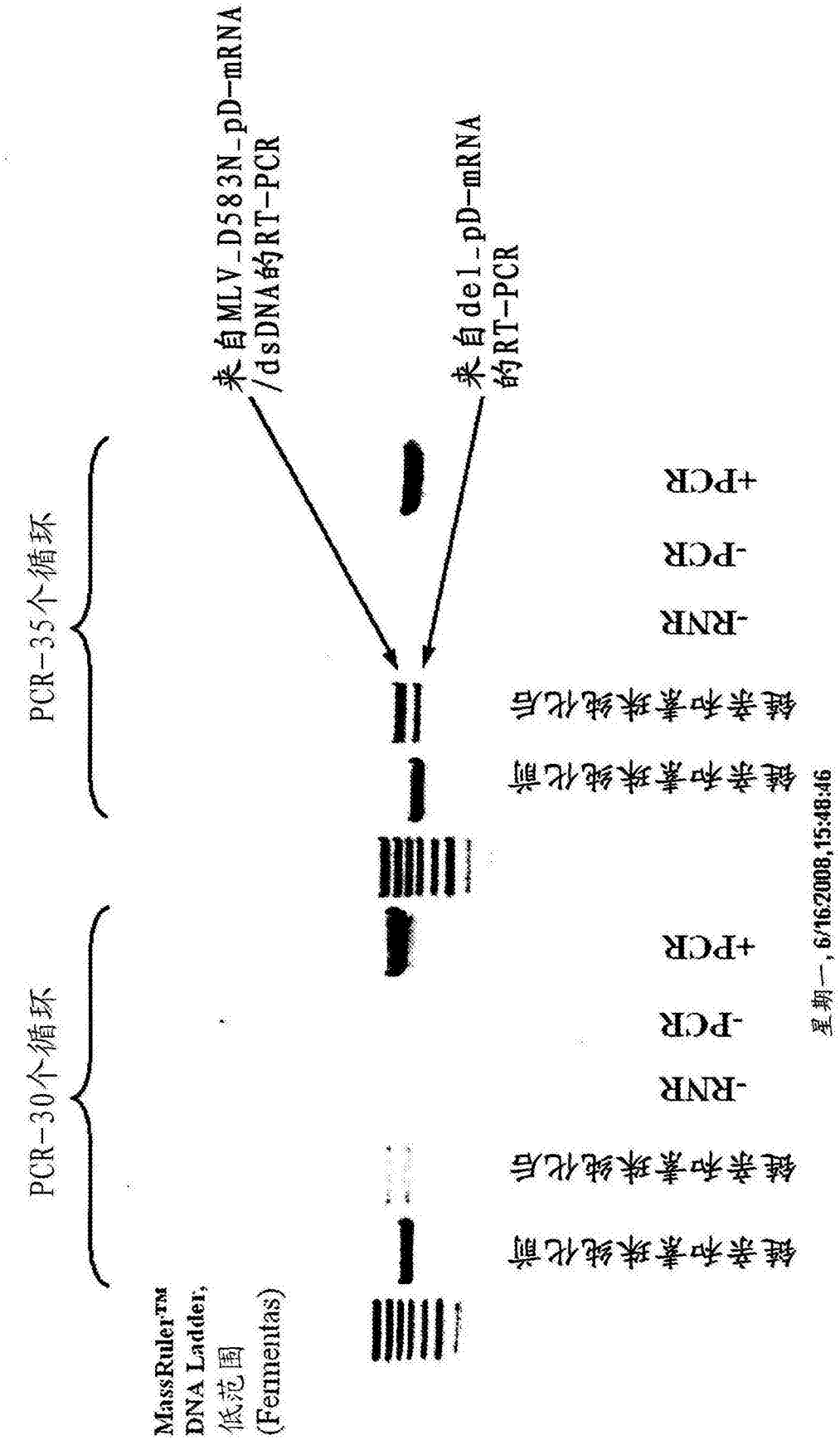

来自MLV_D583N_pD-mRNA /dsDNA的RT-PCR
来自del_pD-mRNA 的RT-PCR
PCR-35个循环
PCR-30个循环
MassRuler™ DNA Ladder, 低范围 (Fermentas)
链亲和素珠珠纯化前
链亲和素珠珠纯化后
-RNR
-PCR
+PCR
星期一, 6/16/2008, 15:48:46

图15

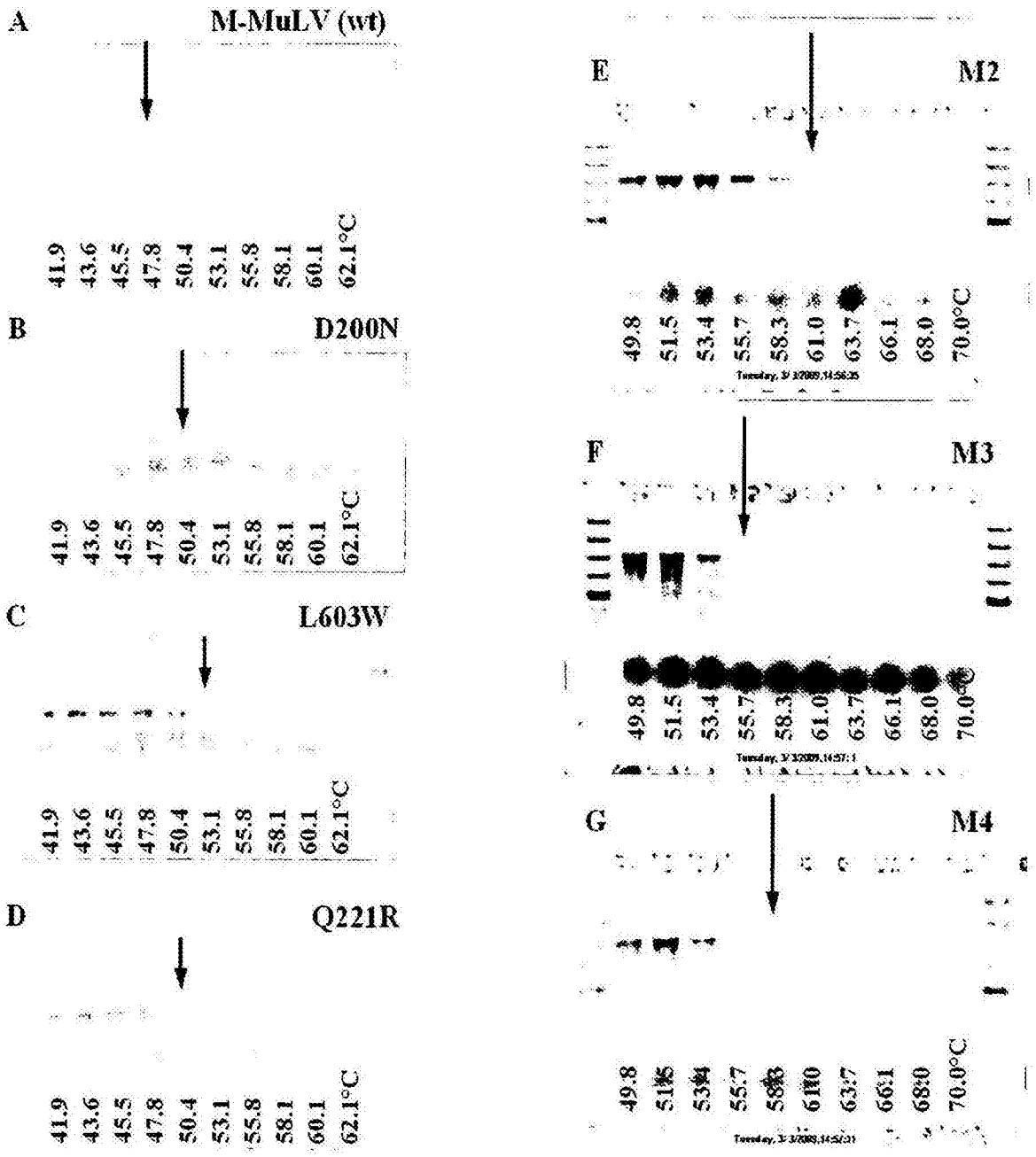

A
M-MuLV (wt)
41.9
43.6
45.5
47.8
50.4
53.1
55.8
58.1
60.1
62.1°C
B
D200N
41.9
43.6
45.5
47.8
50.4
53.1
55.8
58.1
60.1
62.1°C
C
L603W
41.9
43.6
45.5
47.8
50.4
53.1
55.8
58.1
60.1
62.1°C
D
Q221R
41.9
43.6
45.5
47.8
50.4
53.1
55.8
58.1
60.1
62.1°C
E
M2
49.8
51.5
53.4
55.7
58.3
61.0
63.7
66.1
68.0
70.0°C
F
M3
49.8
51.5
53.4
55.7
58.3
61.0
63.7
66.1
68.0
70.0°C
G
M4
49.8
51.5
53.4
55.7
58.3
61.0
63.7
66.1
68.0
70.0°C

图16