



(12) 发明专利

(10) 授权公告号 CN 107406882 B

(45) 授权公告日 2022.03.01

(21) 申请号 201680013462.4

(22) 申请日 2016.04.22

(65) 同一申请的已公布的文献号
申请公布号 CN 107406882 A

(43) 申请公布日 2017.11.28

(30) 优先权数据

62/152,697 2015.04.24 US

62/280,951 2016.01.20 US

(85) PCT国际申请进入国家阶段日
2017.09.01(86) PCT国际申请的申请数据
PCT/US2016/028971 2016.04.22(87) PCT国际申请的公布数据
W02016/172571 EN 2016.10.27(73) 专利权人 基纳生物技术有限公司
地址 美国加利福尼亚州

(72) 发明人 A·O·H·尼格伦

(74) 专利代理机构 上海专利商标事务所有限公司 31100

代理人 陈扬扬 钱文字

(51) Int.Cl.

C12Q 1/6858 (2018.01)

C12Q 1/6851 (2018.01)

(56) 对比文件

CN 101194023 A, 2008.06.04

CN 103333952 A, 2013.10.02

CN 104685064 A, 2015.06.03

CN 103717751 A, 2014.04.09

Mizuki Takenaka等. Multiplex single-base extension typing to identify nuclear genes required for RNA editing in plant organelles.《Nucleic Acids Research》.2008, 第37卷(第2期), 第e13页.

文思远. 单核苷酸多态性基因分型技术原理与进展.《生物技术通讯》.2003, 第218-221页.

审查员 王震

权利要求书4页 说明书53页 附图6页

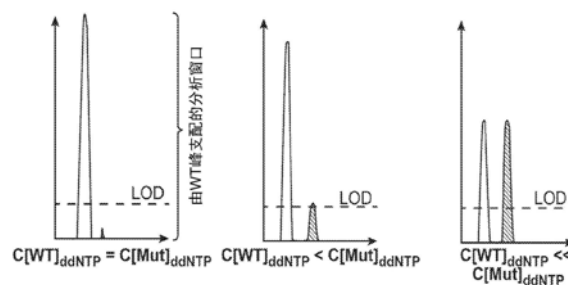
(54) 发明名称

用于次要等位基因和多态性的鉴定和定量的多重化方法

(57) 摘要

本文提供了用于检测含有次要核酸物质和一种或多种主要核酸物质的混合物的样品中次要核酸物质是否存在的产品和方法, 其中次要核酸物质的量(频率或拷贝数)小于主要核酸物质的量。某些方法包括扩增混合物并用链终止剂和与扩增子特异性杂交的延伸引物延伸所得的扩增子, 其中针对主要核酸物质特异性的链终止剂具有小于对于次要核酸物质特异性的链终止剂的浓度。使链终止剂的浓度倾向于相对于主要核酸物质特异性链终止剂有利于对于次要核酸物质特异性的链终止剂的高浓度改善了检测在混合物中以低频率或拷贝数存在的次要核酸物质的检测限(灵敏度)。另外, 从主要核酸物质扩增

子的延伸产物生成的信号可用作阳性对照并且允许相对于混合物中的主要核酸物质对次要核酸物质进行定量。



1. 一种用于非诊断目的的鉴定样品中一种或多种次要核酸物质的存在或不存在的多重化方法,所述样品包括所述一种或多种靶核酸,其中所述方法包括:

(a) 同时扩增所述样品的靶核酸,其中各靶核酸具有主要核酸物质和一种或多种次要核酸物质,其中各次要核酸物质是主要核酸物质的低频率或低拷贝数变体;

(b) 在包括对于一种或多种主要核酸物质特异的链终止剂和一种或多种对于一种或多种次要核酸物质特异的链终止剂的延伸条件下使(a)的扩增的核酸与一种或多种延伸引物接触,其中:

(i) 一种或多种主要核酸物质共有对于主要核酸物质特异且对次要核酸物质非特异的共同链终止剂,

(ii) 一种或多种次要核酸物质中的每一种具有对于次要核酸物质特异并且对于主要核酸物质非特异的链终止剂,其中对于次要核酸物质特异的链终止剂:(A)对于扩增的核酸中的特定次要核酸物质是独特的,并且不被扩增的核酸中的其他次要核酸物质所共有,或(B)一种或多种次要核酸物质中的至少一种与扩增的核酸中的至少一种其他次要核酸物质共有共同的链终止剂,和

(iii) 对于主要核酸物质特异的链终止剂的浓度小于对于一种或多种次要核酸物质特异的每种链终止剂的浓度,

由此引物延伸到相对于主要核酸物质的次要核酸物质中不同的核苷酸位置或通过其中,从而产生分别对应于所述样品中存在的次要核酸物质和主要核酸物质的链终止的延伸产物;和

(c) 分析(b)的延伸产物,从而鉴定一种或多种次要核酸物质的存在或不存在,

其中所述一种或多种次要核酸物质各自以主要核酸物质的频率或拷贝数的0.25%至小于10%,或0.5%至小于10%,或1%至小于10%的频率或拷贝数存在,或者所述一种或多种次要核酸物质代表少于主要核酸物质的30%,20%,15%,10%,8%,5%,4%,3%,2%,1%,0.8%,0.75%,0.5%,0.1%,0.05%或0.01%,并且

其中对于主要核酸物质特异的链终止剂的浓度为每个针对次要核酸物质特异的链终止剂的浓度的1%至20%,或0.1%至10%,或0.01%至5%。

2. 如权利要求1所述的方法,其中核酸样品包括多种次要核酸物质,其是单一主要核酸物质的变体,并且在单一多重化反应中鉴定所述多种次要核酸物质。

3. 如权利要求1所述的方法,其中核酸样品包括多种主要核酸物质,其各自具有一种或多种次要核酸物质,并且所述次要核酸物质在单一多重化反应中鉴定。

4. 如权利要求1-3中任一项所述的方法,还包括在(c)中鉴定一种或多种主要核酸物质的存在或不存在。

5. 一种用于非诊断目的的定量核酸样品中一种或多种次要核酸物质的方法,所述核酸样品包括靶核酸,所述方法包括:

(a) 同时扩增所述样品的靶核酸,其中各靶核酸具有主要核酸物质和一种或多种次要核酸物质,其中一种或多种次要核酸物质是相同主要核酸物质的变体并且各次要核酸物质是主要核酸物质的低频率或低拷贝数变体;

(b) 在延伸条件下使(a)的扩增的核酸与延伸引物接触,所述延伸条件包括对于一种或多种次要核酸物质中的每一种特异的链终止剂,和对于主要核酸物质特异的链终止剂,由

此引物延伸到相对于主要核酸物质的次要核酸物质中不同的核苷酸位置或通过其中,从而产生分别对应于所述样品中存在的次要核酸物质和主要核酸物质的链终止的延伸产物,其中:(1)每一种链终止剂的浓度是已知的;并且(2)对于主要核酸物质特异的链终止剂的浓度小于对于一种或多种次要核酸物质特异的链终止剂的浓度;

(c) 确定对应于一种或多种次要核酸物质中的每一种的延伸产物的量相对于对应于主要核酸物质的延伸产物的量的比例;和

(d) 基于(c)的比例,并且基于相对于对于主要核酸物质特异的链终止剂的浓度的对于一种或多种次要核酸物质特异的链终止剂的浓度,对相对于主要核酸物质的量对次要核酸物质的量进行定量,

其中所述一种或多种次要核酸物质各自以主要核酸物质的频率或拷贝数的0.25%至小于10%,或0.5%至小于10%,或1%至小于10%的频率或拷贝数存在,或者所述一种或多种次要核酸物质代表少于主要核酸物质的30%,20%,15%,10%,8%,5%,4%,3%,2%,1%,0.8%,0.75%,0.5%,0.1%,0.05%或0.01%,并且

其中对于主要核酸物质特异的链终止剂的浓度为每个针对次要核酸物质特异的链终止剂的浓度的1%至20%,或0.1%至10%,或0.01%至5%。

6. 一种用于非诊断目的的定量核酸样品中一种或多种次要核酸物质的多重化方法,所述核酸样品包括所述一种或多种靶核酸,所述方法包括:

(a) 同时扩增所述样品的靶核酸,其中各靶核酸具有主要核酸物质和一种或多种次要核酸物质,其中各次要核酸物质是主要核酸物质的低频率或低拷贝数变体;

(b) 在包括对于一种或多种主要核酸物质特异的链终止剂和一种或多种对于一种或多种次要核酸物质特异的链终止剂的延伸条件下使(a)的扩增的核酸与延伸引物接触,其中:

(i) 一种或多种主要核酸物质共有对于主要核酸物质特异且对次要核酸物质非特异的共同链终止剂,

(ii) 一种或多种次要核酸物质中的每一种具有对于次要核酸物质特异并且对于主要核酸物质非特异的链终止剂,其中对于次要核酸物质特异的链终止剂来说:(A) 其对于扩增的核酸中的特定次要核酸物质是独特的,并且不被扩增的核酸中的其他次要核酸物质所共有,或(B) 一种或多种次要核酸物质中的至少一种与扩增的核酸中的至少一种其他次要核酸物质共有共同的链终止剂,和

(iii) 每一种链终止剂的浓度是已知的并且对于主要核酸物质特异的链终止剂的浓度小于对于一种或多种次要核酸物质特异的各链终止剂的浓度,

由此引物延伸到次要核酸物质中相对于主要核酸物质不同的核苷酸位置或通过该位置,从而产生分别对应于样品中存在的次要核酸物质和主要核酸物质的链终止的延伸产物;

(c) 确定对应于一种或多种次要核酸物质中的每一种的延伸产物的量相对于对应于主要核酸物质的延伸产物的量的比例;和

(d) 基于(c)的比例,并且基于相对于对于主要核酸物质特异的链终止剂的浓度的对于一种或多种次要核酸物质特异的链终止剂的浓度,对相对于主要核酸物质的量的次要核酸物质的量进行定量,

其中所述一种或多种次要核酸物质各自以主要核酸物质的频率或拷贝数的0.25%至

小于10%，或0.5%至小于10%，或1%至小于10%的频率或拷贝数存在，或者所述一种或多种次要核酸物质代表少于主要核酸物质的30%，20%，15%，10%，或8%，5%，4%，3%，2%，1%，0.8%，0.75%，0.5%，0.1%，0.05%，或0.01%，并且

其中对于主要核酸物质特异的链终止剂的浓度为每个针对次要核酸物质特异的链终止剂的浓度的1%至20%，或0.1%至10%，或0.01%至5%。

7.如权利要求5或6所述的方法，其中核酸样品包括多种次要核酸物质，其是主要核酸物质的变体，并且在单一多重化反应中鉴定所述多种次要核酸物质。

8.如权利要求1-3、5和6中任一项所述的方法，其中所述次要和主要核酸物质的序列相差单个碱基，并且引物延伸到不同的单一碱基。

9.如权利要求1-3、5和6中任一项所述的方法，其中所述次要核酸物质的序列包括相对于所述主要核酸物质的序列的插入或缺失。

10.如权利要求1-3、5和6中任一项所述的方法，其中所述一种或多种次要核酸物质是主要核酸物质的单核苷酸多态性SNP变体。

11.如权利要求1-3、5和6中任一项所述的方法，其中所述次要和主要核酸物质分别是相同基因的突变体和野生型等位基因，或其中所述一种或多种次要核酸物质是野生型主要核酸物质的体细胞突变。

12.如权利要求1-3、5和6中任一项所述的方法，其中所述主要核酸物质来自宿主对象，并且所述次要核酸物质来自除所述宿主以外的对象，或者所述次要核酸物质来自宿主对象并且所述主要核酸物质来自除所述宿主以外的对象。

13.如权利要求1-3、5和6中任一项所述的方法，其中所述一种或多种次要核酸物质各自以小于主要核酸物质的频率或拷贝数的10%的频率或拷贝数存在。

14.如权利要求13所述的方法，其中所述一种或多种次要核酸物质各自以主要核酸物质的频率或拷贝数的1%至5%，或1%至4%，或1%至3.5%，或1%至3%的频率或拷贝数存在。

15.如权利要求13所述的方法，其中所述一种或多种次要核酸物质各自以主要核酸物质的频率或拷贝数的2%至小于10%的频率或拷贝数存在。

16.如权利要求13所述的方法，其中所述一种或多种次要核酸物质各自以主要核酸物质的频率或拷贝数的小于5%或小于4%的频率或拷贝数存在。

17.如权利要求16所述的方法，其中所述一种或多种次要核酸物质各自以主要核酸物质的频率或拷贝数的小于2%的频率或拷贝数存在。

18.如权利要求1-3、5和6中任一项所述的方法，其中对于主要核酸物质特异的链终止剂的浓度为每个针对次要核酸物质特异的链终止剂的浓度的小于5%。

19.如权利要求1-3、5和6中任一项所述的方法，其中对于主要核酸物质特异的链终止剂的浓度为每个针对次要核酸物质特异的链终止剂的浓度的0.5%至小于15%。

20.如权利要求1-3、5和6中任一项所述的方法，其中对于主要核酸物质特异的链终止剂的浓度为每个针对次要核酸物质特异的链终止剂的浓度的1%至10%，或1%至7%。

21.如权利要求20所述的方法，其中对于主要核酸物质特异的链终止剂的浓度为每个针对次要核酸物质特异的链终止剂的浓度的1%至2%。

22.如权利要求1-3、5和6中任一项所述的方法，其中所述对于主要核酸物质特异的链

终止剂的浓度与所述对于次要核酸物质特异的链终止剂的浓度的比例基于特定的对于主要核酸物质特异的链终止剂和特定的对于次要核酸物质特异的链终止剂。

23. 如权利要求1-3、5和6中任一项所述的方法,其中所述链终止剂是链终止核苷酸。

24. 如权利要求23所述的方法,其中所述链终止核苷酸独立地选自ddATP、ddGTP、ddCTP、ddTTP和ddUTP。

25. 如权利要求23所述的方法,其中针对一种或多种次要核酸物质特异的链终止核苷酸由一个链终止核苷酸组成。

26. 如权利要求23所述的方法,其中针对一种或多种次要核酸物质特异的链终止核苷酸由2个链终止核苷酸组成。

27. 如权利要求23所述的方法,其中针对一种或多种次要核酸物质特异的链终止核苷酸由3个链终止核苷酸组成。

28. 如权利要求1-3、5和6中任一项所述的方法,其中所述链终止剂包括一种或多种无环终止剂。

29. 如权利要求1-3、5和6中任一项所述的方法,包含(a)中30至45个PCR扩增循环。

30. 如权利要求1-3、5和6中任一项所述的方法,其中在(b)中的所述延伸条件包括多个延伸循环,其中延伸反应循环200至300次。

31. 如权利要求1-3、5和6中任一项所述的方法,其中(b)中的延伸条件包括多个延伸循环,其中延伸反应循环至少65、70、75、80、85、90、95或100次。

32. 如权利要求1-3、5和6中任一项所述的方法,其中所述一种或多种次要核酸物质各自以主要核酸物质的频率或拷贝数的1%的频率或拷贝数存在。

33. 如权利要求1-3、5和6中任一项所述的方法,其中所述链终止剂不包括可检测的标记。

34. 如权利要求33所述的方法,其中所述延伸产物通过质谱进行分析。

35. 如权利要求5或6所述的方法,其中对应于所述一种或多种次要核酸物质中每一种的延伸产物的量相对于对应于所述主要核酸物质的延伸产物的量的比例确定为通过质谱分析生成的次要核酸物质的信号与通过质谱分析生成的主要核酸物质的信号的比例。

36. 如权利要求1-3、5和6中任一项所述的方法,其中所述链终止剂中的一种或多种包括可检测的标记。

37. 如权利要求36所述的方法,还包括检测所述标记,从而鉴定或定量所述一种或多种次要核酸物质。

38. 如权利要求36所述的方法,其中所述标记是荧光标记或染料或是质量标记。

39. 如权利要求38所述的方法,其中所述标记是质量标记并通过质谱进行检测,或所述标记是荧光标记或染料并且通过电泳或实时PCR进行检测。

用于次要等位基因和多态性的鉴定和定量的多重化方法

[0001] 相关申请

[0002] 本申请要求2016年1月20日提交的题为“用于次要等位基因和多态性的鉴定和定量的多重化方法”，发明人为Anders Nygren并且代理人案卷号为AGB-7001-PV2的美国临时申请号62/280951的优先权。本申请还要求2015年4月24日提交的题为“用于次要等位基因和多态性的鉴定和定量的多重化方法”，发明人为Anders Nygren并且代理人案卷号为AGB-7001-PV的美国临时申请号62/152697的优先权。本申请涉及2012年12月18日提交的发明人为Martin Beaulieu和Dirk Johannes van den Boom，标题为“用于高水平多重化聚合酶链反应和均匀质量延伸反应的方法”，代理人案卷号为AGB-2079-CT2的美国专利申请号13/718,758，其是2011年7月28日提交的发明人为Martin Beaulieu和Dirk Johannes van den Boom，标题为“用于高水平多重化聚合酶链反应和均匀质量延伸反应的方法”，代理人案卷号为AGB-2079-CT的美国专利申请号13/193,390，现在美国专利号8,349,566的连续申请，其是2004年7月30日提交的发明人为Martin Beaulieu和Dirk Johannes van den Boom，标题为“用于高水平多重化聚合酶链反应和均匀质量延伸反应的方法”，代理人案卷号为SEQ-2079-UT的美国专利申请号10/903,268，现在美国专利号8,003,317的连续申请，其根据35U.S.C. §119(e) 要求2003年7月31日提交的Martin Beaulieu和Dirk van den Boom的标题为“用于高水平多重化聚合酶链反应和均匀质量延伸反应的方法”，代理人案卷号为17082-087P01 (P2079) 的美国临时申请号60/492,102的优选权。本申请还涉及2012年12月18日提交的发明人为Martin Beaulieu和Dirk van den Boom，标题为“用于高水平多重化聚合酶链反应和均匀质量延伸反应的方法”的国际PCT申请号PCT/US2004/024953。这些申请各自的内容通过引用纳入本文，包括文本、表格和附图。

技术领域

[0003] 该技术部分涉及鉴定和/或定量核酸变体，例如，野生型等位基因的突变变体或多态性。

背景技术

[0004] 检测特异性核酸是诊断医学和分子生物学研究的重要工具。核酸试验可以，例如，鉴定宿主对象中的感染性生物体如细菌和病毒，探测正常基因的表达和鉴定突变基因如癌基因，组织移植前针对相容性对组织分型，匹配法医学的组织或血液样品，分析不同物质的基因之间的同源性和鉴定基因变体的等位基因和多态性。这些应用通常需要在含有较大量的非靶（主要(major)）核酸物质的核酸样品或混合物中检测和/或鉴定较小量的存在的感兴趣的次要(minor) 核酸物质（例如，野生型等位基因的次要等位基因变体，或癌基因）的能力。如果可以多重化方式进行核酸试验，即筛选多种核酸，则可进一步增强这种能力（例如，在效率上）。

[0005] 通常可靠且可重复的检测等位基因间的低频率（拷贝数）变体、多态性或其他突变的现有方法，包括多重化方法，具有小于鉴定低频率变体可能需要的检测灵敏性或检测限

(例如,约10%-15%频率的检测限将无法鉴定低于该范围的突变频率)。较低检测限一般是由于低频率变体检测信号被主要野生型物质的较大检测信号遮蔽。其他方法通过去除“野生型”信号,从而改善低频率变体的检测来克服这一问题。然而,去除野生型信号可能导致难以对核酸样品中变体的相对量进行定量,或者难以明确确认不存在低频率变体。方法,包括多重化方法,其将高检测灵敏度和高准确性结合,可提供对通过之前的方法无法检测到或无法最佳地检测到的某些低频率变体改善的鉴定和/或定量。

发明内容

[0006] 本文提供了方法,包括多重化方法,其将高检测灵敏度和高准确性结合,提供对无法检测到或无法最佳地检测到或定量的某些低频率变体改善的鉴定和/或定量。

[0007] 在某些实施方式中,本文提供了一种用于鉴定核酸群中一种或多种次要核酸物质的存在或不存在的多重化方法,所述核酸群包括所述一种或多种次要核酸物质和一种或多种主要核酸物质的混合物,其中每种次要核酸物质是对应的主要核酸物质的变体,并且其拷贝数小于其对应的主要核酸物质的拷贝数,其中该方法包括:

[0008] (a) 在包括dNTP的扩增条件下用扩增引物对同时扩增混合物的靶区域,由此产生包含主要和次要核酸物质的核酸的扩增混合物;

[0009] (b) 在包括链终止剂的延伸条件下使扩增混合物与延伸引物接触,其中:

[0010] (i) 一种或多种主要核酸物质共有对于主要核酸物质特异且对次要核酸物质非特异的共同链终止剂,和

[0011] (ii) 一种或多种次要核酸物质中的每一种具有对于次要核酸物质特异并且对于主要核酸物质非特异的链终止剂,其中对于次要核酸物质特异的链终止剂: (A) 对于扩增混合物中的特定次要核酸物质是独特的,并且不被扩增混合物中的其他次要核酸物质所共有,或 (B) 一种或多种次要核酸物质中的至少一种与扩增混合物中的至少一种其他次要核酸物质共有共同的链终止剂,

[0012] 由此引物延伸到相对于主要核酸物质的次要核酸物质中不同的核苷酸位置或通过其中,从而产生分别对应于次要核酸物质和主要核酸物质的链终止的延伸产物,其中对于主要核酸物质特异的链终止剂的浓度小于对于一种或多种次要核酸物质特异的每种链终止剂的浓度;和

[0013] (c) 分析 (b) 的延伸产物,从而鉴定一种或多种次要核酸物质的存在或不存在。

[0014] 在某些实施方式中,核酸群包括多种次要核酸物质,其是单一主要核酸物质的变体,并且在单一多重化反应中鉴定所述多种次要核酸物质。在一些实施方式中,该方法的部分 (b) 在一组至少两个反应容器或隔室中进行,其中:

[0015] 第一容器或隔室包括含有针对主要核酸物质特异的链终止剂并且不含针对一种或多种次要核酸物质特异的链终止剂的延伸条件;和

[0016] 每个剩余的容器或隔室包括延伸条件,其包含对一种或多种次要核酸物质特异且共有的单一链终止剂,并且不含对于主要核酸物质特异或针对次要核酸物质特异且不共有的单一链终止剂的链终止剂。在一些实施方式中,每种链终止剂的浓度是已知的。

[0017] 本文还提供了定量核酸群中一种或多种次要核酸物质的方法,所述核酸群包括所述一种或多种次要核酸物质和主要核酸物质的混合物,其中所述次要核酸物质是相同的主

要核酸物质的变体并且各自以比主要核酸物质的拷贝数低的拷贝数存在,其中该方法包括:

[0018] (a) 在包括dNTP的扩增条件下用扩增引物对同时扩增混合物的靶区域,由此产生包含主要和次要核酸物质的核酸的扩增混合物;

[0019] (b) 在延伸条件下使扩增混合物与延伸引物接触,所述延伸条件包括对于(i)一种或多种次要核酸物质中的每一种,和(ii)主要核酸物质特异性链终止剂,由此引物延伸到相对于主要核酸物质的次要核酸物质中不同的核苷酸位置或通过其中,从而产生分别对应于次要核酸物质和主要核酸物质的链终止的延伸产物,其中:(1)每一种链终止剂的浓度是已知的;并且(2)对于主要核酸物质特异的链终止剂的浓度小于对于一种或多种次要核酸物质特异的链终止剂的浓度;

[0020] (c) 确定对应于一种或多种次要核酸物质中的每一种的延伸产物的量相对于对应于主要核酸物质的延伸产物的量的比例;和

[0021] (d) 基于(c)的比例,并且基于相对于主要核酸物质特异的链终止剂的浓度的对于一种或多种次要核酸物质特异的链终止剂的浓度,对相对于主要核酸物质的量的次要核酸物质的量进行定量。

[0022] 本文还提供了定量核酸群中一种或多种次要核酸物质的多重化方法,所述核酸群包括所述一种或多种次要核酸物质和主要核酸物质的混合物,其中每一种次要核酸物质是对应的主要核酸物质的变体并且以比其对应的主要核酸物质的拷贝数低的拷贝数存在,其中该方法包括:

[0023] (a) 在包括dNTP的扩增条件下用扩增引物对同时扩增混合物的靶区域,由此产生包含主要和次要核酸物质的核酸的扩增混合物;

[0024] (b) 在包括链终止剂的延伸条件下使扩增混合物与延伸引物接触,其中:

[0025] (i) 一种或多种主要核酸物质共有对于主要核酸物质特异且对次要核酸物质非特异的共同链终止剂,和

[0026] (ii) 一种或多种次要核酸物质中的每一种具有对于次要核酸物质特异并且对于主要核酸物质非特异的链终止剂,其中对于次要核酸物质特异的链终止剂:(A)对于扩增混合物中的特定次要核酸物质是独特的,并且不被扩增混合物中的其他次要核酸物质所共有,或(B)一种或多种次要核酸物质中的至少一种与扩增混合物中的至少一种其他次要核酸物质共有共同的链终止剂,

[0027] 由此引物延伸到相对于主要核酸物质的次要核酸物质中不同的核苷酸位置或通过其中,从而产生分别对应于次要核酸物质和主要核酸物质的链终止的延伸产物,其中:(1)每一种链终止剂的浓度是已知的;和(2)对于主要核酸物质特异的链终止剂的浓度小于对于一种或多种次要核酸物质特异的链终止剂的浓度;

[0028] (c) 确定对应于一种或多种次要核酸物质中的每一种的延伸产物的量相对于对应于主要核酸物质的延伸产物的量的比例;和

[0029] (d) 基于(c)的比例,并且基于相对于主要核酸物质特异的链终止剂的浓度的对于一种或多种次要核酸物质特异的链终止剂的浓度,对相对于主要核酸物质的量的次要核酸物质的量进行定量。

[0030] 在该方法的些实施方式中,核酸群包括多种次要核酸物质,其是单一主要核酸物

质的变体,并且在单一多重化反应中鉴定该多种次要核酸物质。在某些实施方式中,该方法的部分(b)在一组至少两个反应容器或隔室中进行,其中:

[0031] 第一容器或隔室包括含有针对主要核酸物质特异的链终止剂并且不含针对一种或多种次要核酸物质特异的链终止剂的延伸条件;和

[0032] 每个剩余的容器或隔室包括延伸条件,其包含对一种或多种次要核酸物质特异且共有的单一链终止剂,并且不含对于主要核酸物质特异或针对次要核酸物质特异且不共有共同的单一链终止剂的链终止剂。

[0033] 在本文提供的方法的某些实施方式中,次要和主要核酸物质的序列相差单个碱基,并且引物延伸到或通过不同的单一碱基。在一些实施方式中,次要核酸物质的序列包括相对于主要核酸物质的序列的插入或缺失。在一些实施方式中,一种或多种次要核酸物质是主要核酸物质的单核苷酸多态性(SNP)变体。在某些实施方式中,次要和主要核酸物质分别是相同基因的突变体和野生型等位基因。

[0034] 在本文提供的方法的一些实施方式中,主要核酸物质来自宿主对象,并且次要核酸物质来自除宿主以外的对象。在某些实施方式中,一种或多种次要核酸物质各自以小于主要核酸物质的拷贝数的约10%的拷贝数存在。在某些实施方式中,一种或多种次要核酸物质各自以主要核酸物质的拷贝数的约1%至小于10%的拷贝数存在。

[0035] 在一些实施方式中,一种或多种次要核酸物质各自以主要核酸物质的拷贝数的约2%至小于10%的拷贝数存在。

[0036] 在本文提供的方法的某些实施方式中,对于主要核酸物质特异的链终止剂的浓度为每个针对次要核酸物质特异的链终止剂的浓度的约1%至约20%。在本文提供的方法的某些实施方式中,对于主要核酸物质特异的链终止剂的浓度为每个针对次要核酸物质特异的链终止剂的浓度的约0.1%至约10%。在本文提供的方法的某些实施方式中,对于主要核酸物质特异的链终止剂的浓度为每个针对次要核酸物质特异的链终止剂的浓度的约0.01%至约5%。在一些实施方式中,链终止剂是链终止核苷酸。在实施方式中,链终止核苷酸独立地选自ddATP、ddGTP、ddCTP、ddTTP和ddUTP。

[0037] 在本文提供的方法的一些实施方式中,针对一种或多种次要核酸物质特异的链终止核苷酸由一个链终止核苷酸组成。在某些实施方式中,针对一种或多种次要核酸物质特异的链终止核苷酸由两个链终止核苷酸组成。在一些实施方式中,针对一种或多种次要核酸物质特异的链终止核苷酸由三个链终止核苷酸组成。

[0038] 在本文提供的方法的某些实施方式中,链终止剂包括一种或多种无环终止剂。在本文提供的方法的实施方式中,部分(a)中的扩增包括使用扩增引物的约30至约45个PCR扩增循环。在一些实施方式中,(b)中的延伸条件包括延伸延伸引物的约20至约300个循环。在某些实施方式中,(b)中的延伸条件包括至少50个延伸循环。

[0039] 在本文提供的方法的实施方式中,一种或多种链终止剂和/或延伸引物包括可检测标记。在某些实施方式中,标记是荧光标记或染料。在一些实施方式中,标记是质量标记。在实施方式中,可以检测标记,由此鉴定或定量一种或多种次要核酸物质。在某些实施方式中,标记是质量标记,并通过质谱进行检测。

[0040] 本文提供的方法在某些实施方式中包括在延伸之前(单个碱基或几个碱基,使用链终止剂)进行对次要核酸物质和/或主要核酸物质的扩增。在一些实施方式中,(a)中的扩

增反应条件包括水、基因组DNA、缓冲液、dNTP、引物对、 $MgCl_2$ 和聚合酶,其中 $MgCl_2$ 的浓度与dNTP中每一种的浓度之比选自 $\leq 10:1$ 、 $\leq 9:1$ 、 $\leq 8:1$ 、 $\leq 7:1$ 、 $\leq 6:1$ 或 $\leq 5:1$ 。在某些实施方式中,聚合酶是至少约0.03单位/ μl 浓度的Taq聚合酶。在本文提供的方法的实施方式中,(a)中的扩增反应条件包含约400-700 μM 的各dNTP,约100nM引物对和约1.6至约4.8mM $MgCl_2$ 。在一些实施方式中,序列标签附着于扩增引物对中的一种或多种引物。在一些实施方式中,游离 Mg^{2+} 浓度在1.0-2.0mM之间。

[0041] 在一些实施方式中,已经鉴定了在本文方法中进行的扩增的特定 $MgCl_2$ 浓度,其允许高水平(例如,7-重至50-重或更多重)的多重化PCR和引物延伸反应以及成功分析扩增产物,例如,通过检测荧光信号或通过质谱分析。当选择用于PCR扩增反应中(随后进行质量延伸反应和后续质谱分析)的dNTP和 $MgCl_2$ 的浓度时的考虑是PCR反应混合物的游离 Mg^{2+} 浓度应保持在特定的范围内,其范围足够高以允许粗放PCR扩增,同时足够低以不会不利地影响随后的质量延伸反应和质谱分析。

[0042] 因此,本文提供了通过在扩增条件下同时扩增次要物质的多个核酸靶区域和主要核酸物质的相应区域来对多个多态性位点或次要核酸物质的其它变体进行基因分型的多重化方法,从而由7个或更多个引物对来扩增7个或更多个核酸靶区域的至少60%,以产生含有多态性位点或其他变体的核酸-靶区域的扩增混合物,使扩增的核酸靶区域的混合物在具有针对主要核酸物质特异的至少一种链终止剂和针对次要核酸物质特异的链终止剂的存在下,的引物延伸条件下与7个或更多个基因分型引物(即,延伸引物或UEP)接触,从而引物延伸到或通过相应的多态性位点,其中,对核酸-靶分子内的每个多态性位点存在一种基因分型引物,并且通过质量或其它标记如荧光或电化学标记检测延伸的基因分型引物,其中确定所述7个或更多尝试核酸靶区域的至少60%的基因型。在某些实施方式中,扩增引物对的数量可以选自8个或更多个,9个或更多个,10个或更多个,11个或更多个,12个或更多个,或13个或更多个。

[0043] 在上述每种方法的某些实施方式中,序列标签连接到每个引物对的一个或两个扩增引物的5'末端。在其他实施方式中,使用允许确定(即,判定)至少65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%、高达100%的尝试次要核酸物质的条件(如本文提供的扩增反应条件和/或引物质量延伸反应条件)进行对多个序列变体的多重化检测的方法。本文提供的条件适用于使用多种扩增引物对和来自多种靶核酸的7-重或更多扩增反应的多次多重化反应。此外,如本文所述,所有优化的扩增和/或引物质量延伸基因分型反应都适用于范围从2-重至6-重及其以上的多重化试验。

[0044] 在特定实施方式中,序列标签连接到选自13个或更多个,14个以上,15个以上,16个以上,17个以上,18个以上,19个以上,或20个以上,21个,22个,23个,24个,25个,26个,27个,28个,29个,30个,35个,40个,45个,或50个以上的一级和二级引物对的多个一级和二级扩增引物对。序列标签可以连接到来自每对的一级引物和二级引物中的一个或两个。通常,序列标签连接到每对的一级和二级引物。本文使用的序列标签的长度可以为5至20,5至30,5至40,或5至50个核苷酸,10聚体长度的序列标签在本文提供的方法中特别有用。在多重化扩增反应中,序列标签不必对于每个引物对是相同序列,也不必对于特定扩增对中的一级和二级引物是相同序列。在特定实施方式中,序列标签对于多重化扩增反应中的每个引物是相同的。例如,在某些实施方式中,序列标签是10聚体,并且连接到每个一级和二级引物

的5'末端。在本文提供的方法的具体实施方式中,仅使用单个引物对来扩增每个特定核酸靶(次要核酸物质)区域。

[0045] 在本文提供的方法的实施方式中,上述方法的扩增反应条件包括水,次要和主要核酸物质的混合物,缓冲液,dNTP,一级和二级引物对,MgCl₂,和聚合酶,其中MgCl₂浓度与dNTP中每一种浓度之比选自 $\leq 10:1$ 、 $\leq 9:1$ 、 $\leq 8:1$ 、 $\leq 7:1$ 、 $\leq 6:1$ 、或 $\leq 5:1$ 。在一个具体的实施方式中,MgCl₂的浓度与dNTP中每一种浓度之比为 $\leq 7:1$ 。在其它实施方式中,扩增反应条件包含约400-700 μ M,约500-600 μ M或约500 μ M的每种dNTP,以及约50-250nM引物对。在这些实施方式中,总MgCl₂浓度可以在约2.6mM至约4.8mM MgCl₂之间,约3.0至约4.5mM MgCl₂之间,以及约3.5mM至约4.0mM MgCl₂之间。在实施方式中,当选择dNTP和MgCl₂的浓度时,游离Mg²⁺浓度在约1-2mM之间。如本文所用,游离Mg²⁺浓度=总Mg²⁺浓度(例如,总[MgCl₂]) - 所有4种dNTP的总dNTP浓度(例如,200 μ M各dNTP=800 μ M总[dNTP])。在某些实施方式中,游离Mg²⁺在1.1-1.9mM之间,在1.2-1.8mM之间,在1.3-1.7mM之间,在1.4-1.6mM之间。在一个具体实施方式中,游离Mg²⁺浓度为约1.5mM。对于这些方法中的每一种,扩增热循环的多样性可以是约30至约35、40或45个循环。在一个具体实施方式中,扩增反应条件包含约500 μ M的每种dNTP,约100nM引物对和约3.5mM MgCl₂。对于这些方法中的每一种,聚合酶可以是浓度为0.03单位/ μ l的Taq聚合酶(例如,购自 QIAGEN[®]的 HOTSTARTAQ[®])。在本文提供的方法的具体实施方式中,扩增反应条件不包括加入选自以下添加剂中的一种或任何组合:BSA(牛血清白蛋白)、甘油、DMSO(二甲基亚砷)、尿素或 Q-SOLUTION[®]。

[0046] 延伸引物(UEP)与序列变异相邻杂交,并且混合物还包含dNTP和ddNTP或其它链终止剂例如包括无环终止剂的试剂的预选组合。在dNTP和ddNTP的预选组合中,当ddNTP存在于混合物中时,不存在相同的dNTP。对于这些方法,引物质量延伸热循环的多重性选自至少约50、55、60、65、70、75、80、85、90、95、100、125、150、175、200或更多个循环。在该方法的实施方式中,引物质量延伸反应条件包括对于每个次要核酸物质(C[Mut])特异的约20 μ M至约300 μ M ddNTP或其它链终止剂,通常约20 μ M至约200 μ M ddNTP,约40 μ M至约200 μ M ddNTP,约100 μ M至约200 μ M ddNTP或约50 μ M ddNTP,对于主要核酸物质C[WT]特异的链终止剂的浓度小于对于次要核酸物质(C[Mut])特异的链终止剂浓度的20%,C[WT]通常在C[Mut]的约0.5%至小于约20%之间,C[Mut]的约0.5%至小于约15%,C[Mut]的约1%至约15%,C[Mut]的约1%至约10%,C[Mut]的约2%至约10%或C[Mut]的约1%至约2%和约1 μ M的UEP引物。在某些实施方式中,C[WT]为C[Mut]的约0.1%至约10%或C[Mut]的约0.01%至约10%。在这些引物延伸反应的某些实施方式中,延伸反应条件还包含约0.05至约0.5单位DNA聚合酶/微升。在其他实施方式中,引物质量延伸反应条件还包含约0.1至约0.3单位DNA聚合酶/微升。在其他实施方式中,引物质量延伸反应条件还包含约0.14至约0.2单位DNA聚合酶/微升。在一个实施方式中,引物质量延伸反应条件还包含约0.14单位DNA聚合酶/微升。

[0047] 在某些实施方式中,引物延伸反应具有主要物质链终止剂的存在量为次要物质链终止剂的约0.1%至约10%。在其他实施方式中,针对每个转化(即A至C,A至G,A至T,C至A,C至G,C至T,G至A,G至C,G至T,T至A,T至C和T至G)单独调整主要物质链终止剂百分比,以考虑链终止剂的转化特异掺入的差异。例如,用于优化主要物质链终止剂相对于次要物质链终止

剂的百分比的标准可以是针对特定次要物质频率的次要物质峰与主要物质峰的相对峰高(例如,约10%以下,约5%以下,约2.5%以下或约1%以下)。在某些实施方式中,例如,根据特定转化,相对于次要链终止剂浓度,经调整的主要物质链终止剂百分比可以为约0.1%至约10%,约0.1%至约5%,约1%至约4%或约0.1%,0.2%,0.3%,0.4%,0.5%,0.6%,0.7%,0.8%,0.9%,1.0%,1.1%,1.2%,1.3%,1.4%,1.5%,1.6%,1.7%,1.8%,1.9%,2.0%,2.2%,2.3%,2.4%,2.5%,2.6%,2.7%,2.8%,2.9%,3.0%,3.1%,3.2%,3.3%,3.4%,3.5%,3.6%,3.7%,3.8%,3.9%,4.0%,4.1%,4.2%,4.3%,4.4%,4.5%,4.6%,4.7%,4.8%,4.9%,5.0%,5.1%,5.2%,5.3%,5.4%,5.5%,5.6%,5.7%,5.8%,5.9%,6.0%,6.1%,6.2%,6.3%,6.4%,6.5%,6.6%,6.7%,6.8%,6.9%,7.0%,7.1%,7.2%,7.3%,7.4%,7.5%,7.6%,7.7%,7.8%,7.9%,8.0%,8.1%,8.2%,8.3%,8.4%,8.5%,8.6%,8.7%,8.8%,8.9%,9.0%,9.1%,9.2%,9.3%,9.4%,9.5%,9.6%,9.7%,9.8%或9.9%。

[0048] 在某些实施方式中,例如,根据特定转化,相对于次要物质链终止剂浓度,经调整的主要物质链终止剂百分比可以为约0.01%至约5%,约0.01%至约4%或约0.01%,0.02%,0.03%,0.04%,0.05%,0.06%,0.07%,0.08%,0.09%,0.1%,0.2%,0.3%,0.4%,0.5%,0.6%,0.7%,0.8%,0.9%,1.0%,1.1%,1.2%,1.3%,1.4%,1.5%,1.6%,1.7%,1.8%,1.9%,2.0%,2.2%,2.3%,2.4%,2.5%,2.6%,2.7%,2.8%,2.9%,3.0%,3.1%,3.2%,3.3%,3.4%,3.5%,3.6%,3.7%,3.8%,3.9%,4.0%,4.1%,4.2%,4.3%,4.4%,4.5%,4.6%,4.7%,4.8%,4.9%或5.0%。

[0049] 在一些实施方式中,对于多个扩增引物对中的至少一个引物对,一个引物的浓度低于另一个引物。或者,对于多个引物对中的每一个,一个引物的浓度低于另一个引物。在这样的实施方式中,特定核酸靶区域(次要核酸物质中的区域)的较低浓度引物可以与该核酸靶区域的延伸(UEP)引物的取向相同。此外,扩增步骤的扩增产物可以是单链核酸分子。

[0050] 还提供了用于在延伸之前进行次要和主要核酸物质的多重化扩增的方法,所述方法包括:a)设计7对或更多对一级和二级引物,其中每个引物对扩增次要核酸物质上的特定核酸靶区域;b)形成含有多个引物对和一个或多个次要和主要核酸物质以扩增多个核酸靶区域的混合物;和c)在允许扩增超过60%的7个或更多个核酸靶区域的扩增反应条件下使步骤b)的混合物进行多次热循环,其中扩增反应条件包括dNTP和MgCl₂,并且其中游离Mg²⁺浓度在1.0-2.0mM之间(例如,1.1、1.2、1.3、1.4、1.5、1.6、1.7、1.8、1.9)。在其他实施方式中,一级和二级引物对的数量可以选自8个或更多个,9个或更多个,10个或更多个,11个或更多个,12个或更多个,或13个或更多个。

[0051] 在上述每种方法的某些实施方式中,序列标签连接到每个引物对的一个或两个扩增引物的5'末端。在一些实施方式中,本文提供的扩增反应条件允许扩增一定百分比的7个或更多个核酸靶区域,所述百分比选自7个或更多个核酸靶区域的大于60%,65%,70%,75%,80%,85%,90%,91%,92%,93%,94%,95%,96%,97%,98%或99%,高达100%。本文提供的条件适用于使用多种扩增引物对和来自多种主要和次要核酸物质的混合物的7-重或更多扩增反应的多次多重化反应。

[0052] 下述说明、权利要求和附图中进一步描述某些实施方式。

[0053] 附图简要说明

[0054] 附图示出了该技术的某些非限制性实施方式,并且不一定按比例绘制。

[0055] 图1显示使用本文提供的方法检测相对于主要等位基因以5%频率存在的次要等位基因。

[0056] 图2显示随着时间的推移,将ddNTP链终止剂以野生型特异ddNTP ($C[WT]_{ddNTP}$) 浓度与突变体特异ddNTP ($C[Mut]_{ddNTP}$) 浓度的不同比例掺入对应于野生型或突变型等位基因的单碱基延伸产物中。

[0057] 图3显示使用本文提供的方法检测相对于主要等位基因以5%、2.5%或1.25%频率存在的次要等位基因。

[0058] 图4显示了每个稀释系列样品类型(0%,1%,2.5%和5%)的每个转化(单独图)检测灵敏度(y轴)。

[0059] 图5显示5%黑猩猩(chimp)稀释系列的观察到的黑猩猩频率(y轴)。图代表独特的转化,并且数据点表示在这些转化中的试验中观察到的黑猩猩频率。

[0060] 图6显示不同延伸反应条件(x轴)的试验延伸率(y轴)的方框图。

具体实施方式

[0061] 本文提供了用于鉴定存在于含有一种或多种次要核酸物质和一种或多种主要核酸物质的混合物的样品中的多种次要核酸种中一种或多种存在或不存在的方法。还提供了相对于样品中主要核酸物质的量来对次要核酸物质进行定量(例如,频率或拷贝数)的方法。

[0062] 本文提供的方法可用于分析相对于主要核酸物质以约0.25%至多约50%的频率或拷贝数存在的次要核酸物质,特定范围包括但不限于约或等于0.25%至约或等于45%,40%,35%,30%,25%,20%,15%,10%,9.5%,9%,8.5%,8%,7.5%,7%,6.5%,6%,5.5%,5%,4.5%,4%,3.5%,3%,2.5%,2%,1.5%或1%。在某些实施方式中,次要核酸物质以相对于主要核酸物质的频率或拷贝数约1%或约2%至约10%,9.5%,9%,8.5%,8%,7.5%,7%,6.5%,6%,5.5%,5%,4.5%,4%,3.5%,3%或2.5%的频率或拷贝数存在。在实施方式中,次要核酸物质以相对于主要核酸物质的频率或拷贝数约1%或约2%至小于约10%的频率存在。在实施方式中,次要核酸物质以相对于主要核酸物质的频率或拷贝数约5%的频率存在。

[0063] 本文的方法提供了相对于主要核酸物质的次要核酸物质的检测灵敏度(检测限),同时允许相对于主要核酸物质的存在和/或量的可靠的鉴定和/或定量

[0064] 例如,在诸如iPLEX™或均质MassExtend® (hME)的方法中(参见例如,美国公开专利申请号2013/0237428 A1,美国专利号8,349,566和美国专利8,003,317,其内容通过引用整体并入本文),将次要核酸物质(例如,突变等位基因)和主要核酸物质(例如,野生型等位基因)的混合物进行聚合酶链反应(PCR)扩增,使用一组扩增引物,聚合酶和脱氧核苷酸(dNTP),从而产生野生型和突变物质的扩增子。在用虾碱性磷酸酶(SAP)处理以脱磷酸化未掺入的dNTP之后,使用延伸引物(未延伸的引物或UEP),聚合酶和终止混合物扩增扩增子混合物,其包括链终止剂(例如,双脱氧核苷酸或ddNTP)。UEP与扩增子杂交,并延伸到突变体和野生型物质之间的差异位点(即,延伸在突变体和野生型物质之间的碱基不同的突变位点处的停止以产生单碱基扩展产物或SBE,如iPLEX™)或差异位点过后的几个碱基(例如,2-

3个碱基)(例如,在hME方法中)。然后可以处理所得到的延伸产物(例如,通过在质谱之前脱盐),并基于相对于野生型等位基因的检测信号(例如,质量)的差异来分析突变等位基因的存在。

[0065] 上述iPLEX™和均质MassExtend® (hME)方法在延伸步骤中使用ddNTP的等摩尔混合物来产生对应于野生型和突变体物质的延伸产物。因此,在iPLEX™和均质MassExtend® (hME)方法中,除主要核酸物质以相对于次要核酸物质大量过量存在之外,所有其他因素相等,大多数UEP与主要的核酸物质并且使用对于主要核酸物质特异的链终止剂来延伸。较少的UEP分子可用于杂交和延伸次要核酸物质。这削弱了对应于次要核酸物质的检测信号的量级,其被来自主要核酸物质的主要检测信号所掩盖,并且可能被背景噪声所包含。

[0066] 相比之下,本文提供的方法相对于针对突变体或次要核酸物质特异的链终止剂,使用对野生型或主要核酸物质特异的链终止剂的限制浓度进行延伸步骤。当以低于针对一种或多种次要核酸物质特异的链终止剂的浓度的浓度加入对主要核酸物质特异的链终止剂时,将使用主要核酸物质作为模板延伸较小数量的UEP,从而留下更多数量的可用于使用次要核酸物质作为模板的延伸的UEP。因此,在一系列延伸循环中,对应于次要核酸物质的更多延伸产物积累,从而增加对应于次要核酸物质的检测信号并提高检测的灵敏度(限)。本文提供的方法,其中链终止剂的浓度倾斜以有利于延伸由次要核酸物质产生的扩增子,可用于分析以小于感兴趣样品中的主要核酸物质的频率的频率存在的任何次要核酸物质。在实施方式中,所述方法对于检测以小于主要核酸物质的拷贝数或频率的10%的拷贝数或频率存在的次要核酸物质是有利的,例如以主要核酸物质的拷贝数或频率的约0.25%至小于10%,约0.5%至小于10%,约1%至小于10%或约2%至小于10%。

[0067] 图1显示了本文提供的方法的实施方式。将含有次要核酸物质(少数等位基因)和主要核酸物质(多数等位基因)的混合物的样品根据iPLEX™方法或根据本文提供的方法经过扩增和延伸反应,从而提供用于检测和分析的延伸产物,其中少数等位基因以相对于多数等位基因5%频率存在。最左边的图显示了使用iPLEX™方法的分析结果,其中使用等摩尔浓度的链终止剂。如该图所示,对应于来自多数等位基因的延伸产物的左侧的信号峰值占主导,使得右侧的少数峰降低到背景噪声的水平并且是不可检测的。中图显示当多数等位基因的特异性链终止剂的浓度为少数等位基因的特异性链终止剂的浓度的20% (1/5)时获得的结果。如中图所示,来自少数等位基因延伸产物(右峰)的检测信号的强度相对于来自多数等位基因延伸产物的信号(左峰)的检测信号现在更高且更可见。然而,少数等位基因信号仍然很小,并且接近背景噪声的水平。最右侧的图显示当多数等位基因的特异性链终止剂的浓度为少数等位基因的特异性链终止剂的浓度的约6-7% (十五分之一)时获得的结果。如右图所示,来自少数等位基因延伸产物(右峰)的信号现在与来自多数等位基因的信号(左峰)相当。因此,图1表明,通过使链终止剂的浓度有利于次要核酸物质偏斜,小于10%的频率的次要核酸物质(其不能通过诸如iPLEX™的方法有效检测到)可通过本文提供的方法检测。

[0068] 调整对于主要核酸物质(C[WT])特异的链终止剂的浓度相对于次要核酸物质(C[Mut])特异性链终止剂的浓度,使得两种浓度,C[WT]:C[Mut]的比例不如使次要核酸物质不可检测到的比例那么高。在一些实施方式中,可以根据每个次要核酸物质相对于相应的

主要核酸物质的拷贝数或频率来调整每个次要物质的比例。通常,对于以相对于主要核酸物质10%或更低的频率存在的次要核酸物质,主要核酸物质特异性链终止剂的浓度 $C[WT]$ 小于对于次要核酸物质特异的链终止剂浓度($C[Mut]$)的20%, $C[WT]$ 通常为 $C[Mut]$ 的约0.5%至小于约20%之间, $C[Mut]$ 的约0.5%至小于约15%, $C[Mut]$ 的约1%至约15%, $C[Mut]$ 的约1%至约10%, $C[Mut]$ 的约2%至约10%或 $C[Mut]$ 的约1%,1.5%,2%,2.5%,3%,3.5%,4%,4.5%,5%,5.5%,6%,6.5%,7%,7.5%,8%,8.5%,9%,9.5%或10%。在某些实施方式中, $C[WT]$ 为 $C[Mut]$ 的约0.1%至约10%, $C[Mut]$ 的约0.01%至约10%,或 $C[Mut]$ 的约0.01%,0.05%,0.1%,0.5%,1%,1.5%,2%,2.5%,3%,3.5%,4%,4.5%,5%,5.5%,6%,6.5%,7%,7.5%,8%,8.5%,9%,9.5%或10%。两个浓度的比例不能太低,即 $C[WT]$ 不能远小于 $C[Mut]$ 。当 $C[WT]$ 远低于 $C[Mut]$ 时,主要核酸物质获得的延伸产物的量将较低,导致主要核酸物质的低检测信号。然而,此外,由于次要核酸物质通常以相对于样品中的主要核酸物质的10%以下的量存在,因此对可产生多少次要延伸产物存在限制,即使 $C[Mut]$ 远高于 $C[WT]$,来自次要物质延伸产物的信号也不太可能远高于来自主要物质延伸产物的信号。因此,当 $C[WT]$ 远低于 $C[Mut]$ 时,来自主要核酸和次要核酸物质的信号彼此相当可能太低并且太接近可检测的背景噪音。

[0069] 图2描绘了在各种主要物质特异性ddNTP链终止剂($C[WT]_{ddNTP}$)和次要物质特异性ddNTP链终止剂($C[Mut]_{ddNTP}$)的各种相对浓度下,来自次要核酸物质的检测信号相对于来自主要核酸物质的检测信号的分析。在PCR扩增和SAP处理之后,样品中次要和主要核酸物质的扩增子的量足够用作延伸模板,例如,通过使用ddNTP链终止剂核苷酸和扩增引物(UEP)的单碱基延伸反应。UEP引物的摩尔浓度高于PCR扩增引物的摩尔浓度,允许在延伸期间重复循环进行信号扩增。由于与次要核酸物质相比,对应于主要核酸物质的扩增子大大过量,所以大多数UEP被针对主要核酸物质特异的链终止剂核苷酸($[WT]_{ddNTP}$)延伸,并且因此只有部分UEP将被对于次要核酸物质特异的链终止剂核苷酸($[Mut]_{ddNTP}$)延伸。在第一个延伸循环之后,使用主要核酸物质作为模板延伸大部分UEP引物,降低了下一延伸循环可用的UEP引物的数量。这些延伸的UEP引物,即主要核酸物质的单碱基延伸产物(SBE-WT)也将作为未延伸引物(UEP)的竞争者,因为它们将结合模板主要核酸物质和次要核酸物质扩增子(SBE-WT将与UEP竞争主要和次要核酸物质扩增子,从而减少使用UEP产生的扩展产物的量)。如果主要核酸物质特异性链终止剂ddNTP的浓度($C[WT]_{ddNTP}$)等于对次要核酸物质特异性链终止剂ddNTP的浓度($C[Mut]_{ddNTP}$),并且如果次要核酸物质以低于样品中主要核酸物质的10%的频率或拷贝数存在,则对应于次要核酸物质的信号相对于对应于次要核酸物质的信号可能如此低使其在背景噪音上无法检测到。这在图2的顶图中(从左到右:来自未延伸引物的信号,来自主要核酸物质延伸产物的信号,来自次要核酸物质延伸产物的信号;水平虚线代表背景,即,检测限不能与噪音区分)显示。

[0070] 另一方面,如果 $C[WT]_{ddNTP}$ 小于 $C[Mut]_{ddNTP}$,则样品中较大量的主要核酸物质的UEP引物的消耗较低,允许更多的UEP引物用于在随后的循环中从对应于次要核酸物质的扩增子延伸。由于 $C[Mut]_{ddNTP}$ 明显高于 $C[WT]_{ddNTP}$,反应在每个循环期间完成,因此超过单碱基延伸反应的总循环次数,将用对于次要核酸物质特异的ddNTP延伸高得多数量的UEP,从而相对于对应于主要核酸物质的信号产生更高的信号,并且与背景噪音不同。这在图2的顶部的第二和第三图中(从左到右:来自未延伸引物的信号,来自主要核酸物质延伸产物的信号,

来自次要核酸物质延伸产物的信号;水平虚线代表背景,即,检测限不能与噪音区分)显示。因此,本文提供的方法允许通过使用相对于 $[\text{Mut}]_{\text{ddNTP}}$ 较低浓度的 $[\text{WT}]_{\text{ddNTP}}$,检测以相对于主要核酸物质的频率小于10%的频率存在的次要核酸物质。然而,如果在延伸反应中加入太少的 $[\text{WT}]_{\text{ddNTP}}$ 分子,则主要核酸物质扩增子的延伸将基本上被抑制,当与低频率的少量核酸物质组合时,可以产生对于主要和次要核酸物质的低峰信号。在该方法的实施方式中,来自主要核酸物质的信号可以用作阳性对照,以确保方法的完整性。因此,如果来自主要核酸物质的信号太低,则不再能够将其检测为阳性对照,并且另外,来自次要核酸物质的信号也可能对于准确鉴定而言太低。这在图2的底图中(从左到右:来自未延伸引物的信号,来自主要核酸物质延伸产物的信号,来自次要核酸物质延伸产物的信号;水平虚线代表背景,即,检测限不能与噪音区分)显示。

[0071] 如图3所示,本文提供的方法可用于检测以小于相对于样品中的主要核酸物质的2%,低至1.25%的频率存在的次要核酸物质。图3显示了当样品进行扩增,随后在以下三种条件下延伸时获得的结果的比较:(A)对于主要核酸物质特异的链终止剂ddNTP的浓度(野生型等位基因; $C[\text{WT}]_{\text{ddNTP}}$)等于对于次要核酸物质特异的链终止剂ddNTP的浓度(突变等位基因; $C[\text{Mut}]_{\text{ddNTP}}$);(B) $C[\text{WT}]_{\text{ddNTP}}$ 小于 $C[\text{Mut}]_{\text{ddNTP}}$ 浓度的约20%(即 $C[\text{WT}]_{\text{ddNTP}}:C[\text{Mut}]_{\text{ddNTP}}$ 的比例为0.2);和(C) $C[\text{WT}]_{\text{ddNTP}}$ 小于 $C[\text{Mut}]_{\text{ddNTP}}$ 浓度的约5%(即 $C[\text{WT}]_{\text{ddNTP}}:C[\text{Mut}]_{\text{ddNTP}}$ 的比例为0.05)。如图3所示,当 $C[\text{WT}]_{\text{ddNTP}}:C[\text{Mut}]_{\text{ddNTP}}$ 的比例为1,即两个链终止剂浓度相等(例如,在iPLEX™方法中使用)时,5%、2.5%和1.25%的突变体等位基因频率都低于检测限(LOD)。当 $C[\text{WT}]_{\text{ddNTP}}:C[\text{Mut}]_{\text{ddNTP}}$ 的比例降低到0.2时,可以检测到5%的突变体等位基因频率,而2.5%和1.25%的突变等位基因频率仍然低于检测限。当 $C[\text{WT}]_{\text{ddNTP}}:C[\text{Mut}]_{\text{ddNTP}}$ 的比例进一步降低到0.05时,可以检测到所有三个突变体等位基因频率(5%、2.5%和1.25%)。

[0072] 本文提供的方法还可用于相对于样品中主要核酸物质的量的定量一种或多种次要核酸物质(例如,频率或拷贝数)。例如,如果已知主要核酸物质和次要核酸物质特异链终止剂的相对浓度(即 $C[\text{WT}]$ 和 $C[\text{Mut}]$ 是已知的),则次要和主要核酸物质可使用归一化系数从 $C[\text{WT}]:C[\text{Mut}]$ 的比例中得到。在该方法的实施方式中,延伸反应在每个延伸循环中线性扩增,并且次要和主要核酸物质的相对量与 $C[\text{WT}]:C[\text{Mut}]$ 的比例成反比。因此,可检测量的对应于主要核酸物质的信号的存在为该方法的完整性提供了阳性对照,并且也可用于相对于样品中主要核酸物质的量定量一种或多种次要核酸物质的量(例如,频率或拷贝数)。

[0073] 本文提供的方法可以适用于多重化形式。例如,当分析作为多种主要核酸物质的变体(例如,突变体,等位基因,多态性,缺失变体,插入变体等)的多种次要核酸物质时,可以在一系列反应容器、隔室或容器中进行多重化,其中每个容器具有仅包含一种链终止剂(例如,ddATP,ddCTP,ddUTP,ddGTP或ddTTP)的延伸反应混合物。具有相同特异性链终止剂(例如,这些主要核酸物质的全部单碱基延伸产物终止于ddATP)的所有主要核酸物质可以在一个容器中延伸。类似地,具有相同特异性链终止剂的所有次要核酸物质可以在另一个容器中延伸。使用一系列容器,可以分析来自多种目标次要核酸物质的信号,相对于扩增主要核酸物质的扩增子的容器中使用较低浓度的主要核酸物质特异性链终止剂产生的信号,每个容器具有较高浓度的针对各组次要核酸物质的链终止剂,其共有共同特异性链终止剂。在多重化方法的一个实施方式中,分析了多种次要核酸物质,其全部是相同主要核酸物质的变体。

[0074] 通过本文提供的方法获得的对应于主要和次要核酸物质的延伸产物可以通过多种方法检测。例如,延伸引物(UEP)和/或链终止剂可以用允许检测信号和/或定量信号的任何类型的化学基团或部分标记,包括但不限于,质量标记、放射性分子、荧光分子、抗体、抗体片段、半抗原、碳水化合物、生物素、生物素衍生物、磷光部分、发光部分、电化学发光部分、在氧化或还原时产生电化学信号的部分,例如铁,钆或钽的复合物(参见例如,由Genmark诊断公司(Genmark Diagnostics, Inc)使用的eSensor技术,例如Pierce等, J.Clin.Microbiol., 50(11):3458-3465(2012)中所述)、色度部分和具有可检测的电子自旋共振、电容、介电常数或电导率的部分、或其标记的任何组合。

[0075] 对应于次要和主要核酸物质的标记的延伸产物可以通过多种方法分析,包括但不限于,质谱, MALDI-TOF质谱, 荧光检测, DNA测序凝胶, 自动DNA测序机上的毛细管电泳, 微通道电泳和其他测序方法, 质谱, 飞行时间质谱, 四极杆质谱, 扇形磁场质谱, 电子质谱红外光谱, 紫外光谱, potentiostatic电流分析, 电流/电化学信号测量或通过DNA杂交技术, 包括Southern印迹, Slot印迹, Dot印迹和DNA微阵列, 其中DNA片段可用作“探针”和“靶标”, ELISA, 荧光试验, 荧光共振能量转移(FRET), SNP-IT, GeneChips, HuSNP, BeadArray, TaqMan试验, Invader试验, MassExtend®, 或MassCleave®方法。

[0076] 本文提供的方法可用于例如:(a)快速确定样品中是否存在特定靶序列(例如,包含遗传变异的靶序列);(b)进行混合物分析,例如鉴定混合物和/或其组合物或确定混合物中靶序列的频率(例如,混合群落,准种);(c)检测样品中的序列变异(例如,突变,单核苷酸多态性);(d)进行单倍型试验;(e)进行微生物(例如,病原体)分型;(f)检测样品中微生物靶序列的存在或不存在;(g)识别疾病标志物;(h)检测微卫星;(i)识别短串联重复;(j)识别一种或多种生物体;(k)检测等位变异;(l)确定等位基因频率;(m)确定甲基化模式;(n)进行表观遗传试验;(o)重新序列生物分子的区域;(p)在人类临床研究和医学中进行分析(例如,癌症标志物检测,序列变异检测;对特定药物给药有利或不利的序列特征的检测),(q)进行HLA分型;(r)进行取证分析;(s)进行疫苗质量控制分析;(t)监测治疗;(u)进行矢量特征分析;(v)进行疫苗或生产菌株质量控制和(w)测试菌株特征(x)植物。这些方法也可以用于,例如,各种领域,包括但不限于商业,教育,医疗,农业,环境,疾病监测,军事防御和法医领域。

[0077] 核酸

[0078] 如本文所用的术语“核酸”是指寡核苷酸或多核苷酸,包括但不限于,天然核酸(例如,脱氧核糖核酸(DNA),核糖核酸(RNA)),合成核酸,非天然核酸(例如,肽核酸(PNA)),未修饰的核酸,经修饰的核酸(例如,甲基化DNA或RNA,标记的DNA或RNA,具有一个或多个修饰的核苷酸的DNA或RNA)。以“多核苷酸”提及核酸是指通过共价键连接的两个或更多个核苷酸或核苷酸类似物。核酸可以是适用于本文所述方法的任何类型的核酸。某些实施方式中的核酸可以是DNA(例如,互补DNA(cDNA),基因组DNA(gDNA),质粒和载体DNA等),RNA(例如,病毒RNA,消息RNA(mRNA),短抑制性RNA(siRNA),核糖体RNA(rRNA),tRNA等)和/或DNA或RNA类似物(例如,含有碱基类似物,糖类似物和/或非天然骨架等)。核酸可以是可用于进行本文所述方法的任何形式(例如线性、环状、超螺旋、单链、双链等)。在某些实施方式中,核酸可以是或可以来自,质粒、噬菌体、自主复制序列(ARS)、着丝粒、人工染色体、染色体、细胞、细胞核或细胞的胞质。在一些实施方式中,核酸来自单个染色体(例如,核酸样品可来自从

二倍体生物所得样品的一个染色体)。在胎儿核酸的情况下,核酸可以来自父系等位基因,母系等位基因或母系和父系等位基因。

[0079] 如本文中关于靶核酸、扩增子、引物、序列标签、多核苷酸或寡核苷酸所使用的术语“次要核酸物质”是指在比对核苷酸序列时,与来自另一种核酸“主要核酸物质”的核苷酸序列有一个或多个核苷酸差异的核酸。因此,当比对时,两种物质的序列相差一个或多个核苷酸(例如,约1、2、3、4、5、10、15、20、25、30、35、40、45、50、55、60、65、70、75、80、85、90、95、100或大于100个核苷酸差异)时,第一核酸物质与第二核酸物质不同。在某些实施方式中,通过次要或主要核酸物质或延伸的寡核苷酸物质的PCR扩增获得的核酸物质(例如,次要或主要核酸物质),扩增子物质的数量包括但不限于约2至约10000种核酸物质,约2至约1000种核酸物质,约2至约500种核酸物质,或有时约2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、25、30、35、40、45、50、55、60、65、70、75、80、85、90、95、100、125、150、175、200、225、250、275、300、325、350、375、400、425、450、475、500、600、700、800、900、1000、2000、3000、4000、5000、6000、7000、8000、9000或10000种核酸物质。

[0080] 如本文所用,术语“次要核酸物质”或“靶核酸物质”可以指样品中任何感兴趣的核酸物质。次要核酸物质可包括但不限于(i)在两种或更多种可能的等位基因中的特定次要等位基因,和(ii)具有或不具有特定突变,核苷酸取代,序列变异,重复序列,标志或区别序列的核酸。如本文所用,术语“主要核酸物质”是指通过一个或多个特征与次要核酸物质不同的核酸物质,并且以比次要核酸物质的频率或拷贝数更高的频率或拷贝数存在于样品中。如本文所用,术语“遗传变异”是指通过一个或多个特征不同的核酸物质。如本文所用,术语“变异”是指通过一个或多个特征不同的核酸物质。特征包括,不限于,一个或多个甲基或甲基化状态,一个或多个磷酸基团,一个或多个乙酰基,以及一个或多个核苷酸的一个或多个缺失、添加或取代。一个或多个核苷酸的一个或多个缺失、添加或取代的实例包括但不限于特定突变的存在或不存在,核苷酸取代的存在或不存在(例如,单核苷酸多态性(SNP)),重复序列(例如,二,三,四,五核苷酸重复)的存在或不存在,标记(例如,微卫星)的存在或不存在以及区分序列(例如,区分一种生物体与另一种生物体的序列(例如,区分一种病毒株与其他病毒株的序列))存在或不存在。可以通过任何已知的方法来区分不同的靶核酸,例如,通过质量、结合、可区分标签等,如本文所述。

[0081] 在本文提供的方法中,样品可以含有一种或多种“次要核酸”物质和一种或多种“主要核酸”物质的混合物,或者可以通过组合超过一种含有次要核酸物质和主要核酸物质的样品来产生混合物。次要核酸物质包括主要核酸物质的变体,并且可以包括但不限于野生型(主要核酸)等位基因的突变体(次要核酸),在超过一种宿主中发现的基因的变体(例如,作为正常健康基因(主要核酸)的变体的病毒致癌基因(次要核酸)),多态性,包括单核苷酸多态性(SNP),主要核酸物质的插入,缺失或其他突变形式。相对于主要核酸物质,次要核酸物质在样品(或样品组合)中以约0.25%至至多约50%的频率或拷贝数存在,特定范围包括但不限于约或等于0.25%至约或等于45%,40%,35%,30%,25%,20%,15%,10%,9.5%,9%,8.5%,8%,7.5%,7%,6.5%,6%,5.5%,5%,4.5%,4%,3.5%,3%,2.5%,2%,1.5%或1%。在某些实施方式中,次要核酸物质以相对于主要核酸物质的频率或拷贝数约1%或约2%至约10%,9.5%,9%,8.5%,8%,7.5%,7%,6.5%,6%,5.5%,5%,4.5%,4%,3.5%,3%或2.5%的频率或拷贝数存在。在实施方式中,次要核酸物质以相对

于主要核酸物质的频率或拷贝数约1%或约2%至小于约10%的频率存在。

[0082] 如本文所用,术语“多种”靶核酸、次要核酸物质和/或主要核酸物质是指超过一种靶核酸物质、次要核酸物质和/或主要核酸物质。多种可以是约2至约10000种核酸物质,约2至约1000种核酸物质,约2至约500种核酸物质,或有时约2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、25、30、35、40、45、50、55、60、65、70、75、80、85、90、95、100、125、150、175、200、225、250、275、300、325、350、375、400、425、450、475、500、600、700、800、900、1000、2000、3000、4000、5000、6000、7000、8000、9000或10000种核酸物质,在某些实施方式中。核酸的检测或鉴定导致对次要核酸物质的检测,并且可以指示特定突变、序列变异(突变或多态性)或遗传变异(例如,序列变异、序列差异或多态性)的存在或不存在。核酸的检测或鉴定也通常导致主要核酸物质的检测或鉴定,其可以作为阳性对照和/或定量次要核酸物质的基础。在多种次要核酸和主要核酸物质内,可以检测和/或定量相同或不同的物质;检测和/或定量全部是相同主要核酸物质的次要核酸物质或多种主要核酸物质的变体的多种次要核酸物质。

[0083] 在一些实施方式中,寡核苷酸物质与核酸模板(例如,扩增子)杂交从而形成双链核酸,并且与模板杂交的寡核苷酸物质在本文中称为杂交寡核苷酸物质。在一些实施方式中,杂交的寡核苷酸物质可以包含一个或多个不与模板杂交的核苷酸。例如,杂交的寡核苷酸物质可包含一个或多个错配的核苷酸(例如,非互补核苷酸),并且有时可包含不杂交的核苷酸的5'和/或3'区域。在一些实施方式中,杂交的寡核苷酸物质包括标签(例如,质量可区分的标签、序列标签、发光标签或放射性标签)。在一些实施方式中,杂交的寡核苷酸物质包含捕获剂(例如,生物素或结合对的任何成员)。在一些实施方式中,杂交的寡核苷酸物质包含终止核苷酸。

[0084] 术语“链终止剂”与本文中的“链终止剂试剂”或“链终止剂”可互换使用,是指当加入延伸引物时停止延伸反应的分子。链终止剂可以包括核苷酸类似物,当存在于多核苷酸链中时,其阻止链的进一步延伸。作为链终止核苷酸的示例性的链终止剂包括双脱氧核苷酸,例如ddA(双脱氧腺嘌呤),ddT(双脱氧胸腺嘧啶),ddC(双脱氧胞嘧啶),ddG(二脱氧鸟嘌呤)和ddU(二脱氧尿嘧啶)。

[0085] 如本文所用,术语“核苷酸”是指天然和非天然核苷酸。核苷酸包括但不限于天然存在的核苷单-,二-和三磷酸酯:脱氧腺苷单-,二-和三磷酸酯;脱氧鸟苷单-,二-和三磷酸酯;脱氧胸苷单-,二-和三磷酸酯;脱氧胞苷单-,二-和三磷酸酯;脱氧尿苷单-,二-和三磷酸酯;和脱肌苷单-,二-和三磷酸酯(本文中分别称为dA,dG,dT,dC,dU和dI,或A,G,T,C,U和I)。核苷酸还包括但不限于修饰的核苷酸和核苷酸类似物。修饰的核苷酸和核苷酸类似物包括但不限于,去氮嘌呤(deazapurine)核苷酸,例如7-去氮-脱氧鸟苷(7-脱氮-DG)和7-脱氮-脱氧腺苷(7-脱氮腺苷)单-,二-和三磷酸酯,氘代-脱氧胸苷(氘代-dT)单-,二-和三磷酸酯,甲基化核苷酸,例如,5-甲基脱氧胞苷三磷酸酯,¹³C/¹⁵N标记核苷酸和脱氧肌苷单-,二-和三磷酸酯。可以使用各种功能和附着位置的组合来获得修饰的核苷酸,同位素富集的核苷酸,耗尽的核苷酸,带标签和标记的核苷酸和核苷酸类似物。

[0086] 本文中关于核酸使用的术语“组合物”是指包括一种或多种核酸的有形物品。组合物有时是从来源提取的样品,而且是来源处的所有样品的组合物,并且有时是一种或多种核酸的来源。组合物可以包含核酸。在一些实施方式中,组合物可以包含基因组DNA。在一些

实施方式中,组合物可以包含母体DNA,胎儿DNA或母体和胎儿DNA的混合物。在一些实施方式中,组合物可以包含基因组DNA的片段。在一些实施方式中,组合物可以包含衍生自病毒,细菌,酵母,真菌,哺乳动物或其混合物的核酸。

[0087] 核酸样品可以衍生自一个或多个来源,并且可以含有次要核酸物质和主要核酸物质的混合物。样品也可以组合以产生包括次要核酸物质和主要核酸物质的混合物。例如,可以从生物体、矿物或地质场地(例如,土壤,岩石,矿物沉积物,化石)或法医现场(例如,犯罪现场,违禁品或疑似违禁品)收集样品。因此,源可能是环境的,例如地质,农业,战区或土壤来源。来源还可以来自任何类型的生物体,例如任何植物、真菌、原生生物、原核生物、病毒或动物,包括但不限于人,非人,哺乳动物,爬行动物,牛,猫,狗,山羊,猪,猴,猿,大猩猩,公牛,牛,熊,马,羊,家禽,小鼠,大鼠,鱼,海豚,鲸鱼和鲨鱼,或任何可能具有可检测核酸的动物或生物体。来源还可以指生物体的不同部分,例如,内部、外部、活或非活细胞、组织、流体等。因此,样品可以是“生物样品”,其是指从活来源或以前活来源获得的任何材料,例如,动物如人或其他哺乳动物,植物,细菌,真菌,原生生物或病毒。来源可以是任何形式,包括但不限于固体材料,例如组织,细胞,细胞沉淀,细胞提取物或活检,或生物流体如尿液,血液,唾液,羊水,来自感染或炎症区域的渗出液,或含有颊细胞的口腔清洗液,头发,脑脊髓液和滑液以及器官。样品还可以在与另一样品不同的时间点分离得到,其中各样品来自相同或不同的来源。核酸可来自核酸库,例如cDNA或RNA库。核酸可以是样品中核酸分子的核酸纯化或分离和/或扩增的产物。提供用于本文所述的序列分析方法的核酸可以包含来自一个样品或两个或更多个样品的核酸(例如,来自1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、25、50、75、100、200、300、400、500、600、700、800、900或1000更多个样品)。

[0088] 可以在本文提供的方法期间、之前或之后以各种方式处理核酸。例如,可以将核酸的尺寸减小(例如,剪切,由核酸酶或限制酶消化,去磷酸化,去甲基化),尺寸增加(例如,磷酸化,与甲基化特异试剂反应,附着于可检测标记),用核酸切割抑制剂等处理。

[0089] 在某些实施方式中,核酸可以根据本文提供的方法进行分析,而不进行处理。在一些实施方式中,提供核酸用于在处理后进行本文所述的方法。例如,可从样品提取、分离、纯化或扩增核酸。如本文所用术语“分离”指将核酸从其原始环境中取出(例如,天然发生核酸的天然环境或外源表达核酸的宿主细胞),因此核酸从其原始环境“通过人工”而被改变。与来源样品中的组分含量相比,分离的核酸一般带有较少的非核酸组分(例如,蛋白质、脂质)。包含分离核酸的组合物可以是基本分离的(例如,约90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或大于99%不含非核酸组分)。本文所用术语“纯化”是指提供的核酸与其所衍生自的样品来源相比包含更少的核酸种类。包含核酸的组合物可以是基本纯化的(例如,约90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或大于99%不含其它核酸种类)。

[0090] 核酸可以通过,在某些实施方式中,在为本文所述的方法提供核酸之前产生核酸片段的方法进行处理。在一些实施方式中,经过片段化或剪切的核酸可具有约5-约10,000个碱基对、约100-约1.00个碱基对、约100-500个碱基对或约10、15、20、25、30、35、40、45、50、55、60、65、70、75、80、85、90、95、100、200、300、400、500、600、700、800、900、1000、2000、3000、4000、5000、6000、7000、8000、9000或10000个碱基对的标称、平均或等比中数长度。可

通过本领域已知的任何合适方法产生片段,且核酸片段的平均、等比中数或标称长度可通过选择适当的片段生成方法而加以控制。在某些实施方式中,较短长度的核酸可用来分析几乎不含序列变异和/或含有较大量的已知核苷酸序列信息的序列。在一些实施方式中,较长长度的核酸可用来分析含有较多序列变异和/或含有较少量未知核苷酸序列信息的序列。

[0091] 扩增和延伸

[0092] 在本文提供的方法的实施方式中,可以在某些实施方式中扩增核酸物质(例如,次要和/或主要核酸物质)。如本文所用,术语“扩增”及其语法变体是指产生模板核酸拷贝的方法。例如,核酸模板可以进行线性或指数生成具有与模板的核苷酸序列相同或基本上相同的核苷酸序列或模板的一部分的两个或多个核酸扩增子(拷贝)的过程。在某些实施方式中,核酸扩增通常是特异性的(例如,扩增子具有相同或基本上相同的序列),并且有时可以是非特异性的(例如,扩增子具有不同的序列)。当样品中存在的靶序列的量低时,核酸扩增有时是有益的。通过扩增靶序列并检测合成的扩增子,可以提高试验的灵敏度,因为在用于检测靶核酸的试验开始时需要较少的靶序列。在某些实施方式中,核酸物质(次要或主要核酸物质)有时在杂交延伸寡核苷酸(引物,即UEP)之前不被扩增。

[0093] 扩增条件是已知的,并且可以针对将被扩增的特定核酸选择。扩增条件包括某些试剂,其中一些可以包括但不限于,核苷酸(例如,核苷酸三磷酸酯),修饰的核苷酸,寡核苷酸(例如,用于基于聚合酶的扩增的引物寡核苷酸和用于基于连接酶的扩增的寡核苷酸构建块),一种或多种盐(例如,含镁盐),一种或多种缓冲剂,一种或多种聚合剂(例如,连接酶,聚合酶),一种或多种切口酶(例如,切割双链核酸的一条链的酶)和一种或多种核酸酶(例如,外切核酸酶,内切核酸酶,RNA酶)。可以使用适合于扩增的任何聚合酶,例如具有或不具有核酸外切酶活性的聚合酶,DNA聚合酶和RNA聚合酶,例如这些酶的突变体形式。可以使用适于将一个寡核苷酸的5'连接到另一寡核苷酸的3'端的任何连接酶。扩增条件还可包括某些反应条件,例如,等温或温度循环条件。在扩增过程中循环温度的方法是已知的,例如,通过使用热循环装置。术语“循环”是指利用单个引物或多个引物进行扩增(例如,扩增反应或延伸反应),其中使用温度循环。在一些实施方式中,扩增条件也可以包括乳剂(例如,油),其可以用于形成其中可以扩增单个核酸分子物质的多个反应室。扩增有时是指指数产物生成过程,并且有时是线性产物生成过程。

[0094] 可以扩增单链核酸靶标的链,并且可以扩增双链核酸靶标的一条或两条链。在一些实施方式中,扩增产物(扩增子)长度为约10个核苷酸至约10,000个核苷酸,长度为约10至约1000个核苷酸,长度为约10至约500个核苷酸,长度为10至约100个核苷酸,并且有时长度为约10、11、12、13、14、15、16、17、18、19、20、25、30、35、40、45、50、55、60、65、70、75、80、80、85、90、95、100、125、150、175、200、225、250、275、300、325、350、375、400、425、450、475、500、600、700、800、900或1000个核苷酸。

[0095] 可以选择任何合适的扩增技术和扩增条件用于特定核酸进行扩增。已知的扩增方法包括但不限于聚合酶链反应(PCR)、延伸和连接、连接扩增(或连接酶链反应(LCR))和基于使用Q-β复制酶或模板依赖性聚合酶的扩增方法(参见美国专利公开号US20050287592)。还可以使用链置换扩增(SDA)、嗜热SDA、基于核酸序列的扩增(3SR或NASBA)和转录相关扩增(TAA)。用于进行扩增方法的试剂、装置和硬件是可商购的,并且扩增条件是已知的,并且

可以选择手头的靶核酸。

[0096] 在某些实施方式中,可以通过使用通用引物实现基于聚合酶的扩增。在这些过程中,与一个或多个通用引物杂交的杂交区域被并入模板核酸中。可以将这样的杂交区域并入(i)与靶核酸杂交并延伸的引物,和/或(ii)连接到例如靶核酸或(i)的产物的寡核苷酸(例如,使用连接酶连接)。涉及通用引物的扩增方法可以提供例如仅使用一个或两个扩增引物扩增多种靶核酸的优点。

[0097] 在某些实施方式中,可以在扩增前或扩增后延伸某些次要核酸物质和主要核酸物质。本文所用的术语“延伸”和其语法变体是指延长核酸的一条链。例如,在某些实施方式中,可延伸与次要或主要核酸物质杂交的寡核苷酸或由次要或主要核酸物质产生的扩增子。延伸反应在延伸条件下进行,并且各种这样的条件是已知的并且针对特定应用选择。扩展条件可以包括某些试剂,包括但不限于一种或多种寡核苷酸,延伸核苷酸(例如,核苷酸三磷酸酯(dNTP)),链终止剂或核苷酸(例如,一种或多种双脱氧核苷酸三磷酸酯(ddNTP)或无环终止剂),一种或多种盐(例如,含镁盐),一种或多种缓冲液(例如,使用 β -NAD,曲通X-100),和一种或多种聚合剂(例如,DNA聚合酶、RNA聚合酶)。对于主要核酸物质特异的链终止剂的浓度C[WT]通常小于对次要核酸物质特异的链终止剂的浓度(C[Mut])的约20%,C[WT]通常为C[Mut]的约0.5%至小于约20%,C[Mut]的约0.5%至小于约15%,C[Mut]的约1%至约15%,C[Mut]的约1%至约10%,C[Mut]的约2%至约10%或C[Mut]的约1%,1.5%,2%,2.5%,3%,3.5%,4%,4.5%,5.5%,6%,6.5%,7%,7.5%,8%,8.5%,9%,9.5%或10%。在某些实施方式中,C[WT]为C[Mut]的约0.1%至约10%,C[Mut]的约0.01%至约10%,或C[Mut]的约0.01%,0.05%,0.1%,0.5%1%,1.5%,2%,2.5%,3%,3.5%,4%,4.5%,5%,5.5%,6%,6.5%,7%,7.5%,8%,8.5%,9%,9.5%或10%。

[0098] 在某些实施方式中,延伸可以在等温条件下或在非等温条件下进行(例如,热循环条件)。一种或多种核酸物质可以在延伸反应中延伸,并且各核酸物质的一个或多个分子可以被延伸。核酸可以延伸一个或多个核苷酸,并且在一些实施方式中,延伸产物长度为约10个核苷酸至约10,000个核苷酸,长度为约10至约1000个核苷酸,长度为约10至约500个核苷酸,长度为约10至约100个核苷酸,并且有时长度为约10、11、12、13、14、15、16、17、18、19、20、25、30、35、40、45、50、55、60、65、70、75、80、80、85、90、95、100、125、150、175、200、225、250、275、300、325、350、375、400、425、450、475、500、600、700、800、900或1000个核苷酸。结合终止核苷酸(例如ddNTP),杂交位置或其他因素可以确定寡核苷酸延伸的长度。在某些实施方式中,扩增和延伸过程在相同的检测过程中进行。

[0099] 在一些实施方式中,延伸反应包括重复多个温度循环以扩增反应中延伸产物的量。在一些实施方式中,延伸反应循环2次或更多次。在一些实施方式中,延伸反应循环10次或更多次。在一些实施方式中,延伸反应循环约10、15、20、50、55、60、65、70、75、80、85、90、95、100、200、300、400、500或600或更多次。在一些实施方式中,延伸反应循环20-50次。在一些实施方式中,延伸反应循环20-100次。在一些实施方式中,延伸反应循环20-300次。在一些实施方式中,延伸反应循环200-300次。在某些实施方式中,延伸反应循环至少50、55、60、65、70、75、80、85、90、95或100次。

[0100] 在一些实施方式中,在延伸组合物存在下延伸次要或主要核酸物质,其中核酸物质延伸一个核苷酸。延伸组合物可以包含一种或多种缓冲液,盐,酶(例如,聚合酶,Klenow

等),水,模板(例如,DNA,RNA,扩增子等),引物(例如,寡核苷酸),核苷酸三磷酸酯,甘油,大分子排阻分子和本领域中使用的任何其它添加剂。延伸组合物可以包括终止核苷酸(例如,双脱氧核苷酸(例如ddNTP))或其他链终止剂,非终止或延伸核苷酸(例如,dNTP)或终止核苷酸和非终止核苷酸的混合物。基本上由一种或多种特定终止核苷酸组成的延伸组合物可以含有延伸组合物的任何其它组分(例如,缓冲液,盐,模板,引物等),但除了上述的那些以外,不含任何其它终止核苷酸或核苷酸三磷酸酯(例如,dNTP)。例如,基本上由ddTTP和ddCTP组成的延伸组合物不含ddATP,ddGTP或任何其他dNTP。在一些实施方式中,延伸组合物中的核苷酸仅为终止核苷酸,并且靶核酸延伸一个核苷酸(即,延伸组合物中有时不存在延伸核苷酸)。在一些实施方式中,延伸组合物基本上由终止核苷酸(例如,ddNTP)组成。在本文提供的方法的实施方式中,对于主要核酸物质特异的链终止剂的浓度C[WT]小于对次要核酸物质特异的链终止剂的浓度(C[Mut])的20%,C[WT]通常为C[Mut]的约0.5%至小于约20%,C[Mut]的约0.5%至小于约15%,C[Mut]的约1%至约15%,C[Mut]的约1%至约10%,C[Mut]的约2%至约10%或C[Mut]的约1%,1.5%,2%,2.5%,3%,3.5%,4%,4.5%,5.5%,6%,6.5%,7%,7.5%,8%,8.5%,9%,9.5%或10%。在某些实施方式中,C[WT]为C[Mut]的约0.1%至约10%,C[Mut]的约0.01%至约10%,或C[Mut]的约0.01%,0.05%,0.1%,0.5%1%,1.5%,2%,2.5%,3%,3.5%,4%,4.5%,5%,5.5%,6%,6.5%,7%,7.5%,8%,8.5%,9%,9.5%或10%。在一些实施方式中,在延伸组合物存在下延伸次要或主要核酸物质,其中核酸物质延伸一个核苷酸。延伸组合物可以包含一种或多种缓冲液,盐,酶(例如,聚合酶,Klenow等),水,模板(例如,DNA,RNA,扩增子等),引物(例如,寡核苷酸),核苷酸三磷酸酯,甘油,大分子排阻分子和本领域中使用的任何其它添加剂。延伸组合物可以包括终止核苷酸(例如,双脱氧核苷酸(例如ddNTP))或其他链终止剂,非终止或延伸核苷酸(例如,dNTP)或终止核苷酸和非终止核苷酸的混合物。基本上由一种或多种特定终止核苷酸组成的延伸组合物可以含有延伸组合物的任何其它组分(例如,缓冲液,盐,模板,引物等),但除了上述的那些以外,不含任何其它终止核苷酸或核苷酸三磷酸酯(例如,dNTP)。例如,基本上由ddTTP和ddCTP组成的延伸组合物不含ddATP,ddGTP或任何其他dNTP。在一些实施方式中,延伸组合物中的核苷酸仅为终止核苷酸,并且靶核酸延伸一个核苷酸(即,延伸组合物中有时不存在延伸核苷酸)。在一些实施方式中,延伸组合物基本上由终止核苷酸(例如,ddNTP)组成。在本文提供的方法的实施方式中,对于主要核酸物质特异的链终止剂的浓度C[WT]小于对次要核酸物质特异的链终止剂的浓度(C[Mut])的20%,C[WT]通常为C[Mut]的约0.5%至小于约20%,C[Mut]的约0.5%至小于约15%,C[Mut]的约1%至约15%,C[Mut]的约1%至约10%,C[Mut]的约2%至约10%或C[Mut]的约1%,1.5%,2%,2.5%,3%,3.5%,4%,4.5%,5.5%,6%,6.5%,7%,7.5%,8%,8.5%,9%,9.5%或10%。在某些实施方式中,C[WT]为C[Mut]的约0.1%至约10%,C[Mut]的约0.01%至约10%,或C[Mut]的约0.01%,0.05%,0.1%,0.5%1%,1.5%,2%,2.5%,3%,3.5%,4%,4.5%,5%,5.5%,6%,6.5%,7%,7.5%,8%,8.5%,9%,9.5%或10%。

[0101] 在一些实施方式中,链终止剂或链终止核苷酸包含一个或多个可检测标记。可检测标记包括但不限于,质量标记、放射性分子、荧光分子、抗体、抗体片段、半抗原、碳水化合物、生物素、生物素衍生物、磷光部分、发光部分、电化学发光部分、在氧化或还原时产生电化学信号的部分,例如铁,钆或钆的复合物(参见例如,由Genmark诊断公司(Genmark

Diagnostics, Inc) 使用的eSensor技术, 例如Pierce等, J.Clin.Micribiol., 50 (11) :3458-3465 (2012) 中所述)、色度部分和具有可检测的电子自旋共振、电容、介电常数或电导率的部分、或其标记的任何组合。在一些实施方式中, 链终止剂或链终止核苷酸包含一个可检测标记。在一些实施方式中, 第一链终止剂或链终止核苷酸包括不同于第二链终止剂或链终止核苷酸的可检测标记的可检测标记。在一些实施方式中, 延伸组合物包括一种或多种链终止剂或链终止核苷酸, 其中每种链终止剂或链终止核苷酸包括不同的可检测标记。在一些实施方式中, 延伸组合物包含一种或多种链终止剂或链终止核苷酸, 其中每种含有相同的检测标记。在一些实施方式中, 延伸组合物包括链终止剂或链终止核苷酸和延伸核苷酸(例如dNTP) 并且核苷酸中的一种或多种(例如, 终止核苷酸和/或延伸核苷酸) 包括检测标记。

[0102] 可以选择和利用任何合适的延伸反应。可以采用延伸反应, 例如, 通过将脱氧核苷酸和/或双脱氧核苷酸掺入到与靶核酸中与SNP位点相邻的区域杂交的延伸寡核苷酸来区分SNP等位基因。引物通常用聚合酶延伸。在一些实施方式中, 寡核苷酸仅延伸与SNP位点互补的一个脱氧核苷酸或双脱氧核苷酸。在一些实施方式中, 寡核苷酸可以通过掺入dNTP延伸, 并通过ddNTP终止, 或者在某些实施方式中通过ddNTP掺入而终止但没有dNTP延伸。在一些实施方式中, 可以使用以下来进行延伸: 未修饰的延伸寡核苷酸和未修饰的双脱氧核苷酸, 未修饰的延伸寡核苷酸和生物素化的双脱氧核苷酸, 含有脱氧肌苷的延伸寡核苷酸和未修饰的双脱氧核苷酸, 含有脱氧肌苷的延伸寡核苷酸和生物素化的双脱氧核苷酸, 通过生物素化的双脱氧核苷酸延伸或通过生物素化的脱氧核苷酸和/或未修饰的双脱氧核苷酸延伸。

[0103] 在一些实施方式中, 寡核苷酸物质可以在杂交条件下与邻近遗传变异或变体的模板(例如, 次要或主要核酸物质) 杂交(例如, 寡核苷酸物质的3'末端可位于遗传变异位点的5'末端, 并且可距离遗传变异位点的5'末端0至10个核苷酸)。在靶核酸的遗传变异位点上可能存在几种变体。遗传变异有时是单核苷酸多态性(SNP) 或单核苷酸变体。几个单核苷酸变体可以存在于位于杂交的寡核苷酸3'的模板靶标上的单个碱基位置。几个单核苷酸变体可以因模板靶上位于杂交寡核苷酸物质的3'的位置处的单碱基而不同。在一些实施方式中, 寡核苷酸物质在变体位置延伸一个核苷酸。在一些实施方式中, 根据存在的变体的数量, 寡核苷酸可以由五种终止核苷酸(例如ddATP, ddUTP, ddTTP, ddGTP, ddCTP) 中的任何一种延伸。主要核酸物质及其次要核酸物质变体或相应的扩增子可以作为模板, 并且部分地可以确定延伸反应中哪个终止核苷酸被加入到寡核苷酸中。主要核酸物质可以具有两种或更多种次要核酸物质变体。在一些实施方式中, 主要核酸物质具有两种、三种或四种次要核酸物质变体。

[0104] 在本文提供的方法的多重化形式中, 可以在单个反应容器中分析主要核酸物质及其一种或多种次要核酸物质变体, 并且对于主要核酸物质特异的链终止剂不同于对一种或多种次要核酸物质特异的链终止剂。在某些实施方式中, 次要核酸物质全部由与主要核酸物质特异性链终止剂不同的相同链终止剂终止。在一些实施方式中, 至少一种次要核酸物质具有不同于其它次要核酸物质的链终止剂的链终止剂。在某些实施方式中, 反应中每种次要核酸物质的链终止剂彼此不同, 也不同于对于主要核酸物质特异的链终止剂。

[0105] 在本文提供的方法的一些多重化形式中, 可以在一系列反应容器中分析超过一种

主要核酸物质和超过一种次要核酸物质的混合物。具有相同特异链终止剂的所有主要核酸物质在相同的反应容器中进行延伸(以产生,例如,使用链终止剂的单碱基延伸产物),并且在实施方式中,容器中唯一的链终止剂是对容器中主要核酸物质有特异(并且共有)的链终止剂。类似地,具有相同特异性链终止剂的所有次要核酸物质在相同的反应容器中被延伸,并且在实施方式中,容器中唯一的链终止剂是对容器中的次要核酸物质有特异性(并且共有)的链终止剂。

[0106] 在本文提供的方法,包括方法的多重化形式的实施方式中,对于主要核酸物质特异的链终止剂的浓度(C[WT])可小于对次要核酸物质特异的链终止剂的浓度(C[Mut])的20%,C[WT]通常为C[Mut]的约0.5%至小于约20%,C[Mut]的约0.5%至小于约15%,C[Mut]的约1%至约15%,C[Mut]的约1%至约10%,C[Mut]的约2%至约10%或C[Mut]的约1%,1.5%,2%,2.5%,3%,3.5%,4%,4.5%,5.5%,6%,6.5%,7%,7.5%,8%,8.5%,9%,9.5%或10%。在某些实施方式中,C[WT]为C[Mut]的约0.1%至约10%,C[Mut]的约0.01%至约10%,或C[Mut]的约0.01%,0.05%,0.1%,0.5%,1%,1.5%,2%,2.5%,3%,3.5%,4%,4.5%,5%,5.5%,6%,6.5%,7%,7.5%,8%,8.5%,9%,9.5%或10%。

[0107] 可以通过可存在于一个或多个延伸引物(UEP)、链终止剂以及一个或多个核苷酸(如果存在于延伸反应中)上的检测标记来鉴定得到的延伸产物。标记可包括但不限于,质量标记、放射性分子、荧光分子、抗体、抗体片段、半抗原、碳水化合物、生物素、生物素衍生物、磷光部分、发光部分、电化学发光部分、在氧化或还原时产生电化学信号的部分,例如铁,钆或钆的复合物(参见例如,由Genmark诊断公司(Genmark Diagnostics, Inc)使用的eSensor技术,例如Pierce等, *J. Clin. Microbiol.*, 50(11):3458-3465(2012)中所述)、色度部分和具有可检测的电子自旋共振、电容、介电常数或电导率的部分、或其标记的任何组合。

[0108] 对应于次要和主要核酸物质的标记的延伸产物可以通过多种方法分析,包括但不限于,质谱, MALDI-TOF质谱, 荧光检测, DNA测序凝胶, 自动DNA测序机上的毛细管电泳, 微通道电泳和其他测序方法, 质谱, 飞行时间质谱, 四极杆质谱, 扇形磁场质谱, 电子质谱红外光谱, 紫外光谱, potentiostatic电流分析, 电流/电化学信号测量或通过DNA杂交技术, 包括Southern印迹, Slot印迹, Dot印迹和DNA微阵列, 其中DNA片段可用作“探针”和“靶标”, ELISA, 荧光试验, 荧光共振能量转移(FRET), SNP-IT, GeneChips, HuSNP, BeadArray, TaqMan试验, Invader试验, MassExtend®, 或 MassCleave® 方法。

[0109] 在一些实施方式中,次要核酸物质相对于主要核酸物质的相对量(例如,频率或拷贝数)可以通过其检测信号相对于主要核酸物质特异性链终止剂的浓度与次要核酸物质特异性链终止剂的浓度之比的比例,使用归一化系数来确定。在该方法的实施方式中,延伸期间的主要和次要核酸物质的扩增是线性的并且与试验中次要核酸物质,主要核酸物质的分子的相对量及其链终止剂的相对浓度(比例)成比例。在一些实施方式中,通过使用系数将次要核酸物质的信号与主要核酸物质的信号的比例归一化来对次要核酸物质的量(例如,拷贝数、浓度、百分比)进行定量。该系数与对于主要核酸物质特异的链终止剂或核苷酸的浓度相比对于次要核酸物质特异的链终止剂或核苷酸的浓度的分数成反比(即,主要核酸物质特异性链终止剂相对于次要核酸物质特异性链终止剂的分数越低,系数越大)。在本文提供的方法的实施方式中,以足以产生可检测信号的量加入对于主要核酸物质特异的链终

止剂,其可用作鉴定是否存在次要核酸物质的阳性对照并且还可以作为量化样品中次要核酸物质的相对量(例如,频率或拷贝数)的基础。

[0110] 在一些实施方式中,一种次要核酸物质可以比其它次要核酸物质更大的丰度存在。在一些实施方式中,可以调节对各种次要核酸物质中的每一种特异的链终止剂的相对浓度,以优化对应于每种次要核酸物质的检测信号的量级。在一些实施方式中,主要核酸物质包括第一,第二和第三变体次要核酸物质,其中第二变体以比第一和第三变体更大的丰度存在。在一些实施方式中,主要核酸物质包括第一,第二、第三和第四变体次要核酸物质,其中第二变体以比第一、第三和第四变体更大的丰度存在。以更大的丰度表示的变体通常以更高的浓度存在,或者当与另一种变体相比时由更多数量的分子(例如,拷贝)表示。更高的浓度可以是2倍以上。在一些实施方式中,更高的浓度为10倍或更多。在一些实施方式中,更高的浓度为100倍、1000倍或10000倍或更多。在一些实施方式中,主要核酸物质以比次要核酸物质高100倍或更高的浓度存在。在一些实施方式中,次要核酸物质代表少于主要核酸物质的30%,20%,15%,10%,8%,5%,4%,3%,2%,1%,0.8%,0.75%,0.5%,0.1%,0.05%,0.01%或更少。在一些实施方式中,次要核酸物质占主要核酸物质的约5%至约0.75%。在一些实施方式中,次要核酸物质代表少于组合中总核酸物质的30%,20%,15%,10%,8%,5%,4%,3%,2%,1%,0.8%,0.75%,0.5%,0.1%,0.05%,0.01%或更少。

[0111] 在一些实施方式中,在延伸组合物中存在(或在一些实施方式中不存在)的终止核苷酸确定将哪个终止核苷酸加入到寡核苷酸中。在一些实施方式中,延伸组合物包含一个或多个终止核苷酸(例如,ddNTP)。在一些实施方式中,延伸组合物包含一个或多个终止核苷酸和一个或多个非终止核苷酸(例如,dNTP)。在一些实施方式中,延伸组合物仅包含对应于特定次要核酸物质或主要核酸物质的终止核苷酸,因此仅允许延伸该物质。在一些实施方式中,在含有多种次要核酸物质的样品中,延伸组合物仅包括对应于希望检测的次要核酸物质(变体)的那些终止核苷酸。

[0112] 本文所用的术语“信噪比”是指通过在使用检测过程(例如,质谱)时量化信号相对于噪声的强度的比例来定量测量信号质量。在一些实施方式中,一个光谱上的强度峰比由另一光谱上相同分析物(例如,延伸的寡核苷酸物质)产生的低强度峰具有更大的信噪比。在一些实施方式中,由主要核酸物质(例如,野生型等位基因、第二变体、野生型变体)衍生的延伸的寡核苷酸物质产生噪音。在一些实施方式中,由次要核酸物质(例如,突变变体,突变等位基因,SNP)衍生的延伸的寡核苷酸物质产生的信号被更丰富的次要或主要延伸的寡核苷酸物质产生的噪音所遮蔽。本文中的“信噪比”中使用的术语“信号”是指延伸的寡核苷酸物质的信号峰的强度。在一些实施方式中,术语“信噪比”中使用的术语“信号”通常是指来自次要核酸物质的延伸的寡核苷酸物质的信号峰的强度(例如,突变变体,突变等位基因,SNP)。在一些实施方式中,以允许增强由次要核酸物质的延伸产生的信号的浓度加入允许延伸主要核酸物质(例如,野生型等位基因)的终止核苷酸,同时仍然提供主要核酸物质信号用作对照,并作为量化次要核酸物质的基础。

[0113] 本文所用的术语“灵敏度”是指当使用检测过程(例如,质谱)时可以给定的信噪比检测的分析物的量。在一些实施例中,可以通过降低背景或噪音水平来提高灵敏度。在一些实施方式中,由主要核酸物质(例如,野生型等位基因、野生型变体)衍生的延伸的寡核苷酸

物质产生噪音。在一些实施方式中,当从来自主要核酸物质的延伸的寡核苷酸物质产生的信号减少时,灵敏度增加。

[0114] 任何合适类型的核苷酸都可以并入扩增产物或延伸产物中。在一些实施方式中,核苷酸可以是天然存在的核苷酸,终止核苷酸或非天然存在的核苷酸(例如,核苷酸类似物或衍生物)。在一些实施方式中,某些核苷酸可以包含可检测的标记和/或用于FRET检测的结合对或荧光标记对的成员(例如,该对中的一个成员可以在通过延伸并入UEP的终止核苷酸上,而该对的另一个成员可以在延伸产物寡核苷酸上的其他位置)。

[0115] 可以对包含通过扩增方法产生的扩增子的溶液或含有由延伸方法产生的延伸产物的溶液进行进一步处理。例如,可将溶液与从未掺入扩增子或延伸产物的游离核苷酸中除去磷酸部分的试剂接触。这种试剂的一个示例是磷酸酶(例如,碱性磷酸酶,例如虾碱性磷酸酶)。扩增子和延伸产物还可以与固相结合,可以被洗涤,可以与除去末端磷酸(例如,暴露于磷酸酶)的试剂接触,可以与除去末端核苷酸的试剂(例如,外切核酸酶)接触,可以与切割的试剂(例如,内切核酸酶,核糖核酸酶)接触等。

[0116] 本文所用术语“寡核苷酸”是指通过共价键连接的两个或多个核苷酸或核苷酸类似物。寡核苷酸具有任何方便的长度,并且在一些实施方式中,长度为约5至约200个核苷酸,长度为约5至约150个核苷酸,长度为约5至约100个核苷酸,长度为约5至约75个核苷酸或长度为约5至约50个核苷酸,并且有时长度为约5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、25、30、35、40、45、50、55、60、65、70、75、80、80、85、90、95、100、125、150、175、或200个核苷酸。寡核苷酸可包括脱氧核糖核酸(DNA),核糖核酸(RNA),天然存在和/或非天然存在的核苷酸或其组合,及其任何化学或酶修饰(例如,甲基化DNA,修饰的核苷酸的DNA)。寡核苷酸的长度有时短于扩增子或靶核酸的长度,但不一定比用于扩增的引物或多核苷酸短。寡核苷酸通常包含与扩增子、靶核酸或其互补序列互补或基本互补的核苷酸子序列或杂交序列(例如,比对时与扩增子或靶核酸互补序列约95%、96%、97%、98%、99%或超过99%相同)。寡核苷酸可包含与扩增子、靶核酸或其互补序列不互补或基本上不互补的核苷酸子序列(例如,在与扩增子互补或基本互补的引物中核苷酸子序列的3'或5'末端处)。某些实施方式中的寡核苷酸可含有可检测分子(例如,标签,荧光团,放射性同位素,比色剂,颗粒,酶等)和/或(在某些实施方式中)结合对成员(例如,生物素/亲和素,生物素/链霉亲和素)。

[0117] 本文所用的术语“溶液中”是指液体,例如含有一种或多种核酸的液体。溶液中的核酸和其它组分可以分散在整个溶液中,并且溶液通常包含水(例如,水溶液)。溶液可包含任何方便数量的寡核苷酸物质,并且通常至少存在与待检测的扩增子物质或靶核酸物质相同数量的寡核苷酸物质。

[0118] 本文所用的术语“杂交序列”是指能够与扩增子、靶核酸或其互补序列特异性杂交的寡核苷酸中的核苷酸序列。杂交序列易于设计和选择,并且其长度可适于与如本文所述的溶液中的扩增子、靶序列或其互补序列杂交。在一些实施方式中,各寡核苷酸中的杂交序列的长度为约5至约200个核苷酸(例如,长度为约5至10,约10至15,约15至20,约20至25,约25至30,约30至35,约35至40,约40至45,或约45至50,约50至70,约80至90,约90至110,约100至120,约110至130,约120至140,约130至150,约140至160,约150至170,约160至180,约170至190,约180至200个核苷酸)。

[0119] 本文所用的术语“杂交条件”是指具有互补核苷酸序列的两个核酸可以彼此相互作用的条件。杂交条件可以是高严格性,中等严格性或低严格性,并且这些不同程度的严格性的条件是已知的。通常根据感兴趣的应用选择允许扩增和/或延伸的杂交条件。

[0120] 本文使用的术语“与一个扩增子或靶核酸特异性杂交”是指基本上与一个扩增子物质或靶核酸物质杂交,并且基本上不与溶液中的其它扩增子物质或靶核酸物质杂交。特异性杂交排除错配,使得例如可以设计寡核苷酸以与特定等位基因特异性杂交,并且仅与该等位基因杂交。与等位基因均匀匹配或互补的寡核苷酸将与该等位基因特异性杂交,而如果存在一个或多个碱基错配,则不会发生杂交。

[0121] 本文所用的术语“杂交位置”是指另一核酸杂交的扩增子或靶核酸上的特定位置。在某些实施方式中,寡核苷酸的末端相邻于或基本上相邻于扩增子物质或靶核酸物质上具有与另一个扩增子物质或靶核酸物质不同序列的位点。当位点和寡核苷酸末端之间没有核苷酸时,寡核苷酸末端与该位点“相邻”。在某些实施方式中,当位点和寡核苷酸末端之间存在1、2、3、4、5、6、7、8、9或10个核苷酸时,寡核苷酸末端与位点“基本相邻”。

[0122] 可区分标记和释放

[0123] 如本文所用,术语“可区分标记”和“可区分标签”是指可以彼此区分并用于鉴定标签所附着的核酸的标记或标签的类型。本文的多重化方法可选择并使用多种类型的标记和标签。例如,寡核苷酸、氨基酸、小有机分子、发光分子、光吸收分子、光散射分子、发光分子、同位素、酶等可用作可区分标记或标签。在某些实施方式中,具有不同长度,不同质荷比,不同电泳迁移率(例如,毛细管电泳迁移率)和/或不同质量的寡核苷酸、氨基酸、和/或小分子有机分子也可用作可区分标记或标签。因此,荧光团、放射性同位素、比色剂、发光剂、化学发光剂、光散射剂等可用作标记。标记的选择取决于所需灵敏度、与核酸偶联的难易程度、稳定性要求和可用仪器。本文中关于可区分标记和标签使用的术语“可区分特征”是指可以与另一种标记或标签区分的一种标记或标签的任何特征(例如,本文所述的质量和其他)。在一些实施方式中,可以选择和/或设计可区分标记和标签的标记组成以在质谱仪中产生最佳的飞行行为(flight behavior),并允许以高多重化水平区分标记和标签。

[0124] 对于本文使用的方法,特定的靶(主要或次要)核酸物质,扩增子物质和/或延伸的寡核苷酸物质通常与可区分的可检测标记物质配对,使得特定标记或标签物质的检测直接鉴定特定组合物中的特定靶标次要或核酸物质、扩增子物质和/或延伸的寡核苷酸物质的存在和/或对其定量。因此,可以使用标记物质的一个可区分的特征,例如,以鉴定组合物中的一个靶核酸物质,因为该特定可区分的特征对应于特定的靶核酸。标记和标签可以通过任何已知方法并在任何位置(例如,寡核苷酸的5'处)连接到核酸(例如,寡核苷酸)。因此,如本文所用,对于每个特定的标记物质,对于每个特定的靶核酸物质来说,“特异性对应”是指与一种靶物质配对的一种标记物质。当检测到标记物质的存在时,检测与该标记物质相结合的靶核酸物质的存在和/或在某些实施方式中对其定量。

[0125] 本文中关于可区分标签或标记(统称为“标记”)使用的术语“物质(species)”是指与另一种标记可检测地区分的一种标记。在某些实施方式中,标记物质的数量包括但不限于约2至约10000种标记物质,约2至约500,000种标记物质,约2至约100,000,约2至约50000,约2至约10000和约2至约500种标记物质,或有时约2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、25、30、35、40、45、50、55、60、65、70、75、80、80、85、90、95、100、125、

150、175、200、225、250、275、300、325、350、375、400、425、450、475、500、600、700、800、900、1000、2000、3000、4000、5000、6000、7000、8000、9000、10000、20000、30000、40000、50000、60000、70000、80000、90000、100000、200000、300000、400000或500000种标记物质。

[0126] 本文所用的术语“质量可区分标记”是指以质量作为特征区分的标记。可以选择和使用各种质量可区分的标记,例如,复合体、氨基酸和/或多联体。核苷酸串(例如,核酸,复合体),氨基酸串(例如,肽,多肽,复合体)和/或多联体的不同长度和/或组成可以通过质量来区分并用作标记。任何数量的单元都可以在质量可区分的标记中使用,并且这些单元的上限和下限部分取决于用于检测和区分这些标签的系统的分辨率和质量窗口。因此,可以部分地基于用于检测和区分标记的检测器的分辨率和质量窗口来选择质量可区分标记的长度和组成。

[0127] 本文所用的术语“复合体”是指一组单体单元的组成,并且不是单体单元的特定顺序。对于核酸,术语“复合体”是指以碱基作为单体单元的核酸的碱基组合物。每种类型的碱基的数量可以由 B_n 表示(即, $A_a C_c G_g T_t$,其中 $A_0 C_0 G_0 T_0$ 表示“空”复合体或不含碱基的复合体)。天然复合体是所有组分单体单元(例如,对于核酸是碱基并且对于多肽是氨基酸)大于或等于零的复合体。在某些实施方式中, A, C, G 或 T 中的至少一个等于1或更大(例如, $A_0 C_0 G_1 T_0$, $A_1 C_0 G_1 T_0$, $A_2 C_1 G_1 T_2$, $A_3 C_2 G_1 T_5$)。为了比较序列以确定序列变异,在本文提供的方法中,可以通过用于处理数据的算法产生含有负数的单体单元的“非天然”复合体。对于多肽,复合体是指多肽片段的氨基酸组成,其中类似地表示每种类型的氨基酸的数量。复合体物质可对应于多个序列。例如,聚合物 $A_2 G_3$ 对应于序列AGGAG, GGGAA, AAGGG, GGAGA等。通常,存在对应于序列的唯一复合体,但是超过一个序列可以对应于相同复合体。在某些实施方式中,一个复合体物质与(例如,对应于)一个靶核酸物质、扩增子物质和/或寡核苷酸物质配对。在本文的实施方式中(例如, $A_0 C_0 G_5 T_0$ 和 $A_0 C_5 G_0 T_0$ 是不同的和质量可区分的复合体物质),不同的复合体物质具有不同的碱基组成,和可区分的质量。在一些实施方式中,一组复合体物质的碱基组成不同并具有相同的长度。在某些实施方式中,一组复合体物质的碱基组成和长度不同。

[0128] 用作质量可区分标记的核苷酸复合体可以具有任何长度,其中所有复合体物质可以可检测地区分,例如长度为约1至15、5至20、1至30、5至35、10至30、15至30、20至35、25至35、30至40、35至45、40至50、或25至50,或有时约55、60、65、70、75、80、85、90、85或100个核苷酸。用作质量可区分标记的肽或多肽复合体可以具有任何长度,其中所有复合体物质可以可检测地区分,例如长度为约1至20、10至30、20至40、30至50、40至60、50至70、60至80、70至90个或80至100个氨基酸。如上所述,复合体中单元数量的限制通常受到用于区分多联体物质的检测方法的分辨率和质量窗口的限制。

[0129] 术语“多联体”和“串联体”在本文中同义使用(统称为“多联体”),并且是指包含彼此连接的两个或更多个单元的分子(例如,通常连续连接;在某些实施方式中有时分支)。在一些实施方式中,多联体有时是核酸和/或人造聚合物。在一些实施方式中,多联体可以包含相同类型的单元(例如,同多联体),并且有时多联体可以包含不同类型的单元(例如,异多联体)。多联体可包含任何类型的单元,包括核苷酸单元、氨基酸单元、小有机分子单元(例如,三苯甲基)、特定核苷酸序列单元、特定氨基酸序列单元等。在一个实施方式中,三个特定序列单元ABC的同多联体是ABCABCABC。只要可以将各多联体物质与其他物质区分开,多联体可以包含任何数量的单元。例如,在一些实施方式中,三苯甲基多联体物质可以包含

约1、2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、25、30、35、40、45、50、55、60、65、70、75、80、85、90、95、100、125、150、175、200、225、250、275、300、325、350、375、400、425、450、475、500、600、700、800、900或1000个三苯甲基单元。

[0130] 在某些实施方式中,可以从核酸产物(例如,延伸寡核苷酸)释放可区分的标记。可区分标记和核酸之间的连接可以是可被转录并切割,切割并允许检测释放的一种或多种标记,从而鉴定和/或定量核酸产物的任何类型(例如,Ehrich等的题为“靶标特异性复合体及使用方法”的美国专利申请公开号为US20050287533A1)。这种连接和切割连接的方法(“切割条件”)是已知的。在某些实施方式中,标记可以与其所连接的分子的其它部分分离。在一些实施方式中,标记(例如,聚合物)从更大的核苷酸串(例如,延伸的寡核苷酸)切割。连接的非限制性示例包括可由核酸酶(例如,核糖核酸酶,内切核酸酶)切割的连接;可由化学物质切割的连接;可通过物理处理切割的连接;和可被光切割的光可切割接头(例如,邻硝基苄基,6-硝基维酰氨基氧羰基,2-硝基苄基)。当使用发射光的检测系统(例如,基质辅助激光解吸电离(MALDI)质谱涉及激光发射光)时,光可切割接头提供了优点,因为切割和检测在单个步骤中结合并进行。

[0131] 在某些实施方式中,标记可以是较大单元的一部分,并且可以在检测之前与该单元分离。例如,在某些实施方式中,标记是较大核苷酸序列中的一组连续核苷酸,并且从较大的核苷酸序列切下标记。在这样的实施方式中,标记通常位于其所在的核苷酸序列或核酸的一个末端处。在一些实施方式中,标记或其前体位于转录盒中,转录盒包括与编码标记的前体序列可操作地连接的启动子序列。在后一个实施方式中,启动子有时是产生包含标记或由标记组成的RNA的RNA聚合酶募集启动子。包含标记的RNA可以在检测之前被切割以释放标记(例如,用RNA酶)。

[0132] 在某些实施方式中,可区分的标记或标签不从延伸的寡核苷酸切下,并且在一些实施方式中,可区分的标记或标签包含捕获剂。在某些实施方式中,检测可区分的特征包括检测延伸的寡核苷酸的存在或不存在,并且在一些实施方式中,延伸的寡核苷酸包含捕获剂。

[0133] 检测 and 多重化程度

[0134] 本文所用的术语“检测”标签是指鉴定标记物质。任何合适的检测装置可用于区分样品中的标记物质。适用于检测大量可区分标记的检测装置包括,但不限于,某些质谱和凝胶电泳装置。质谱形式的示例包括但不限于,基质辅助激光解吸/电离飞行时间(MALDI-TOF)质谱(MS),MALDI正交TOF MS(OTOF MS;二维),激光解吸质谱(LDMS),电喷雾(ES)MS,离子回旋共振(ICR)MS和傅里叶变换MS。本文所述的方法可以容易地应用于其中分析物挥发和离子化的质谱形式(“电离MS”,例如,MALDI-TOF MS,LDMS,ESMS,线性TOF,OTOF)。正交离子提取MALDI-TOF和轴向MALDI-TOF可以产生相对高的分辨率,从而相对高的多重化水平。适用于检测发光,光吸收和/或光散射标记的检测装置包括但不限于,某些光检测器和光学检测器(例如,用于荧光,化学发光,吸收和/或光散射标签)。

[0135] 本文提供的方法允许高通量检测或发现存在于含有一种或多种主要核酸物质的样品或样品组合中的多种次要核酸物质。相较各单独的次要核酸物质必须在分开质谱或其他分析中进行,多重化提供的优势在于能在少至单个质谱或其他检测系统(如荧光或电信号)中鉴定多种次要核酸物质(例如,一些具有不同的序列变化)。在一些实施方式中,本文

提供的方法适用于高通量、高度自动化的用于以高速度和精确度分析序列变化的方法。在一些实施方式中,本文的方法可在单一反应中以高水平多重化。当多态性基因座处的基因型未知,并且在一些实施方式中,基因座上的基因型已知时,可以应用多重化。

[0136] 在某些实施方式中,多重化的靶(次要和/或主要)核酸物质的数量包括但不限于约2-1000,并且有时约1-3、3-5、5-7、7-9、9-11、11-13、13-15、15-17、17-19、19-21、21-23、23-25、25-27、27-29、29-31、31-33、33-35、35-37、37-39、39-41、41-43、43-45、45-47、47-49、49-51、51-53、53-55、55-57、57-59、59-61、61-63、63-65、65-67、67-69、69-71、71-73、73-75、75-77、77-79、79-81、81-83、83-85、85-87、87-89、89-91、91-93、93-95、95-97、97-101、101-103、103-105、105-107、107-109、109-111、111-113、113-115、115-117、117-119、121-123、123-125、125-127、127-129、129-131、131-133、133-135、135-137、137-139、139-141、141-143、143-145、145-147、147-149、149-151、151-153、153-155、155-157、157-159、159-161、161-163、163-165、165-167、167-169、169-171、171-173、173-175、175-177、177-179、179-181、181-183、183-185、185-187、187-189、189-191、191-193、193-195、195-197、197-199、199-201、201-203、203-205、205-207、207-209、209-211、211-213、213-215、215-217、217-219、219-221、221-223、223-225、225-227、227-229、229-231、231-233、233-235、235-237、237-239、239-241、241-243、243-245、245-247、247-249、249-251、251-253、253-255、255-257、257-259、259-261、261-263、263-265、265-267、267-269、269-271、271-273、273-275、275-277、277-279、279-281、281-283、283-285、285-287、287-289、289-291、291-293、293-295、295-297、297-299、299-301、301-303、303-305、305-307、307-309、309-311、311-313、313-315、315-317、317-319、319-321、321-323、323-325、325-327、327-329、329-331、331-333、333-335、335-337、337-339、339-341、341-343、343-345、345-347、347-349、349-351、351-353、353-355、355-357、357-359、359-361、361-363、363-365、365-367、367-369、369-371、371-373、373-375、375-377、377-379、379-381、381-383、383-385、385-387、387-389、389-391、391-393、393-395、395-397、397-401、401-403、403-405、405-407、407-409、409-411、411-413、413-415、415-417、417-419、419-421、421-423、423-425、425-427、427-429、429-431、431-433、433-435、435-437、437-439、439-441、441-443、443-445、445-447、447-449、449-451、451-453、453-455、455-457、457-459、459-461、461-463、463-465、465-467、467-469、469-471、471-473、473-475、475-477、477-479、479-481、481-483、483-485、485-487、487-489、489-491、491-493、493-495、495-497、497-501种物质或更多。

[0137] 通过多重化试验实现分辨质谱的设计方法可以包括引物和寡核苷酸设计方法,试剂如链终止剂的相对浓度,检测标记的选择和其他反应设计方法。就多通路试验中的引物与寡核苷酸设计而言,单通路反应采用相同的引物设计总体方针,例如避免假引发和引物二聚体,只是多通路反应涉及更多引物。另外,对于质谱分析,在某一试验质谱中的分析物峰从任意试验的产物中充分解析,所述试验为包括暂停峰(pausing peak)和任意其它副产物峰的多通路试验。此外,分析物峰最佳地落入用户指定的质量窗口内,例如在5,000-8,500Da的范围内。在一些实施方式中,可以相对于给定SNP链的靶序列设计延伸寡核苷酸。在这样的实施方式中,长度通常在可以是例如用户指定的限制之间(例如,17至24个碱基或17-26个碱基),并且通常不包含靶序列中不确定的碱基。杂交强度有时通过计算序列依赖性熔化(或杂交/解离)温度 T_m 来测量。由于其发夹潜力、假引发潜力、引物二聚体潜力、低复

杂度区域和有问题的子序列(如GGGG),特定的引物选择可能不允许或相对于引物的其他选择不利。用于设计延伸寡核苷酸的方法和软件(例如,根据这些标准)是已知的,并且包括,例如SpectroDESIGNER(塞昆纳姆公司(Sequenom))。

[0138] 如本文所使用的,术语“检出率”或“判定率”是指相对于尝试获得的检出的数量获得的检出的数量(例如,确定的基因型或突变体)。换言之,对于12-重反应,如果10种基因型最终通过进行本文提供的方法确定,则10次检出已经获得,检出率为10/12。不同的事件可能导致特定尝试试验的失败,并导致检出率低于100%。偶尔,在用于终止dNTP和ddNTP的混合的情况下,例如,在引入一种非终止核苷酸(即,dNTP)后,通过暂停聚合酶可以发生不适合的延伸产物。多态性位点处的这种错误终止和正确终止的引物扩展反应之间的质量差异有时太小而不能一致地解析,如果使用不适当的终止混合物,则可能导致错判。在正确终止和错误终止之间的质量差异(即由暂停引起的)以及正确的终止和盐加合物以及正确的终止和非特异性并入之间的质量差异通常被最大化以减少错判的数量。

[0139] 可以通过在一个或多个试验中评估获得的检出的数量(例如,正确或准确评估)和/或假阳性和/或假阴性事件的数量来确定多重化试验准确度。也可以通过比较每个多重化试验中评估的目标的相应单重试验的准确性评估准确性。在某些实施方式中,可以使用一种或多种方法来确定判定率。例如,可以结合用于进行检出的自动化或计算机方法来使用手动方法,并且在一些实施方式中,可以将每种方法的检出率相加以计算总体检出率。在某些实施方式中,当多重化两种或更多种靶核酸(次要和/或主要核酸物质)(例如50或更多种靶核酸)时,检出率的准确度可以为约99%或更大,98%,97%,96%,95%,94%,93%,92%,91%,90%,89%,87-88%,85-86%,83-84%,81-82%,80%,78%-79%或76-77%。在一些实施方式中,包括约2至200种靶物质的多重化试验中每种靶物质的检出率大于或等于80%或更多(例如,81%,82%,83%,84%,85%,86%,87%,88%,89%,90%,91%,92%,93%,94%,95%,96%,97%,98%,99%以上)。

[0140] 在某些实施方式中,错误率可以基于检出率或准确率来确定。例如,错误率可能是错误的检出数。在一些实施方式中,例如,错误率可以是检出率或精确率的不到100%。错误率也可以称为“失败率”。识别假阳性和/或假阴性可以再调整检出率和错误率。在某些实施方式中,运行更多试验还可以帮助鉴定假阳性和/或假阴性,从而调整检出和/或错误率。在某些实施方式中,当多重化作为主要和/或次要核酸物质(例如,五十种或更多种靶核酸物质)的两个或更多种靶核酸时,错误率可以为约1%或更少,2%,3%,4%,5%,6%,7%,8%,9%,10%,11%,12%,13%,14%,15%,16%,17%,18%,19%,20%,21%,22%,23%,24%或25%。

[0141] 应用

[0142] 以下是本文描述的复用技术的非限制性应用的示例。

[0143] 1. 检测序列变异(例如,遗传变体)

[0144] 提供了用于鉴定疾病及其标志物的基因组基础的改进方法。可以通过本文提供的方法鉴定的序列变异(例如,遗传变体)候选物包括含有作为多态性的序列变异的序列。多态性包括天然存在,体细胞序列变异和突变引起的变异。多态性包括但不限于:其中局部区域中的一个或多个核苷酸在个体到个体之间变化的序列微变体,可以从一个核苷酸到百万个碱基大小变化的插入和缺失,以及重复数量变化的微卫星或核苷酸重复。核苷酸重复序

列包括同源重复序列,例如二核苷酸,三核苷酸,四核苷酸或更大的重复序列,其中重复多次的相同序列,以及发现序列基序重复的异源重复序列。对于给定的基因座,核苷酸重复的数量可以根据个体而变化。

[0145] 多态标记或位点是发生差异的基因座。这样的位点可以小至一个碱基对 (SNP)。多态性标记包括但不限于,限制性片段长度多态性 (RFLP),可变数量的串联重复序列 (VNTR),高变区,微卫星,二核苷酸重复序列,三核苷酸重复序列,四核苷酸重复序列和其他重复模式,简单序列重复和插入元件,如Alu。多态形式也表现为基因的不同孟德尔等位基因。可以通过蛋白质,蛋白质修饰,RNA表达修饰,DNA和RNA甲基化,改变基因表达和DNA复制的调节因子,以及基因组核酸或细胞器核酸中任何其它改变表现的差异来观察多态性。

[0146] 此外,许多基因具有多态性区域。由于个体具有多态性区域的几个等位基因变体中的任何一个,因此可以基于基因多态性区域的等位基因变体的类型来鉴定个体。这可以用于,例如,法医目的。在其他情况下,了解个体具有的等位基因变体的种类至关重要。例如,某些基因,例如主要组织相容性复合物 (MHC) 基因中的等位基因差异涉及骨髓运输中的移植物排斥或移植物抗宿主病。因此,非常需要开发用于确定基因或遗传病变多态性区域的等位基因变体的种类的快速、灵敏和准确的方法。本文提供的方法或试剂盒可用于通过测定对象的一个或多个基因或染色体中一个或多个多态性区域的一个或多个等位基因变体的种类来对对象进行基因型分型。使用本文提供的方法对对象进行基因分型可用于法医或种类测试目的,并且多态性区域可以存在于线粒体基因中,或可以是短串联重复。

[0147] 单核苷酸多态性 (SNP) 通常是双倍体系,即个体可以具有任何特定标记的两个等位基因。这意味着与微卫星标记相比,每个SNP标记的信息含量相对较低,其可能具有10个以上的等位基因。单核苷酸多态性也倾向于具有非常的人口特异性;在一个群体中多态性的标记可能在另一个群体中不具有非常的多态性。大约每千碱基发现的SNP (参见Wang等, (1998) Science 280:1077-1082) 提供了产生非常高密度遗传图谱的潜力,这对于开发用于基因或感兴趣区域的单倍型系统是非常有用的,并且由于SNP的性质,它们实际上可能是与正在研究的疾病表型相关的多态性。SNP的低突变率也使得它们成为研究复杂遗传性状的良好标记。

[0148] 基因组学的重点许多是SNP的鉴定,这出于种种原因而很重要。它们允许间接测试 (单倍体的关联) 和直接测试 (功能变体)。它们是最丰富和稳定的遗传标记。常见的疾病最好由常见的遗传改变来解释,并且人群中的天然变异有助于理解疾病、治疗和环境相互作用。

[0149] 体细胞突变的敏感检测对癌症研究界尤其有价值,癌症研究界的兴趣是鉴定肿瘤起始和增殖的遗传决定因素。从敏感方法获得的信息也可用于分析突变以预测患者的结局,并通知相关治疗选项。在一些实施方式中,需要敏感的检测方法,其可以检测到小于或等于其对应野生型序列的5%的遗传变体。在一些实施方式中,实现了可以检测小于或等于野生型的1%的检测方法。在一些实施方式中,可以实施检测小于或等于野生型的5%、4%、3%、2%、1%、0.8%、0.75%、0.5%、0.1%、0.05%或0.01%的检测方法。此外,在产前诊断中,这种类型的方法可以阐明子宫内的父系衍生突变。

[0150] 在一些实施方式中,等位基因分析可以通过从携带一种或多种感兴趣的体细胞突变 (例如,SNP,疾病标记等及其组合) 的核酸靶标产生延伸的寡核苷酸来进行。在一些实施

方式中,检测代表携带体细胞突变的等位基因的释放的延伸的寡核苷酸的存在或不存在可用作筛选靶群体中特定突变的存在或不存在的快速方法。在涉及从突变等位基因产生延伸的寡核苷酸的某些实施方式中,延伸的寡核苷酸可以在合适的突变等位基因产生延伸的寡核苷酸产物时被检测到。

[0151] 2. 鉴定疾病标记物

[0152] 本文提供了用于快速和准确鉴定作为疾病遗传标记的序列变异的方法,其可用于诊断或确定疾病的预后。以遗传标记物表征的疾病可以包括但不限于,动脉粥样硬化,肥胖症,糖尿病,自身免疫性疾病和癌症。所有生物体中的疾病都具有遗传成分,无论是遗传的还是由身体对环境胁迫的反应产生的,如病毒和毒素。正在进行的基因组研究的最终目标是使用这些信息来开发新的方法来鉴定、治疗和潜在地治愈这些疾病。第一步是筛选疾病组织,并确定个体样本水平的基因组变化。这些“疾病”标记的鉴定取决于检测基因组标记物变化以鉴定错误基因或序列变体的能力。基因组标记物(所有遗传基因座,包括单核苷酸多态性(SNP),微卫星和其他非编码基因组区域,串联重复序列,内含子和外显子)可用于鉴定所有生物体,包括人类。这些标记物提供了一种方法,不仅可以鉴定人群,而且可以根据其对疾病、药物治疗、对环境因素的抵抗力等因素的反应来允许群体分层。在一些实施方式中,疾病标志物有时是突变,并且可以是相对罕见的等位基因,例如针对野生型等位基因背景的体细胞突变(例如,癌组织与正常组织,突变型病毒型与正常病毒型(例如,HIV))。在一些实施方式中,罕见的等位基因或突变代表小于野生型的5%、4%、3%、2%、1%、0.8%、0.75%、0.5%、0.1%、0.05%或0.01%。在一些实施方式中,罕见的等位基因或突变可以表示小于野生型的1%。

[0153] 3. 微生物鉴定

[0154] 本文提供用于鉴定微生物和病毒的属、种、株、克隆或亚型的过程或方法。微生物和病毒选自多种生物体,包括但不限于细菌、真菌、原生动物、纤毛虫和病毒。微生物不限于特定属、种、株、亚型或血清型或任何其他分类。可以通过相对于一个或多个参考序列或样品确定靶微生物序列中的序列变异来鉴定微生物和病毒。参考序列可以从例如来自相同或不同属、种、株或血清型或任何其它分类,或来自宿主原核或真核生物体或任何混合群体的其它微生物获得。

[0155] 病原体(例如,细菌或病毒)的鉴定和分型在传染病的临床管理中至关重要。微生物的精确种类不仅用于将疾病状态与健康状态区分开,而且也是确定感染来源及其传播以及抗生素或其他抗微生物治疗是否是最适于治疗的基础。另外可以监控治疗。传统的病原体分型方法使用了各种表型特征,包括生长特征、颜色、细胞或集落形态、抗生素敏感性、染色、气味、血清分型、生化分型和与特异性抗体的反应性,以鉴定微生物(例如,细菌)。所有这些方法都需要疑似病原体的培养,这些病原体有一些严重的缺点,包括高材料和劳动力成本,工作人员接触的危险,由于误操作导致的假阳性和由于活细胞数量较少导致的假阴性或由于对许多病原体的严格培养要求。此外,培养方法需要相当长的时间来实现诊断,并且由于这种感染的潜在的危及生命的性质,通常在获得结果之前开始抗微生物治疗。一些生物体不能保持在培养物中或显示出非常缓慢的生长速率(例如,结核分枝杆菌达到6-8周)。

[0156] 在许多情况下,病原体以少量存在和/或与构成正常菌群的生物体非常相似,并且

可能无法通过上述方法与无害菌株区分。在这些情况下,确定致病菌株的存在可能需要通过本文提供的分子分型方法提供的较高分辨率。

[0157] 4. 检测指示感染的病毒或细菌核酸序列的存在

[0158] 本文提供的方法可用于通过鉴定相对于一个或多个参考序列存在于病毒或细菌核酸序列中的序列变异来确定指示感染的病毒或细菌核酸序列的存在。参考序列可以包括但不限于从感染性生物体,相关非感染性生物体获得的序列或来自宿主生物的序列。

[0159] 病毒、细菌、真菌和其他感染性生物体含有与宿主细胞中所含序列不同的不同核酸序列,包括序列变体。靶DNA序列可以是外来遗传序列的一部分,例如入侵微生物的基因组,包括例如,细菌及其噬菌体、病毒、真菌、原生动物等。本文提供的方法特别适用于区分微生物的不同变体或菌株(例如,致病性,较少致病性,耐药性和非抗性等),以便例如选择适当的治疗干预。感染人类和动物并且可以通过公开的方法检测的致病病毒的示例包括但不限于,逆转录病毒科(例如,人免疫缺陷病毒如HIV-1(也称为HTLV-III, LAV或HTLV-III/LAV; Ratner等, Nature, 313:227-284(1985); Wain Hobson等, Cell, 40:9-17(1985), HIV-2 (Guyader等, Nature, 328:662-669(1987); 欧洲专利公开号0269520; Chakrabarti等, Nature, 328:543-547(1987); 欧洲专利申请0655501)和其它分离物例如HIV-LP(国际公布号W0 94/00562); 小核糖核酸病毒科(例如,脊髓灰质炎病毒,甲型肝炎病毒(Gust等, Intervirology, 20:1-7(1983)); 肠道病毒,人类柯萨奇病毒,鼻病毒,艾柯病毒); 杯状病毒科(例如,导致胃肠炎的毒株); 披膜病毒科(例如,马脑炎病毒,风疹病毒); 黄病毒科(例如,登革热病毒,脑炎病毒,黄热病毒); 冠状病毒科(例如,冠状病毒); 弹状病毒科(例如,水泡性口炎病毒,狂犬病病毒); 丝状病毒科(例如,埃博拉病毒); 副粘病毒科(例如,副流感病毒,腮腺炎病毒,麻疹病毒,呼吸道合胞病毒); 正粘病毒科(例如,流感病毒); 布加病毒科(例如,汉坦病毒,伯加病毒,白蛉病毒和内罗毕病毒); 沙粒病毒科(出血热病毒); 呼肠病毒科(例如,呼肠病毒,轮状病毒和轮状病毒); 双核糖核酸病毒科; 肝炎病毒科(乙肝病毒); 细小病毒科(细小病毒); 细小病毒科(大多数腺病毒); 乳头状瘤病毒(乳头瘤病毒,多瘤病毒); 腺病毒科(大多数腺病毒); 疱疹病毒科(1型单纯疱疹病毒(HSV-1)和HSV-2,水痘带状疱疹病毒,巨细胞病毒,疱疹病毒); 痘病毒科(天花病毒,痘苗病毒,痘病毒); 虹彩病毒科(例如,非洲猪瘟病毒); 和未分类的病毒(例如,海绵状脑病的病原体,δ型肝炎的媒介(被认为是乙型肝炎病毒的有缺陷的卫星),非A型非B型肝炎的媒介(1类=内部传播; 2类=胃肠外传播,即丙型肝炎); 诺沃克和相关病毒,以及星状病毒。

[0160] 感染性细菌的示例包括但不限于幽门螺杆菌(*Helicobacter pylori*),布氏疏螺旋体(*Borrelia burgdorferi*),嗜肺军团菌(*Legionella pneumophila*),分枝杆菌属(例如,结核分枝杆菌(*M. tuberculosis*),鸟分枝杆菌(*M. avium*),胞内分枝杆菌(*M. intracellulare*),堪萨斯州分枝杆菌(*M. kansasii*),戈登分枝杆菌(*M. goodii*)),沙门氏菌,金黄色葡萄球菌(*Staphylococcus aureus*),淋病奈瑟氏球菌(*Neisseria gonorrhoeae*),脑膜炎奈瑟氏球菌(*Neisseria meningitidis*),单核细胞增生利斯特氏菌(*Listeria monocytogenes*),化脓链球菌(*Streptococcus pyogenes*) (A组链球菌),无乳链球菌链球菌(*Streptococcus agalactiae*) (B组链球菌),链球菌属(绿色链球菌),粪链球菌(*Streptococcus faecalis*),牛链球菌(*Streptococcus bovis*),链球菌属(厌氧种),肺炎链球菌(*Streptococcus pneumoniae*),致病性弯曲杆菌属,肠球菌属,流感嗜血杆菌

(*Haemophilus influenzae*), 炭疽芽孢杆菌 (*Bacillus anthracis*), 白喉棒状杆菌 (*Corynebacterium diphtheriae*), 棒状杆菌属, 红斑丹毒丝菌 (*Erysipelothrix rhusiopathiae*), 产气荚膜梭菌 (*Clostridium perfringens*), 破伤风梭菌 (*Clostridium tetani*), 大肠杆菌 (*Escherichia coli*), 产气肠杆菌 (*Enterobacter aerogenes*), 肺炎克雷伯菌 (*Klebsiella pneumoniae*), 多杀巴斯德氏菌 (*Pasturella multocida*), 拟杆菌属, 具核梭杆菌 (*Facterobacterium nucleatum*), 链霉菌链球菌 (*Streptobacillus moniliformis*), 苍白密螺旋体 (*Treponemapallidum*), 极细密螺旋体 (*Treponema pertenuis*), 钩端螺旋体属和放线杆菌 (*Actinomyces israeli*) 以及包括抗生素抗性变体的任何变体。

[0161] 感染性真菌的示例包括但不限于, 新型隐球菌 (*Cryptococcus neoformans*), 荚膜组织胞浆菌 (*Histoplasma capsulatum*), 粗球孢子菌 (*Coccidioides immitis*), 皮炎芽生菌 (*Blastomyces dermatitidis*), 沙眼衣原体 (*Chlamydia trachomatis*), 白色念珠菌 (*Candida albicans*)。其他感染性生物体包括原生生物如恶性疟原虫和弓形虫。

[0162] 5. 抗生素概况

[0163] 本文提供的方法可以提高涉及耐药性, 包括抗生素耐药性的核苷酸变化的检测的速度和准确性。已经鉴定了参与抗异烟肼、利福平、链霉素、氟喹诺酮和乙硫异烟胺的遗传基因座 [Heym等, *Lancet* 344:293 (1994) 和 Morris等, *J. Infect. Dis.* 171:954 (1995)]。常规使用异烟肼 (inh) 和利福平 (rif) 以及吡嗪酰胺和乙胺丁醇或链霉素的组合作为对结核分枝杆菌确诊病例的第一道攻击 [Banerjee等, *Science* 263:227 (1994)]。这种耐药菌株的出现日益增加, 需要开发快速试验来检测它们, 从而降低遵循无效和可能有害的治疗的费用和社区健康危害。鉴定涉及耐药性的一些遗传基因座有助于采用突变检测技术快速筛选导致耐药性的核苷酸变化。此外, 该技术有助于治疗监测和跟踪微生物群体结构以及治疗期间的监督监测。此外, 可以进行混合种群的相关性和监督监测。

[0164] 6. 单元分型

[0165] 本文提供的方法可用于检测单元分型 (haplotyping)。在任何二倍体细胞中, 在任何基因或其他染色体片段上都有两个单元型, 其含有至少一个区别差异。在许多充分研究的遗传系统中, 与单核苷酸变异相比, 单元型与表型更强相关。因此, 确定单元型对于理解各种表型的遗传基础是有价值的, 包括疾病倾向或易感性, 对治疗干预的反应, 以及医学、畜牧业和农业中其他感兴趣的表型。

[0166] 本文提供的单元型方法允许从个体的两个同源染色体之一中选择序列的一部分, 并且在该序列部分上选择基因型链接的SNP。单元型的直接分辨可以产生增加的信息含量, 改善任何相关疾病基因的诊断或鉴定与这些疾病相关的联系。

[0167] 7. 微卫星

[0168] 本文提供的方法允许快速、明确检测微卫星序列变异。微卫星 (有时称为可变数量的串联重复序列或VNTR) 是1至7个或更多个碱基的短串联重复核苷酸单元, 其中最突出的是二、三和四核苷酸重复。微卫星在基因组DNA中每100,000bp存在 (J.L.Weber和P.E.Can, *Am. J. Hum. Genet.* 44,388 (1989); J.Weissenbach等, *Nature* 359,794 (1992))。例如, CA二核苷酸重复约占0.5%的人线粒体外基因组; CT和AG重复一起约占0.2%。CG重复是罕见的, 最可能是由于CpG岛的调节功能。微卫星相对于长度而言是高度多态的, 广泛分布在整个基因

组上,主要富集在非编码序列中,并且它们在基因组内的功能是未知的。微卫星在法医应用中可能是重要的,因为种群将维持该种群特有的各种微型卫星,并且与其他不互配的种群不同。

[0169] 微卫星内的许多变化可能是沉默的,但有些可能导致基因产物或表达水平的显著变化。例如,在一些肿瘤中,基因编码区中发现的三核苷酸重复序列受到影响(CT Caskey等,Science 256,784(1992)),并且微卫星的改变可导致导致癌症倾向的遗传不稳定性(P.J.McKinnen,Hum.Genet.175,1997(1987);J.German等,Clin.Genet.35,57(1989))。

[0170] 8.短串联重复

[0171] 本文提供的方法可以用于相对于例如,在不含STR区域的人类基因组中的参考序列鉴定人类基因组的某些靶序列中的短串联重复(STR)区域。STR区域是与任何疾病或病症无关的多态性区域。人类基因组中的许多基因座包含多态性短串联重复(STR)区域。STR基因座包含长度为3至7个碱基对的短重复序列元件。估计有20万个预期的三聚体和四聚体STR,其在人类基因组中每15kb出现一次频繁存在(参见例如,国际PCT申请号W0 9213969 A1,Edwards等,Nucl.Acids Res.19:4791(1991);Beckmann等,(1992)Genomics 12:627-631)。这些STR基因座中的近一半是多态性的,提供了丰富的遗传标记物来源。在特定基因座处的重复单元数的变化是观察到的可变核苷酸串联重复(VNTR)基因(Nakamura等,(1987)Science 235:1616-1622);和微卫星基因座(Jeffreys等,(1985)Nature 314:67-73),其含有较长重复单元,和微卫星或二核苷酸重复基因座(Luty等,(1991)Nucleic Acids Res.19:4308;Litt等,(1990)Nucleic Acids Res.18:4301;Litt等,(1990)Nucleic Acids Res.18:5921;Luty等,(1990)Am.J.Hum.Genet.46:776-783;Tautz(1989)Nucl.Acids Res.17:6463-6471;Weber等,(1989)Am.J.Hum.Genet.44:388-396;Beckmann等,(1992)Genomics 12:627-631)的序列变异。VNTR分型是微生物分型中非常成熟的工具,例如,结核分枝杆菌(MIRU分型)。

[0172] STR基因座的示例包括但不限于人CD4基因座中的五核苷酸重复序列(Edwards等,Nucl.Acids Res.19:4791(1991));人类芳香酶细胞色素P-450基因(CYP19;Polymeropoulos等,Nucl.Acids Res.19:195(1991))中的四核苷酸重复序列;人类凝血因子XIIIA亚单位基因(F13A1;Polymeropoulos等,Nucl.Acids Res.19:4306(1991))中的四核苷酸重复序列;F13B基因座中的四核苷酸重复序列(Nishimura等,Nucl.Acids Res.20:1167(1992));人类c-les/fps,原癌基因(FES;Polymeropoulos等,Nucl.Acids Res.19:4018(1991))中的四核苷酸重复序列;LFL基因中的四核苷酸重复序列(Zuliani等,Nucl.Acids Res.18:4958(1990));人胰腺磷脂酶A-2基因(PLA2;Polymeropoulos等,Nucl.Acids Res.18:7468(1990))的三核苷酸重复序列变异;VWF基因中的四核苷酸重复序列变异(Ploos等,Nucl.Acids Res.18:4957(1990));和人类甲状腺过氧化物酶(hTPO)基因座中的四核苷酸重复序列(Anker等,Hum.Mol.Genet.1:137(1992))。

[0173] 9.生物体鉴定

[0174] 基因的多态性STR基因座和其他多态性区域是序列变异,是人类鉴定,父系和母系测试,遗传图谱,移植和遗传争议,双胞胎接合测试,人类近亲繁殖检测,人类培养细胞质量控制,人类遗骸的鉴定,以及法医学中精液样本,血迹,微生物和其他物质的检测的极端有用的标记物。这种基因座也是商业动物育种和谱系分析以及商业植物育种中的有用标记

物。通过使用多态性DNA标记物的连锁分析可以鉴定植物作物和动物的经济重要的性状。本文提供了用于确定这种基因座种类的有效和准确方法。

[0175] 10. 检测等位基因变异

[0176] 本文提供的方法允许高通量、快速和准确检测等位基因变体。等位基因变异的研究不仅涉及复杂背景中特异性序列的检测,还涉及具有少量或单个核苷酸差异的序列之间的区别。通过PCR检测等位基因特异性变体的一种方法是基于当模板链与引物的3'末端之间存在错配时,Taq聚合酶难以合成DNA链的事实。可以通过使用仅与一个可能等位基因完全匹配的引物来检测等位基因特异性变体;与其他等位基因的错配起到阻止引物延伸的作用,从而阻止该序列的扩增。本文的方法也适用于关联研究,拷贝数变异,疾病标记物的检测和用于分型的SNP组等。

[0177] 11. 确定等位基因频率

[0178] 本文描述的方法对于鉴定其在群体内频率变化作为年龄、种族、性别或某些其他标准的函数的一种或多种遗传标记物而言是有价值的。例如,ApoE基因型的年龄依赖性分布是本领域已知的(参见,例如,Schechter等,(1994) *Nature Genetics* 6:29-32)。已知在一些程度上与疾病相关的序列变异的频率也可用于检测或监测疾病状态的进展。例如,脂蛋白脂肪酶基因的N291S多态性(N291S)导致丝氨酸取代在氨基酸密码子291处的天冬酰胺,导致与增加男性动脉硬化风险,特别是心肌梗塞相关的高密度脂蛋白胆固醇(HDL-C)水平降低(参见,Reymer等,(1995) *Nature Genetics* 10:28-34)。此外,确定等位基因频率的变化可以允许鉴定先前未知的序列变异并最终鉴定涉及疾病发作和进展的基因或途径。

[0179] 12. 表观遗传学

[0180] 本文提供的方法可用于研究靶核酸或蛋白质相对于参考核酸或蛋白质的变化,该变化不基于序列,例如,碱基或氨基酸的种类,所述碱基或氨基酸是天然存在的单体单元核酸或蛋白质。例如,本文提供的方法可用于识别与序列无关的特征如甲基化模式,修饰的碱基或氨基酸的存在,或靶分子与参考分子之间的较高级结构的差异,以产生片段,其在序列独立位点处被切割。表观遗传学是基于基因表达差异而不是基因序列差异的遗传信息研究。表观遗传改变是指不能通过核酸序列变化解释的基因功能变化或高级核酸结构变化的有丝分裂和/或减数分裂上可遗传的变化。受到表观遗传变异或变化的特征的示例包括,但不限于,动物中的DNA甲基化模式,组蛋白修饰和Polycomb-trithorax组(Pc-G/tx)蛋白复合物(参见,例如,Bird,A., *Genes Dev.*,16:6-21 (2002))。

[0181] 表观遗传变化通常(虽然不一定)会导致基因表达的变化,其通常(虽然不一定)是可遗传的。例如,如下文进一步讨论的,甲基化模式的变化是癌症和其他疾病发展和进展中的早期事件。在许多癌症中,由于异常甲基化,某些基因被不适当地关闭或打开。甲基化模式抑制或激活转录的能力可以被遗传。Pc-G/trx蛋白复合物,如甲基化,可以以可遗传的方式抑制转录。Pc-G/trx多蛋白组装靶向基因组的特定区域,其有效地冻结基因的胚胎基因表达状态,无论基因是活性还是非活性基因,并通过发育稳定地传播该状态。Pc-G/trx组蛋白质靶向和结合基因组的能力仅影响基因组中所含基因的表达水平,而不影响基因产物的性质。本文提供的方法可以与特异性切割试剂或特异性延伸反应一起使用,所述特异性延伸反应鉴定靶序列相对于基于与序列无关的变化例如表观遗传变化的参考序列的变化。

[0182] 13. 甲基化模式

[0183] 本文提供的方法可用于检测作为靶序列表观遗传变化的序列变异,例如靶序列中甲基化模式的变化。细胞甲基化分析是一个新兴的研究学科。甲基对胞嘧啶的共价加成主要存在于CpG二核苷酸(微卫星)上。尽管不位于启动子区域的CpG岛的功能仍有待探索,但启动子区域中的CpG岛是特别感兴趣的,因为其甲基化状态调节相关基因的转录和表达。启动子区域的甲基化导致基因表达的沉默。这种沉默是永久性的,并通过有丝分裂过程持续。由于其在基因表达中的重要作用,DNA甲基化对发育过程,印记和X染色体失活以及肿瘤发生、衰老以及寄生DNA的抑制均有影响。甲基化被认为参与许多广泛的肿瘤如肺、乳腺和结肠癌以及白血病的癌发生。甲基化和蛋白质功能障碍(长Q-T综合症)或代谢疾病(暂时性新生儿糖尿病,2型糖尿病)之间也存在关联。

[0184] 基因组DNA的亚硫酸氢盐处理可用于分析DNA内甲基化胞嘧啶残基的位置。用亚硫酸氢盐处理核酸将胞嘧啶残基脱氨基成尿嘧啶残基,而甲基化的胞嘧啶保持不变。因此,通过将未用亚硫酸氢盐处理的靶核酸的序列与本文提供的方法中用亚硫酸氢盐处理的核酸的序列进行比较,可以推断出甲基化核酸中的甲基化程度以及胞嘧啶被甲基化的位置。

[0185] 通过使用具有甲基化特异识别位点如HpaII和MSPI的限制性酶,通过限制性内切核酸酶反应进行甲基化分析成为可能。基本原理是某些酶被识别序列中的甲基化胞嘧啶阻断。一旦完成了这种分化,可以使用本文提供的方法进行所得片段的后续分析。

[0186] 这些方法可以与亚硫酸氢盐限制性分析(COBRA)联用。用亚硫酸氢盐处理引起扩增PCR产物中BstUI识别位点的丢失,与未处理的样品相比,其导致在分析中出现新的可检测片段。本文提供的方法可与甲基化位点的特异切割结合使用,以提供关于靶核酸序列中甲基化模式的快速可靠的信息。

[0187] 14.再测序

[0188] 来自各种生物体的可用基因组序列信息的显著增加的量增加了对允许大规模比较序列分析将序列信息与功能、表型或种类相关联的技术的需求。用于比较序列分析的这些技术的应用可以是广泛的,包括SNP发现和病原体的序列特异鉴定。因此,再测序和高通量突变筛选技术对于鉴定疾病背后的突变以及潜在差异药物反应的遗传变异性至关重要。

[0189] 为了满足这些需要,已经开发了几种方法。目前用于高通量DNA测序的技术包括使用电泳和激光诱导荧光检测的DNA测序仪。基于电泳的测序方法对于检测杂合子有固有限制,并且受到GC压缩的影响。因此,在不使用电泳的情况下生成数字数据的DNA测序平台可以克服这些问题。基质辅助激光解吸/离子化飞行时间质谱(MALDI-TOF MS)测量核酸片段并输出数字数据。本文提供的方法允许相对于参考序列检测序列相同性和序列变化的高通量、高速度和高精度。这种方法使得可以常规地使用MALDI-TOF MS测序进行准确的突变检测,如筛选与乳腺癌发展相关的BRCA1和BRCA2中的始祖突变(founder mutation)。

[0190] 15.疾病爆发监测

[0191] 在全球运输和旅游的时代,需要密切监测病原性疫情爆发,以防止其全球传播并实现控制。通过高通量技术进行的基于DNA的分型使得在相对较短的时间内能够快速地进行样品通量,如爆发情况(例如,医院环境中的监测,早期预警系统)所要求的。监测依赖于使用的微生物标记物区域,但可以方便监测属、种、株或亚型特异性水平。这些方法可用于生物防治、临床和药物监测和宏基因组学应用(例如,肠道菌群分析)。治疗进展或失败的这种监测在美国专利号7,255,992、美国专利号7,217,510、美国专利号7,226,739和美国专利

号7,108,974中描述,其通过引用并入本文。

[0192] 16. 疫苗质量控制和生产克隆质量控制

[0193] 本文提供的方法可用于控制重组生产克隆(不限于疫苗)的种类,其可以是疫苗或例如,胰岛素或任何其他生产克隆或生物或医疗产品。

[0194] 17. 药学中用于生产控制和质量的微生物监测

[0195] 本文提供的方法可用于通过,例如,检测这些产物中某些微生物靶核酸的存在或不存在来控制药物产品的质量。

[0196] 试剂盒

[0197] 在一些实施方式中,提供了用于实施本文所述方法的试剂盒。试剂盒经常包含含有一种或多种本文所述组分的一个或多个容器。试剂盒在任意数量的独立容器、包、管、小管、多孔板等中包含一种或多种组分,或者组分可以在此类容器中以不同组合合并。例如一种或多种下列组分可以包含在试剂盒中:(i)一种或多种核苷酸(例如,终止的核苷酸和/或非终止的核苷酸);其中一种或多种可以包括检测标记(ii)一种或多种包含捕获剂的核苷酸;(iii)一种或多种寡核苷酸,其中一种或多种可以包括检测标记(例如,扩增引物,一种或多种延伸引物(UEP),包含标签的寡核苷酸,包含捕获剂的寡核苷酸);(iv)一种或多种酶(例如,聚合酶,内切核酸酶,限制性酶等);(v)控制组分(例如,控制基因组DNA,引物,合成模板,靶核酸等)(vi)一种或多种缓冲液和(vii)印刷物(例如,说明,标记等)。在试剂盒的实施方式中,终止核苷酸的相对量存在于溶液中或以相对的量存在,使得在根据提供的说明溶解时,对于主要核酸物质特异的链终止核苷酸的浓度(C[WT])小于对于次要核酸物质特异的链终止核苷酸的浓度(C[Mut])。在实施方式中,对于主要核酸物质特异的链终止剂的浓度C[WT]小于对次要核酸物质特异的链终止剂的浓度(C[Mut])的20%,C[WT]通常为C[Mut]的约0.5%至小于约20%,C[Mut]的约0.5%至小于约15%,C[Mut]的约1%至约15%,C[Mut]的约1%至约10%,C[Mut]的约2%至约10%或C[Mut]的约1%,1.5%,2%,2.5%,3%,3.5%,4%,4.5%,5.5%,6%,6.5%,7%,7.5%,8%,8.5%,9%,9.5%或10%。在某些实施方式中,C[WT]为C[Mut]的约0.1%至约10%,C[Mut]的约0.01%至约10%,或C[Mut]的约0.01%,0.05%,0.1%,0.5%,1%,1.5%,2%,2.5%,3%,3.5%,4%,4.5%,5%,5.5%,6%,6.5%,7%,7.5%,8%,8.5%,9%,9.5%或10%。

[0198] 试剂盒有时与处理联用,并且能包含实行一种或多种方法的说明书和/或一种或多种组合物的说明。试剂盒可以用于进行本文所述的方法。说明书和/或说明可以是有形形式(如纸等)或电子形式(如有形介质(如压缩盘)等上的计算机可读文件),并且可以包含在试剂盒附件(kit insert)中。试剂盒还可以包括提供这样的说明或描述的互联网位置的书面描述。

[0199] 实施例

[0200] 下述实施例说明但不限制本技术。

[0201] 实施例1-PCR扩增

[0202] 在本文提供的方法的实施方式中,使用合适的扩增引物对包含一种或多种主要核酸物质和一种或多种次要核酸物质的混合物的样品或样品组合进行PCR扩增。示例性的扩增条件如下所示,例如,美国公开专利申请号2013/0237428 A1,美国专利号8,349,566和美国专利号8,003,317以及Oeth等, SEQUENOM® 的申请说明文件号8876-006,R04,于

2006年11月10日出版,其内容通过引用全文纳入本文。例如,扩增条件可以如下设置:

[0203] 表1-PCR混合物

试剂	终浓度	体积(μ l)
	每 5 μ l 反应	(1 rxn)
超纯 H ₂ O	N/A	1.850
10X 含 MgCl ₂ 的 PCR 缓冲液	1.25X	0.625
引物混合物(各 500 nM)	100 nM	1.000
dNTP 混合物(25 mM)	500 μ M	0.100
样品 DNA (主要/次要核酸物质; 5-10 ng/ μ l)	5-10 ng/rxn	1.000
Hotstar Taq® 聚合酶 (5U/ μ l)	0.5 U/rxn	0.100

[0205] 将表1的组分混合并温和涡旋。PCR循环的示例性条件如下:94℃持续15分钟;45个循环,94℃持续20秒/56℃持续30秒/72℃持续1分钟;72℃持续3分钟;冷却至4℃。

[0206] 实施例2-用虾碱性磷酸酶(SAP)处理

[0207] 扩增后,未掺入的dNTP通过用SAP处理脱磷酸化。使用表2所示的试剂组合制备“SAP混合物”:

[0208] 表2-SAP混合物

试剂	体积(μ l)
(1 rxn)	
超纯 H ₂ O	1.330
10X SAP 缓冲液	0.170
SAP 酶 (1U/ μ l)	0.500

[0210] 将根据表2制备的SAP混合物(总体积:2 μ l)加入到来自实施例1的5 μ l PCR反应中。以这种方式处理的每个5 μ l PCR反应物温和混合或涡旋。将处理的样品在标准热循环仪中如下孵育:37℃持续40分钟;85℃持续5分钟;冷却至4℃。

[0211] 实施例3-UEP引物延伸

[0212] 向来自实施例2的每个经SAP处理的样品中加入2 μ l延伸混合物,其示例性的一种

如下表3所示制备：

[0213] 表3-延伸混合物

	试剂	9 μ l 反应中的终浓度	体 积 (μ l) (1 rxn)
	超纯 H ₂ O	N/A	0.755
[0214]	iPLEX 缓冲液 Plus (10X)	0.222X	0.200
	UEP 引物混合物(7 μ M 低质量；14 μ M 高质量)	0.625 μ M 低质量；1.25 μ M 高质量	0.804
	ddNTP 混合物	50-200 μ M C[Mut] _{ddNTP} ； 1-20 μ M C[WT] _{ddNTP}	0.200
	iPLEX 酶	1X	0.041

[0215] 将表1的组分混合并温和涡旋。延伸循环的示例性条件可以如下：在94℃持续30秒的标准热循环仪中进行2-步骤200短循环程序；40个循环的94℃持续5秒/（52℃持续5秒/80℃持续5秒-循环5次，总共200个短循环）；72℃持续3分钟；冷却至4℃。200个短循环程序使用两个循环周期，在40个循环的周期内的一个5循环，从而产生200循环程序。

[0216] 实施例4-MALDI-TOF分析

[0217] 根据实施例3形成的延伸产物的脱盐通过加入6mg CLEAN树脂（塞昆纳姆）来实现。使用纳米分配器，将15nl的裂解反应自动分配到预先装载有基质

（SpectroCHIP[®]，SEQUENOM[®]）的硅芯片上。使用MassARRAY Compact Analyzer（MALDI-TOF质谱仪）获取质谱。将样品转移到384-孔SPECTROCHIP[®]（SEQUENOM[®]；加利福尼亚州圣迭戈）上。整个SPECTROCHIP[®]微芯片转移到BRUKER/SEQUENOM[®]质谱仪中，可以自动测量样品。分析正离子并累积约100个单发光谱（例如，5个光栅位置X20发/位置）。所有样品在线性飞行时间模式下使用延迟离子提取和20kV的总加速电压进行分析。有关说明，请参见SEQUENOM[®]（加利福尼亚州圣迭戈）的“MASSARRAY[®] Nanodispenser用户指南”中的“将引物质量延伸反应产物分配到SPECTROCHIP[®]上（Dispensing Primer Mass Extension Reaction Products onto SPECTROCHIP[®]）”一章。MASSARRAY[®] Typer系统（Typer 3.0版）用于从SPECTROCHIP[®]中获取光谱，如SEQUENOM[®]（加利福尼亚州圣迭戈）的“MASSARRAY[®] Typer用户指南”中的“SPECTROACQUIRE”一章所述。

[0218] 实施例5-使用倾斜的ddNTP浓度的示例性方案和结果

[0219] 如上所述,本文提供的方法结合了样品中次要核酸物质(如次要等位基因)的检测灵敏度,该方法具有相对于主要核酸物质(如主要等位基因)定量其频率、数量或拷贝数的能力。这是通过以下来实现:在延伸反应中调节对于主要核酸物质特异的链终止剂(例如,ddNTP)的浓度范围,使得其小于对于次要核酸物质特异的链终止剂的浓度,从而增加对来自次要核酸物质的信号的检测限,但不低至使得主要核酸物质的信号降低到背景噪声水平,从而排除其用作阳性对照(以确保方法的完整性,即,当检测到对应于次要核酸物质的信号的存在或不存在时,确保观察到的结果是真实的)或作为定量次要核酸物质的相对量(例如,频率,拷贝数)的基础。

[0220] 示例性的方案如下:

[0221] 试验设计

[0222] 每个试验由三种引物、两种PCR引物和一种单碱基延伸引物组成。扩增子(核酸长度可以变化,但通常推荐长度低于150bp,以确保样品,如循环的无细胞DNA和从FFPE(福尔马林固定的石蜡包埋的组织)中分离的降解的DNA中成功扩增)。将质量标签加入到引物的5'末端以将未掺入的PCR引物移出分析质量窗口或允许使用下一代测序进行验证。延伸探针设计也有几个要求。首先,延伸产物的质量必须被质量差异充分间隔以确保试验之间没有冲突(例如,信号的重叠)。其次,多重化是基于设计的反应,使得每个反应容器限于结合相同特异性链终止剂核苷酸的核酸物质。例如,在“A”多重化反应中,反应中所有核酸物质的链终止剂核苷酸均为ddA,所有使用该终止剂核苷酸的试验可以合并在一起。通过允许反向设计获得设计的额外自由度,即,在这种情况下探测相反链上的序列可以研究“T”,因为探针将靶向相反链上的“A”。

[0223] PCR扩增

[0224] 以20μL的总体积进行PCR,用10μL的DNA模板,补充10μL的主混合物,其由补充有1mM MgCl₂,125μM dNTP,0.125U尿嘧啶-DNA糖基化酶(美国马萨诸塞州伊普斯维奇的新英格兰生物实验室公司(New England Biolabs)),4U Taq聚合酶和100nM的各PCR引物组成。反应最初在30℃下孵育10分钟,然后在94℃下孵育2分钟。进行45个循环的94℃持续30秒,56℃持续30秒,和72℃持续1分钟的PCR。以72℃孵育5分钟的最后孵育完成PCR。5μL扩增产物通过添加0.24x虾碱性磷酸酶(SAP)缓冲液中的2μL 0.5U SAP,总体积为7μL,在37℃下持续40分钟,然后将SAP酶在85℃下变性10分钟来调节。除非另有说明,使用的所有试剂均来自阿金纳生物科学公司(Agena Bioscience,Inc.)。

[0225] 单碱基延伸

[0226] 通过加入2μL由0.2x延伸缓冲液,200μM次要等位基因变体核苷酸扩增子和不同浓度的浓度范围为0.25μM-20μM的主要等位基因变体核苷酸扩增子,不同浓度的延伸引物,和0.14U iPLEX®Pro 酶组成的主混合物。在9μL的总体积中进行单碱基延伸反应。反应参数包括在94℃下初始孵育30秒,之后40个循环的在94℃持续5秒钟和52℃持续5秒然后80℃持续5秒的五个嵌套循环。单碱基延伸以在72℃孵育3分钟完成。

[0227] 捕获和数据获取

[0228] 在纳米分配之前,用5μL(3mg)阴离子交换树脂浆料对产物进行调理以进行脱盐。最后,使用RS1000纳米分配器将分析物分配到Spectrochip®II 固体支持物上。使用

MassARRAY® 4仪器通过MALDI-TOF质谱获得数据。

[0229] 根据本文提供的方法的分析的示例性结果示于图1-3。如图1所示,将含有次要核酸物质(少数等位基因)和主要核酸物质(多数等位基因)的混合物的样品根据iPLEX™方法或根据本文提供的方法经过扩增和延伸反应,其中少数等位基因以相对于多数等位基因5%频率存在。最左边的图显示了使用iPLEX™方法的分析结果,其中使用等摩尔浓度的链终止剂。如该图所示,对应于来自多数等位基因的延伸产物的左侧的信号峰值是如此的主导,使得右侧的少数峰降低到背景噪声的水平并且是不可检测的。中图显示当多数等位基因特异的链终止剂的浓度为少数等位基因特异的链终止剂的浓度的20% (1/5) 时获得的结果。如中图所示,来自少数等位基因延伸产物(右峰)的检测信号的强度相对于来自多数等位基因延伸产物的信号(左峰)的检测信号现在更高且更可见。然而,少数等位基因信号仍然很小,并且接近背景噪声的水平。最右侧的图显示当多数等位基因特异的链终止剂的浓度为少数等位基因特异的链终止剂的浓度的约6-7% (十五分之一) 时获得的结果。如右图所示,来自少数等位基因延伸产物(右峰)的信号现在与来自多数等位基因的信号(左峰)相当。因此,图1表明,通过使链终止剂的浓度有利于次要核酸物质偏斜,小于10%的频率的次要核酸物质(其不能通过诸如iPLEX™的方法有效检测到)可通过本文提供的方法检测。

[0230] 另外,如图3所示,本文提供的方法可用于检测以小于相对于样品中的主要核酸物质的2%,低至1.25%的频率存在的次要核酸物质。图3显示了当样品进行扩增,随后在以下三种条件下延伸时获得的结果的比较:(A) 对于主要核酸物质特异的链终止剂ddNTP的浓度(野生型等位基因; $C[WT]_{ddNTP}$) 等于对于次要核酸物质特异的链终止剂ddNTP的浓度(突变等位基因; $C[Mut]_{ddNTP}$); (B) $C[WT]_{ddNTP}$ 小于 $C[Mut]_{ddNTP}$ 浓度的约20% (即 $C[WT]_{ddNTP}:C[Mut]_{ddNTP}$ 的比例为0.2); 和 (C) $C[WT]_{ddNTP}$ 小于 $C[Mut]_{ddNTP}$ 浓度的约5% (即 $C[WT]_{ddNTP}:C[Mut]_{ddNTP}$ 的比例为0.05)。如图3所示,当 $C[WT]_{ddNTP}:C[Mut]_{ddNTP}$ 的比例为1,即两个链终止剂浓度相等(例如,在iPLEX™方法中使用)时,5%、2.5%和1.25%的突变体等位基因频率都低于检测限(LOD)。当 $C[WT]_{ddNTP}:C[Mut]_{ddNTP}$ 的比例降低到0.2时,可以检测到5%的突变体等位基因频率,而2.5%和1.25%的突变等位基因频率仍然低于检测限。当 $C[WT]_{ddNTP}:C[Mut]_{ddNTP}$ 的比例进一步降低到0.05时,可以检测到所有三个突变体等位基因频率(5%、2.5%和1.25%)。

[0231] 实施例6-iPLEX+可行性研究

[0232] 在MassARRAY系统上实施的iPLEX化学提供了对单核苷酸多态性(SNP),体细胞突变和拷贝数变异(CNV)的准确且灵敏的检测。不受理论的限制,iPLEX化学对体细胞突变次要变体检测的敏感性为10%以上。10%次要等位基因频率阈值是源自MassARRAY平台属性的iPLEX方法的实际限制。MassARRAY是一款特征为80cm线性TOF分析仪的MALDI-TOF质谱仪,其最大动态范围为50X(光谱中最小到最高峰)。通过离子信号量和噪音量进一步降低该50X动态范围。例如,iPLEX体细胞突变应用已采用保守的10X动态范围阈值(10%次要变体频率)。iPLEX+化学试图将现有iPLEX化学的敏感度提高到低于10%次要变体的水平。

[0233] 在体细胞突变应用中,突变频率被确定为次要突变峰的强度与优势野生型(WT)峰的强度比或次要突变峰的百分比。在任一情况下,WT峰的强度用作测量突变峰的背景。例如,考虑到MassARRAY仪器的理论50X动态范围和实际的10X动态范围,假设的5%次要等位基因将以低于检测限(LOD)阈值的水平被检测到,因此不会报告为高置信度突变。

[0234] 通过消耗与常规iPLEX相比的WT离子峰值的量来研究常规iPLEX的动态范围限制

的改变。iPLEX的这个修改版本称为iPLEX+。

[0235] 在iPLEX反应期间,通过在200个延伸反应循环的过程中并入模板特异性单终止剂核苷酸(非环核苷酸)将未延伸的引物核苷酸(UEP)转化成产物。延伸反应是双分子反应,其中产物的量(观察到的峰)与WT或次要峰的DNA模板浓度与终止剂核苷酸的量的乘积成比例:

$$[0236] \quad [WT_{产物}] \sim [DNA_{WT}] \times [WT_{终止剂}],$$

$$[0237] \quad [次要_{产物}] \sim [DNA_{次要}] \times [次要_{终止剂}].$$

[0238] 鉴于WT DNA模板的较高浓度和相同浓度的常规iPLEX的WT和次要终止剂,检测到的峰的相对强度将与DNA模板量成比例。降低WT峰值的量并且使平衡“偏向”次要产物的一种方法是与次要终止剂核苷酸相比降低WT终止剂核苷酸的量。iPLEX+是常规iPLEX的WT-耗尽形式,其中四个终止核苷酸的先前等摩尔混合物现在已经针对WT等位基因耗尽(例如,耗尽到总量的1%以下的水平)。通过降低WT-特异性终止剂的量,延伸反应的平衡偏向形成次要产物。次要峰的富集将根据反应可用的WT终止剂的量而变化。

[0239] 进行了一项可行性研究,以确定称为iPLEX+的WT消耗的iPLEX方法(WT-消耗的终止剂核苷酸的混合物)是否将允许低于10%的次要等位基因检测。使用覆盖所有潜在WT-突变情况(转化)的综合试验组的方法,使用模型系统(其中0%,1%,2.5%和5%以稀释系列的人类-黑猩猩DNA来产生)来研究iPLEX+的性能。模型系统用正交技术(离子激流(Ion Torrent)PGM)验证。进行了三轮综合研究,其中针对超过300个试验的iPLEX+组进行了询问,反映了复制整个稀释系列的多个技术。在每个可行性轮次之后,对几个延伸反应参数进行了评估,以提高整个组的性能。其他目标包括确定iPLEX+方法的灵敏度,确定iPLEX+是否是定量的,并建立生物化学的初始参数(终止剂核苷酸混合物,终止剂核苷酸量和延伸反应参数)。该研究表明,iPLEX+方法允许以近100%灵敏度检测5%突变体,对2.5%突变体灵敏度高于90%,对1%突变体的平均灵敏度为35-61%(取决于方法的条件,50%为平均灵敏度)。

[0240] 模型系统

[0241] 用于模拟低频突变和询问所有可能的WT-突变体等位基因组合的模型系统是人和黑猩猩(Pan Troglodytes)基因组DNA的混合物。在人hg19参考组合和黑猩猩2013组装物之间进行全基因组序列比对,以选择具有人和黑猩猩之间的单个稳定核苷酸错配的短直立同源序列。这些地区成为iPLEX+试验候选物。应用于直向同源区域的另一组标准是它们必须在5个外显子或更大的ORF内映射到人外显子,并且同一染色体上的区域之间的距离必须高于5000bp。合格试验候选的最终数量为5,822。

[0242] 试验设计

[0243] Agena Bioscience试验设计套件2.0被用于设计使用来自人-黑猩猩模型的候选物的iPLEX+试验。使用最大多重化水平为15的“体细胞突变”设置实施试验设计。iPLEX+呈现不是通用iPLEX试验设计的一部分的多重化约束。终止混合物的WT-消耗的组合物产生了四种核苷酸之一对于“通道”中所有试验是WT的情况。所有具有WT主要等位基因的多重试验被称为A-通道或A-转化。A转变具有耗尽的延伸所有主要WT产物的A核苷酸,和延伸次要产物的C、G和T核苷酸。试验设计在设计前用具有专门试验方向性的输入序列进行,以消除具有互补核苷酸(试验设计者的共同特征)的反向试验设计。对四个通道(WT-A,C,G,T)分别进

行了试验设计,并将多重化结合到最终设计中。最终的设计具有12个可能的WT-次要转化中的所有功能,每个WT终止剂有三个转化(例如,A-通道具有三个WT-次要转化a/C,a/G,a/T)。组合的试验设计在24次多重化中具有334个试验,并且具有由至少20个试验表示的所有可能的转化。

[0244] 生物样品

[0245] 使用合并的人类基因组HapMap01(科里尔公司(Coriell))DNA样品和来自单个个体的Max(科里尔公司)的黑猩猩DNA产生生物样品。在所有人-黑猩猩混合物中,人类DNA是提供WT等位基因的主要组分,而黑猩猩DNA用作次要组分。每个样品(纯或混合物)的DNA总量设定为约3300个基因组DNA拷贝,对应于每个样品10ng DNA。

[0246] 用iPLEX的试验验证

[0247] 使用100%人样品,100%黑猩猩样品和50%人-黑猩猩混合物,使用常规iPLEX化学验证334个试验。进行常规iPLEX的初步验证以确定表现不佳的试验。表现不佳的试验是具有低延伸率(通常低于80%)的试验和具有较差的模板特异性的试验,其表现为试验偏差(非模板等位基因的非特异性延伸)。试验延伸率计算为所有试验峰上试验产物峰的分数,包括UEP:

$$[0248] \quad \text{延伸率} = \frac{\sum_{j \in \text{产物}}}{\text{UEP} + \sum_{j \in \text{产物}}}$$

[0249] 进行iPLEX验证以确定由于技术或生物学原因而失败的试验。这些原因包括设计不良(引物特异性差,退火特性差,多重化不良等),生物学(PCR或延伸引物位置上的扩增子内的序列变异或突变SNP,试验引物的非特异性退火)或技术原因(试验位于盐加合物的位置,试验引物的质量差)。在334个设计的试验中,15个试验具有低于80%的延伸率,因此被标记为差。除了延伸率之外,常规的iPLEX数据用于确定歪斜或具有非特异性扩增的试验。由于试验询问人类和黑猩猩DNA中的稳定区域,所以期望的是100%的人和100%的黑猩猩样品将是纯合的。然而,在杂合等位基因的位置上,41个试验具有10%或更高的次要等位基因频率。例如,在100%黑猩猩样品中杂合的试验显示G等位基因(人)的非模板延伸,并导致杂合检出,其中仅预期纯合的T等位基因。因此,发现总共49个独特的试验具有差的延伸率或非特异性延伸。设计-组(design-to-panel)的成功率是85%,这在设计的试验的成功率>80%的ADS 2.0验收标准以内。在49个不良试验中,从进一步的分析中去除24个失败的试验(非常差的延伸率<0.5和非模板等位基因延伸)。

[0250] iPLEX+第一轮

[0251] iPLEX+实验的第一轮由以下研究组成。将人/黑猩猩DNA合并成稀释系列,以代表0%,1%,2.5%和5%黑猩猩的贡献。0%黑猩猩样品为100%人类WT对照样品。在所有四种终止混合物(a/CGT,c/AGT,g/CAT和t/ACG)中将WT终止剂核苷酸分数设为1%。每个稀释系列在16个技术重复中被询问。在确定方法性能时评估以下标准:

[0252] • 在0%黑猩猩样品中的试验延伸,

[0253] • 在0%黑猩猩样品中的黑猩猩频率,

[0254] • 试验灵敏度,

[0255] • 转化性能和灵敏度,

[0256] • 总体灵敏度。

[0257] 前两个标准被用作质量控制测量以去除在实验过程中失败的试验。为了评估单独试验的灵敏度,它必须在0%黑猩猩(100%人类WT DNA样品)中延伸以产生最小可检测的WT峰(SNR为5或更高)。在0%黑猩猩没有可靠的WT信号下,由于缺乏与基线的比较,不能评估1%,2.5%和5%黑猩猩的试验性能。WT样品中第二标准-观察到的黑猩猩次要等位基因,用于去除由于试验设计而非iPLEX+方法而失败的非特异性试验。

[0258] 鉴于iPLEX+的WT-消耗的原理,预计在100%WT时,iPLEX+试验的延伸率比5%黑猩猩和常规iPLEX的延伸率要低。常规iPLEX的平均试验延伸率为0.95,而iPLEX+0%黑猩猩样品的延伸率为0.40,并且iPLEX+5%黑猩猩样品为0.78。核苷酸终止剂混合物的WT-消耗的制剂使平衡偏向次要等位基因产物,因此,没有次要产物的100%人类WT样品显示出低得多的试验延伸率。可行性研究的任务之一是建立在100%WT样品中确保试验扩增的WT核苷酸浓度。在评估第一轮的方法性能后,96个试验在100%人类WT样品中没有延伸,36个试验在WT样品中具有非特异性的黑猩猩延伸。因此,只有57%的试验是有功能的。

[0259] 在去除较差的试验后,估计试验性能和组灵敏度。用于估计试验性能的主要标准是可与0%黑猩猩WT样品区分的来自1%、2.5%和5%黑猩猩稀释度的实验观察到的黑猩猩频率。黑猩猩频率来自两个峰之间,人类等位基因峰和黑猩猩等位基因峰的等位基因比率。等位基因峰比率表示为如下的数字基因型:

$$[0260] \quad \text{数字基因型} = \frac{\text{强度}_{\text{低质量}}}{(\text{强度}_{\text{低质量}} + \text{强度}_{\text{高质量}})} \quad (\text{高质量下标})$$

[0261] 其中强度_{低质量}是低质量等位基因的峰强度并且强度_{高质量}是高质量等位基因的峰强度。数字基因型表示当黑猩猩等位基因处于低质量时的黑猩猩次要等位基因频率。在黑猩猩等位基因处于高质量的情况下,黑猩猩频率计算为1-数字基因型。

[0262] 评估了16项技术重复的每个试验的性能。将数据(观察到的芯片频率)绘制为每个试验的晶格图,y轴作为观察到的黑猩猩频率,并且x轴作为稀释系列,WT样品为0%黑猩猩。根据是否可将样品以统计学显著的程度与0%黑猩猩区分开来评估数据。所使用的统计数据是使用Bonferroni校正进行多次比较的t检验,并且将临界α值设置为0.95。试验chr2_AtoC_104可以显著区分2.5%和5%黑猩猩样品与0%黑猩猩样品。以下两种试验,chr20_AtoC_118和chr4_AtoC_146只能区分5%与0%黑猩猩,而chr1_AtoG_115试验可以区分所有稀释度与0%黑猩猩。通过使用每个试验的显著性结果计算iPLEX+的每个转化和整体灵敏度。图4显示了iPLEX+的每个转化灵敏度。计算每种稀释水平(0%,1%,2.5%和5%黑猩猩)的灵敏度,作为与该稀释水平在统计学上显著不同(0%黑猩猩)的试验的平均数。例如,图4的左上角显示,1%黑猩猩的灵敏度为0,2.5%黑猩猩的敏感度为0.33,并且5%黑猩猩的敏感度为1.0,基于代表A到C转化的试验。这意味着在该转化的任何试验中,1%不能与0区分,在33%的时间上可以区分2.5%稀释度,在100%的时间上可以区分5%稀释度。每个稀释系列的全组灵敏度对于1%为0.61,对于2.5%为0.88,并且对于5%黑猩猩频率样品为1.0。

[0263] iPLEX+灵敏度根据试验和转化而不同。观察到挑战iPLEX+的定量方面的重要的试验特异性和转化特异性偏倚。图5显示了使用5%稀释黑猩猩DNA进入人类的不同转化中的试验所观察到的黑猩猩频率。

[0264] 在可行性研究的第1轮中,显示iPLEX+的WT核苷酸消耗方法可用于成功降低WT等位基因信号,并将试验灵敏度提高到低于10%的水平。根据该数据,iPLEX+可用于以下出色灵敏度检测低于10%的次要变体,对于5%突变体为100%,对于2.5%突变体为88%,并且对于1%突变体为61%。第1轮方法的一些缺点包括WT样品中差的试验延伸和试验特异性核苷酸掺入。为了克服低试验延伸,降低产物强度并减轻一些试验特异性掺入偏倚,进行一轮优化,其中核苷酸终止混合物(对于WT核苷酸设为1%,对于第1轮中突变体等位基因为等摩尔)针对各转化进行优化。

[0265] 核苷酸终止剂优化

[0266] 使用5%黑猩猩稀释度的技术重复进行核苷酸终止剂优化,得到每个试验获得16个数据点(14个重复和2个NTC)。评估以下终止混合物:2%,4%,8%和16%WT核苷酸。核苷酸终止剂优化的目标是确定每个转化的WT核苷酸终止剂的最佳比例。优化标准是5%黑猩猩次要等位基因峰与WT峰的相对峰高。优化目标是产生终止剂混合物,导致WT和5%黑猩猩峰值相等的峰高。

[0267] 不同浓度的WT核苷酸产生黑猩猩次要产物的不同反应。将WT终止剂的量从1%增加到16%可以减少所有转化的观察到的次要等位基因分数。将WT核苷酸部分从8%增加到16%导致5%黑猩猩次要等位基因的信号减少,这使得iPLEX+与常规iPLEX表现相似。与以前的iPLEX+实验一样,每个WT百分比的数据点的一般扩散表明转化特异性等位基因的并入。然而,2%WT核苷酸量是平均检测5%黑猩猩次要等位基因作为0.5黑猩猩等位基因频率的最佳量。

[0268] 计算每个转化的调整的WT核苷酸百分比(基于该转化的平均值)以产生0.5个黑猩猩等位基因频率。这些数据见表4。调整的WT核苷酸百分比(底行)被选择用于下一轮iPLEX+可行性。

[0269] 表4

[0270]

百分比 WT	AtoC	AtoG	AtoT	CtoA	CtoG	CtoT	GtoA	GtoC	GtoT	TtoA	TtoC	TtoG
1	0.63	0.68	0.68	0.7	0.74	0.71	0.71	0.57	0.48	0.68	0.72	0.69
2	0.38	0.46	0.44	0.54	0.57	0.53	0.58	0.39	0.33	0.49	0.5	0.48
4	0.22	0.28	0.19	0.4	0.42	0.35	0.4	0.22	0.2	0.37	0.32	0.34
8	0.16	0.21	0.11	0.29	0.3	0.23	0.27	0.12	0.12	0.25	0.17	0.21
16	0.13	0.17	0.07	0.2	0.21	0.13	0.19	0.06	0.09	0.2	0.12	0.18
黑猩猩等位基因频率	1.5	2.15	1.9	3	3.5	2.8	3.25	1.15	1	2.5	2.5	2.5

[0271]

率的预期 % wt(黑猩猩/(黑猩猩+人))												
---------------------------	--	--	--	--	--	--	--	--	--	--	--	--

[0272] iPLEX+第二轮

[0273] 除终止混合物以外,第二轮iPLEX+可行性研究与第一轮的布局和范围相同。第1轮在其通用核苷酸混合物中使用1%WT核苷酸终止剂进行,而第2轮使用转化特异性WT核苷酸量进行(表4)。iPLEX+可行性的第二轮由以下研究组成。将人/黑猩猩DNA合并成稀释系列,以代表0%,1%,2.5%和5%黑猩猩的贡献。每个稀释系列在16个技术重复中被询问。如第1轮的情况,在确定方法性能时评估了以下标准:

- [0274] • 在0%黑猩猩样品中的试验延伸,
- [0275] • 在0%黑猩猩样品中的黑猩猩频率,
- [0276] • 试验灵敏度,
- [0277] • 转化性能和灵敏度,
- [0278] • 总体灵敏度。

[0279] 与第1轮相似,在第2轮中鉴定出表现不佳的试验。44个试验(与第1轮中的96个相比)在100%WT样品中没有产生足够的信号。这是第1轮的2.2倍的改进,完全是由于第2轮中的WT核苷酸浓度增加。除了失败的试验外,33个试验在100%WT样品中延伸了黑猩猩等位基因(与第1轮中的36个相比)。与第1轮中的136个试验(组的43%)相比,总共77个试验(组的25%)在第2轮中较差。从进一步分析中除去试验。

[0280] 研究了增加WT核苷酸贡献对延伸率的影响。(0%黑猩猩,5%黑猩猩和常规iPLEX)。没有观察到与第1轮中的延伸率的显著变化。0%黑猩猩样品的平均延伸率为0.41(与第1轮中0.40相比),5%黑猩猩样品的平均延伸率为0.63(与第1轮中0.78相比)。

[0281] 使用与第1轮相同的方法,检查了每个转化的灵敏度和组的整体灵敏度。观察到每个转化水平的适度变化,然而,第2轮组的整体表现与第1轮相似。

[0282] 第2轮小组的总体灵敏度为1%黑猩猩为0.52(与第1轮中0.61相比),2.5%黑猩猩为0.91(与第1轮中0.88相比),并且5%黑猩猩为0.98(与第1轮中1.0相比)。此外,如在第1轮中那样,观察到显著程度的试验特异性延伸,导致在每个稀释系列观察到的次要等位基因部分的高变异性。因此,转化特异性WT核苷酸方法的主要益处是与第1轮(96个试验)中相比,在0%处较少的试验在第2轮(44个试验)中没有延伸。

[0283] 模型系统的正交验证

[0284] 在第1轮之后,进行核苷酸优化和第2轮研究了一组实验以验证iPLEX+开发中使用的人-黑猩猩模型系统。建立实验(如模型系统中所述)以询问在人和黑猩猩基因组中不同的靶区域,并且产生稀释系列样品以代表在主要人类等位基因背景上的0%黑猩猩,1%黑猩猩,2.5%黑猩猩和5%黑猩猩次要等位基因频率。使用离子激流PGM(生命技术公司(Life Technologies))进行一组实验以验证目标区域并验证稀释度。通过使用iPLEX PCR混合物和dNTPS扩增iPLEX(+)设计的扩增子来制备用于测序的文库。该PCR反应中推荐的gDNA输入为拷贝10-20ng/多重化。使用BioAnalyzer DNA 1000试剂盒测量每个多重化中的DNA量并合并成等摩尔部分。针对离子激流输入 NEBNext®快速DNA片段化和文库制备组的扩增子文库对于各多重化为50ng。离子激流文库的制备涉及末端修复,适配器连接P1和A适配器,切口平移和小的6x循环PCR。单独步骤之后的清理是使用Ampure XP珠。使用BioAnalyzer DNA 1000试剂盒进行库QC和定量,并且输入到离子激流-PGM OT2 200试剂盒的文库为26pmol。使用离子激流OT2ES机器进行模板化ISP的富集。使用离子激流200测序试剂盒和240个核苷酸的流,在PGM系统上进行了这些模板化ISP的测序。使用单个318或318v2芯片对

每个模型进行测序。通过将测序读数与包含iPLEX(+)设计的靶扩增子的基因组坐标的BED文件进行比对来进行数据分析。使用离子激流覆盖分析和变体检出者插件来自动计算和验证目标表示和次要变体频率。

[0285] 证实了使用5%黑猩猩稀释库的310个目标(90%的目标)中的离子数据278.5%黑猩猩文库的中值黑猩猩次要等位基因变体频率为4.9%，中值目标覆盖率为10,000倍。分析的最低覆盖率截止值为5,000倍。2.5%稀释系列文库导致189个目标被确认(61%)，中值次要等位基因变体频率为3.6%，中值覆盖率为8,200倍。

[0286] 每个稀释库包含小于5%的异常值，具有比预期的5%和2.5%高得多的黑猩猩等位基因频率。检查5%和2.5%稀释系列靶标之间的相关性明显发现在5%和2.5%数据中发现的变体异常值属于相同目标试验。

[0287] 1%离子PGM运行中没有鉴定到1%的变体主要是由于芯片负载不良而没有产生足够的目标覆盖来鉴定1%的黑猩猩变体。0%的黑猩猩文库没有按预期产生任何变体。

[0288] 通过PGM数据证实，人类-黑猩猩文库的总体质量是很好的，90%的来自5%黑猩猩稀释度的试验目标通过PGM确认达到几乎相同的4.9%的中值次要变体频率。文库异常值属于相同试验目标的证据表明，这些伪影是由于试验设计中的挑战而不是iPLEX+方法。

[0289] 核苷酸量调整

[0290] 第2轮iPLEX+可行性工作持续改进了WT样品中iPLEX+试验的延伸率。以转化特异性方式增加WT核苷酸浓度导致由于WT中差的延伸而失败的试验数量减少2.2。另外，在使用WT样品询问iPLEX+组的研究中考虑了延伸反应的以下参数，并研究了它们对WT样品中的试验延伸率的影响：

[0291] • PCR退火时间-延伸期间的退火时间增加将允许聚合酶(Thermosequenase)更多次“找到”消耗的WT核苷酸。

[0292] • 总核苷酸浓度-保持与第2轮相同的WT核苷酸比例，但允许较高的总量将增加WT掺入率。

[0293] • 300循环PCR-将循环次数从200增加到300将产生更多的产物和更高强度的等位基因峰。

[0294] • 增加聚合酶的量-聚合酶浓度的增加会增加产物的量，由于更快的周转率并导致更高强度的等位基因峰。

[0295] 图6显示了如上所述的每种条件的试验延伸速率的箱体图。

[0296] 在0%黑猩猩WT样品中产生最高试验延伸率的条件是#1,#2,#5和#6(在x轴上从左到右)。这四个条件的恒定组分是2X核苷酸量。2X核苷酸量代表核苷酸的总浓度的两倍，WT转化-特异性比例保持第2轮中的比例。

[0297] iPLEX+第三轮

[0298] iPLEX+可行性研究的第三轮也是最后一轮利用了先前确定的转化特异性WT核苷酸浓度和2X总核苷酸浓度。与第1轮和第2轮相似，iPLEX+可行性第3轮由以下研究组成。将人/黑猩猩DNA合并成稀释系列，以代表0%，1%，2.5%和5%黑猩猩的贡献。每个稀释系列在16个技术重复中被询问。如第1轮的情况，在确定方法性能时评估了以下标准：

[0299] • 在0%黑猩猩样品中的试验延伸，

[0300] • 在0%黑猩猩样品中的黑猩猩频率，

[0301] • 试验灵敏度,

[0302] • 转化性能和灵敏度,

[0303] • 总体灵敏度。

[0304] 由于0%黑猩猩WT样品和在WT样品中非特异性延伸黑猩猩的33个样品中延伸不良,因此37个试验失败。与第2轮(44个试验)相比,存在进一步改善。在第3轮中所得的差试验分数为23%(与第2轮中的25%相比)。还观察到通过使用2N总终止剂核苷酸的试验延伸率的适度改善。在0%黑猩猩样品中,0%和5%黑猩猩样品中的平均试验延伸率为0.53和0.72。这相比第2轮结果有所改善(0%黑猩猩为0.41,并且5%黑猩猩样品为0.63)。如第1和第2轮那样,确定了每个转化的灵敏度和组的整体灵敏度。观察到每个转化水平的适度变化,然而,第3轮组的整体表现再次与前两轮的总体表现相似,但1%黑猩猩灵敏度除外。

[0305] 第3轮小组的总体灵敏度为1%黑猩猩为0.35(与第2轮中0.52相比),2.5%黑猩猩为0.89(与第2轮中0.91相比),并且5%黑猩猩为0.99(与第1轮中0.98相比)。1%黑猩猩样品的3轮试验的灵敏度从0.52降低到0.35。不受理论的限制,观察到的终止剂核苷酸浓度增加对1%黑猩猩样品的组灵敏度降低(第1轮的0.61,第2轮的0.52,和第3轮的0.35)的两种可能解释如下:

[0306] • 增加WT核苷酸百分比和总核苷酸浓度导致主要WT等位基因的峰强度增加。这又导致该方法执行类似于常规iPLEX,其中1%次要等位基因信号是低强度噪音峰。

[0307] • 增加WT核苷酸百分比和总核苷酸浓度将不良延伸试验的数量从第1轮的96减少到第2轮的44和第3轮的37。在第2轮和第3轮中几乎没有通过QC的“最近较差”试验降低1%黑猩猩灵敏度,因为它们具有较低的信噪比,但是当计算整体灵敏度时,它们被考虑在内。

[0308] 实施例7-最优核苷酸浓度

[0309] iPLEX+反应中每种可能转化的最佳核苷酸浓度如表5所示。还显示了WT/Mut(主要等位基因/次要等位基因)比例。

[0310] 表5

[0311]

混合物	核苷酸	反应中的浓度 (mM)	Wt/Mut 比例
A	A	0.007	
	C	2.222	0.003
	G	0.311	0.021
	T	0.889	0.008
C	A	2.222	0.006
	C	0.013	
	G	0.381	0.035
	T	0.889	0.015
G	A	1.111	0.013
	C	1.111	0.013

[0312]

	G	0.014	
	T	1.111	0.013
T	A	1.111	0.0125
	C	1.111	0.0125
	G	2.778	0.005
	T	0.0139	

[0313] 实施例8-iPLEX+验证研究

[0314] 进行了一项研究以验证iPlex+试验。测试了NRAS和EGFR基因(模型)中体细胞突变检测的功能、灵敏度和再现性。

[0315] 试验功能

[0316] 实验针对5%突变体使用200拷贝/u1突变体。从该起始值滴定所有其他次要变体。野生型贡献为4000拷贝/u1。

[0317] • 使用野生型模板和标准iPLEX终止混合物评估试验的合适野生型延伸。评估无法延伸合适的野生型核苷酸的试验的非特异性相互作用并且,如果合适重新设计或移到另一重中。

[0318] • 评估试验的对非模板对照中的非特异性延伸,并将其移到另一重中。

[0319] • 使用定制的核苷酸混合物评估试验在5%突变体模型中合适延伸。观察到所有合适的变体基因型。

[0320] • 评估了试验的外切核酸酶活性和附加信号。

[0321] • 通过峰SNR (信噪比) 评估试验。在开发期间,三个试验显示在峰SNR中重复的差的性能。检查每个试验的原因。提出的原因是通过新设计和/或重再分配减弱。

[0322] 试验灵敏度

[0323] 初步研究调查了在5%或低于5% (初始核苷酸混合物) 下实现次要变体检测所必需的核苷酸比例。样品一式三份运行,WT样品一式三份运行,各重也一样。初始核苷酸混合物没有提供所需的灵敏度。高野生型信号导致突变体比例低,这不允许与其他突变类别区分 (数据未显示)。

[0324] 为了提高对突变体等位基因的灵敏度,WT核苷酸贡献降低。该评估使用WT和NTC样品的5%、2%和1%次要变体。各样品一式三份运行。野生型的减少成功地增加了突变体比例以允许更大的次要变体鉴别,然而几个试验的WT峰值的SNR下降到用于检出峰的截止值 (SNR为5) 以下 (数据未显示)。

[0325] WT核苷酸浓度在调整的核苷酸混合物中略有增加,以增加WT SNR。该评估使用与先前的实验相同的标准和次要变体来研究WT核苷酸的减少。WT样品一式两份运行,并且各重一个NTC。将混合物的WT贡献提高到原始混合物和降低的WT组合之间的浓度,成功地调整了许多具有差的WT信号或差的变体类型之间鉴别的试验 (数据未显示)。评估所有三种核苷酸混合试验的数据表明使用WT和突变体之间的定制核苷酸比例。

[0326] 用与前两次实验相同的十二个对照一式四份进行定制核苷酸混合物的评估。WT和NTC也按照每重一式四份运行。定制的核苷酸比例混合物成功地提供可以分化的突变比,同时保持足够的WT SNR以允许由软件识别和检出峰 (SNR值大于5) (参见表6)。定制的核苷酸混合物 (表6) 表示表5的最佳核苷酸浓度混合物的进一步精修。

[0327] 表6

[0328]

混合物	核苷酸	反应中的 浓度 (mM)	Wt/Mut 比
A	A	0.007	
	C	2.222	0.003
	G	0.311	0.021
	T	0.889	0.008
C	A	2.222	0.006
	C	0.013	
	G	0.381	0.035
	T	0.889	0.015
G	A	1.111	0.013

[0329]		C	1.111	0.013
		G	0.014	
		T	1.111	0.013
	T	A	1.111	0.0003
		C	1.111	0.0003
		G	2.778	0.0001
		T	0.0003	

[0330] 再现性

[0331] 灵敏度实验中使用的各次要变体模型 (5%, 2.5%, 1% 和 WT) 以 24 次重复 (96 次重复的 WT) 进行评估。各模型次要变体类型产生统计学显著的不同分布 (数据未显示)。

[0332] 实施例9-实施方式的非限制性示例

[0333] 下文提供了本技术的某些实施方式的非限制性例子。

[0334] A1. 一种用于鉴定核酸群中一种或多种次要核酸物质的存在或不存在的多重化方法, 其包括所述一种或多种次要核酸物质和一种或多种主要核酸物质的混合物, 其中每种次要核酸物质是对应的主要核酸物质的变体, 并且其拷贝数小于其对应的主要核酸物质的拷贝数, 其中所述方法包括:

[0335] (a) 在包括 dNTP 的扩增条件下用扩增引物对同时扩增混合物的靶区域, 由此产生包含主要和次要核酸物质的核酸的扩增混合物;

[0336] (b) 在包括链终止剂的延伸条件下使扩增混合物与延伸引物接触, 其中:

[0337] (i) 一种或多种主要核酸物质共有对于主要核酸物质特异且对次要核酸物质非特异的共同链终止剂, 和

[0338] (ii) 一种或多种次要核酸物质中的每一种具有对于次要核酸物质特异并且对于主要核酸物质非特异的链终止剂, 其中对于次要核酸物质特异的链终止剂: (A) 对于扩增混合物中的特定次要核酸物质是独特的, 并且不被扩增混合物中的其他次要核酸物质所共有, 或 (B) 一种或多种次要核酸物质中的至少一种与扩增混合物中的至少一种其他次要核酸物质共有共同的链终止剂,

[0339] 由此引物延伸到相对于主要核酸物质的次要核酸物质中不同的核苷酸位置或通过其中, 从而产生分别对应于次要核酸物质和主要核酸物质的链终止的延伸产物, 其中对于主要核酸物质特异的链终止剂的浓度小于一种或多种次要核酸物质特异的每种链终止剂的浓度; 和

[0340] (c) 分析 (b) 的延伸产物, 从而鉴定一种或多种次要核酸物质的存在或不存在。

[0341] A2. 如实施方式 A1 所述的方法, 其中核酸群包括多种次要核酸物质, 其是单一主要核酸物质的变体, 并且在单一多重化反应中鉴定该多种次要核酸物质。

[0342] A3. 如实施方式 A1 所述的方法, 其中 (b) 在一组至少 2 个反应容器或隔室中进行, 其中:

[0343] 第一容器或隔室包括含有针对主要核酸物质特异的链终止剂并且不含针对一种

或多种次要核酸物质特异的链终止剂的延伸条件;和

[0344] 每个剩余的容器或隔室包括延伸条件,其包含对一种或多种次要核酸物质特异且共有的单一链终止剂,并且不含对于主要核酸物质特异或针对次要核酸物质特异且不共有的单一链终止剂的链终止剂。

[0345] A4.如实施方式A1-A3中任一项所述的方法,其中每个链终止剂的浓度是已知的。

[0346] B1.一种定量核酸群中一种或多种次要核酸物质的方法,所述核酸群包括所述一种或多种次要核酸物质和主要核酸物质的混合物,其中所述次要核酸物质是相同的主要核酸物质的变体并且各自以比主要核酸物质的拷贝数低的拷贝数存在,所述方法包括:

[0347] (a) 在包括dNTP的扩增条件下用扩增引物对同时扩增混合物的靶区域,由此产生包含主要和次要核酸物质的核酸的扩增混合物;

[0348] (b) 在延伸条件下使扩增混合物与延伸引物接触,所述延伸条件包括对于(i)一种或多种次要核酸物质中的每一种,和(ii)主要核酸物质特异性链终止剂,由此引物延伸到相对于主要核酸物质的次要核酸物质中不同的核苷酸位置或通过其中,从而产生分别对应于次要核酸物质和主要核酸物质的链终止的延伸产物,其中:(1)每一种链终止剂的浓度是已知的;并且(2)对于主要核酸物质特异的链终止剂的浓度小于对于一种或多种次要核酸物质特异的链终止剂的浓度;

[0349] (c) 确定对应于一种或多种次要核酸物质中的每一种的延伸产物的量相对于对应于主要核酸物质的延伸产物的量的比例;和

[0350] (d) 基于(c)的比例,并且基于对于一种或多种次要核酸物质特异的链终止剂的浓度,相对于主要核酸物质的量对次要核酸物质的量进行定量,所述对于一种或多种次要核酸物质特异的链终止剂的浓度相对于主要核酸物质特异性链终止剂的浓度。

[0351] C1.一种定量核酸群中一种或多种次要核酸物质的多重化方法,所述核酸群包括所述一种或多种次要核酸物质和主要核酸物质的混合物,其中每一种次要核酸物质是对应的主要核酸物质的变体并且以比其对应的主要核酸物质的拷贝数低的拷贝数存在,所述方法包括:

[0352] (a) 在包括dNTP的扩增条件下用扩增引物对同时扩增混合物的靶区域,由此产生包含主要和次要核酸物质的核酸的扩增混合物;

[0353] (b) 在包括链终止剂的延伸条件下使扩增混合物与延伸引物接触,其中:

[0354] (i) 一种或多种主要核酸物质共有对于主要核酸物质特异且对次要核酸物质非特异的共同链终止剂,和

[0355] (ii) 一种或多种次要核酸物质中的每一种具有对于次要核酸物质特异并且对于主要核酸物质非特异的链终止剂,其中对于次要核酸物质特异的链终止剂:(A)对于扩增混合物中的特定次要核酸物质是独特的,并且不被扩增混合物中的其他次要核酸物质所共有,或(B)一种或多种次要核酸物质中的至少一种与扩增混合物中的至少一种其他次要核酸物质共有共同的链终止剂,

[0356] 由此引物延伸到相对于主要核酸物质的次要核酸物质中不同的核苷酸位置或通过其中,从而产生分别对应于次要核酸物质和主要核酸物质的链终止的延伸产物,其中:(1)每一种链终止剂的浓度是已知的;和(2)对于主要核酸物质特异的链终止剂的浓度小于对于一种或多种次要核酸物质特异的链终止剂的浓度;

[0357] (c) 确定对应于一种或多种次要核酸物质中的每一种的延伸产物的量相对于对应于主要核酸物质的延伸产物的量的比例;和

[0358] (d) 基于(c)的比例,并且基于对于一种或多种次要核酸物质特异的链终止剂的浓度相对于对于主要核酸物质特异的链终止剂的浓度,相对于主要核酸物质的量对次要核酸物质的量进行定量。

[0359] C2.如实施方式B1或实施方式C1所述的方法,其中核酸群包括多种次要核酸物质,其是单一主要核酸物质的变体,并且在单一多重化反应中鉴定该多种次要核酸物质。

[0360] C3.如实施方式C1所述的方法,其中(b)在一组至少2个反应容器或隔室中进行,其中:

[0361] 第一容器或隔室包括含有针对主要核酸物质特异的链终止剂并且不含针对一种或多种次要核酸物质特异的链终止剂的延伸条件;和

[0362] 每个剩余的容器或隔室包括延伸条件,其包含对一种或多种次要核酸物质特异且共有的单一链终止剂,并且不含对于主要核酸物质特异或针对次要核酸物质特异且不共有的共同的单一链终止剂的链终止剂。

[0363] D1.如实施方式A1-A4中任一项所述的方法,其中所述次要和主要核酸物质的序列相差单个碱基,并且引物延伸到或通过不同的单一碱基。

[0364] D2.如实施方式A1-A4、B1、C1-C3和D1中任一项所述的方法,其中所述次要核酸物质的序列包括相对于所述主要核酸物质的序列的插入或缺失。

[0365] D3.如实施方式A1-A4、B1、C1-C3、D1和D2中任一项所述的方法,其中一种或多种次要核酸物质是主要核酸物质的单核苷酸多态性(SNP)变体。

[0366] D4.如实施方式A1-A4、B1、C1-C3和D1-D3中任一项所述的方法,其中所述次要和主要核酸物质分别是相同基因的突变体和野生型等位基因。

[0367] D5.如实施方式A1-A4、B1、C1-C3和D1-D4中任一项所述的方法,其中所述主要核酸物质来自宿主对象,并且所述次要核酸物质来自除所述宿主以外的对象,或者所述次要核酸物质来自宿主对象并且所述主要核酸物质来自除所述宿主以外的对象。

[0368] D6.如实施方式A1-A4、B1、C1-C3和D1-D5中任一项所述的方法,其中所述一种或多种次要核酸物质各自以小于主要核酸物质的拷贝数的约10%的拷贝数存在。

[0369] D7.如实施方式D6所述的方法,其中所述一种或多种次要核酸物质各自以主要核酸物质的拷贝数的约1%至小于10%的拷贝数存在。

[0370] D8.如实施方式D7所述的方法,其中所述一种或多种次要核酸物质各自以主要核酸物质的拷贝数的约2%至小于10%的拷贝数存在。

[0371] D9.如实施方式D7所述的方法,其中所述一种或多种次要核酸物质各自以主要核酸物质的拷贝数的约5%或更少的拷贝数存在。

[0372] D10.如实施方式D7所述的方法,其中所述一种或多种次要核酸物质各自以主要核酸物质的拷贝数的约2%或更少的拷贝数存在。

[0373] D10.1.如实施方式D7所述的方法,其中所述一种或多种次要核酸物质各自以主要核酸物质的拷贝数的约1%的拷贝数存在。

[0374] D11.如实施方式A1-A4、B1、C1-C3和D1-D10.1中任一项所述的方法,其中对于主要核酸物质特异的链终止剂的浓度为每个针对次要核酸物质特异的链终止剂的浓度的约1%

至约20%。

[0375] D12.如实施方式A1-A4、B1、C1-C3和D1-D10.1中任一项所述的方法,其中对于主要核酸物质特异的链终止剂的浓度为每个针对次要核酸物质特异的链终止剂的浓度的约0.1%至约10%。

[0376] D13.如实施方式D12所述的方法,其中对于主要核酸物质特异的链终止剂的浓度为每个针对次要核酸物质特异的链终止剂的浓度的约0.1%至约4%。

[0377] D13.1.如实施方式A1-A4、B1、C1-C3和D1-D10.1中任一项所述的方法,其中对于主要核酸物质特异的链终止剂的浓度为每个针对次要核酸物质特异的链终止剂的浓度的约0.01%至约10%。

[0378] D13.2.如实施方式D13.1所述的方法,其中对于主要核酸物质特异的链终止剂的浓度为每个针对次要核酸物质特异的链终止剂的浓度的约0.01%至约4%。

[0379] D14.如实施方式A1-A4、B1、C1-C3和D1-D13.2中任一项所述的方法,其中所述对于主要核酸物质的链终止剂的浓度基于特定的主要物质/次要物质转化而变化。

[0380] D15.如实施方式A1-A4、B1、C1-C3和D1-D14中任一项所述的方法,其中所述链终止剂是链终止核苷酸。

[0381] D16.如实施方式D15所述的方法,其中所述链终止核苷酸独立地选自ddATP、ddGTP、ddCTP、ddTTP和ddUTP。

[0382] D17.如实施方式D15或D16所述的方法,其中针对一种或多种次要核酸物质特异的链终止核苷酸由一个链终止核苷酸组成。

[0383] D18.如实施方式D15或D16所述的方法,其中针对一种或多种次要核酸物质特异的链终止核苷酸由2个链终止核苷酸组成。

[0384] D19.如实施方式D15或D16所述的方法,其中针对一种或多种次要核酸物质特异的链终止核苷酸由3个链终止核苷酸组成。

[0385] D20.如实施方式A1-A4、B1、C1-C3和D1-D19中任一项所述的方法,其中所述链终止剂包括一种或多种无环终止剂。

[0386] D21.如实施方式A1-A4、B1、C1-C3和D1-D20中任一项所述的方法,包含(a)中约30至约45个PCR扩增循环。

[0387] D22.如实施方式A1-A4、B1、C1-C3和D1-D21中任一项所述的方法,其中(b)中的延伸条件包含约20至约300个循环。

[0388] D23.如实施方式D22所述的方法,其中(b)中的延伸条件包括至少50个循环。

[0389] D24.如实施方式A1-A4、B1、C1-C3和D1-D23中任一项所述的方法,其中一种或多种链终止剂包括可检测的标记。

[0390] D25.如实施方式D24所述的方法,其中所述标记是荧光标记或染料。

[0391] D26.如实施方式D24所述的方法,其中所述标记是质量标记。

[0392] D27.如实施方式D24-D26中任一项所述的方法,还包括检测所述标记,从而鉴定或定量所述一种或多种次要核酸物质。

[0393] D28.如实施方式D27所述的方法,其中所述标记是质量标记,并通过质谱进行检测。

[0394] D28.1.如实施方式D27所述的方法,其中所述标记是荧光标记或染料并且通过电

泳或实时PCR进行检测。

[0395] D29.如实施方式A1-A4、B1、C1-C3和D1-D28.1中任一项所述的方法,其中(a)中的扩增反应条件包括水、基因组DNA、缓冲液、dNTP、引物对、 MgCl_2 和聚合酶,其中 MgCl_2 的浓度与dNTP中每一种的浓度之比选自 $\leq 10:1$ 、 $\leq 9:1$ 、 $\leq 8:1$ 、 $\leq 7:1$ 、 $\leq 6:1$ 或 $\leq 5:1$ 。

[0396] D30.如实施方式D29所述的方法,其中所述聚合酶是至少约0.03单位/ μl 浓度的Taq聚合酶。

[0397] D31.如实施方式A1-A4、B1、C1-C3和D1-D30中任一项所述的方法,其中(a)中的扩增反应条件包含约400-700 μM 的各dNTP,约100nM引物对和约1.6至约4.8mM MgCl_2 。

[0398] D32.如实施方式A1-A4、B1、C1-C3和D1-D31中任一项所述的方法,其中序列标签附着于扩增引物对中的一种或多种引物。

[0399] D33.如实施方式A1-A4、B1、C1-C3和D1-D32中任一项所述的方法,其中游离 Mg^{2+} 浓度是1.0-2.0mM。

[0400] ***

[0401] 本文中引用的各专利、专利申请、出版物和文献的全部内容均通过引用纳入本文。对上述专利、专利申请、出版物和文献的引用并不表示承认上述任何内容是相关的现有技术,也并不表示承认这些出版物或文献的内容或日期。

[0402] 可以对上述内容进行改变而不背离本技术的基本方面。尽管参考一个或多个具体实施方式充分详细描述了本技术,但是本领域普通技术人员应认识到可对本申请中具体公开的实施方案进行改变,而这些改良和改进在本技术的范围和精神内。

[0403] 本文中适当地说明性描述的技术可在没有任何本文未具体公开的元素的情况下实施。因此,例如,在本文的各个例子中,术语“包括”、“基本由.....组成”和“由.....组成”中的任何一个都可用其它两个中的任意一个代替。已经使用的术语和表达用作说明而非限制性的术语,此类术语和表达的使用并不排除对所显示和所描述的特征或其部分的任何等价物,以及在要求权利的本技术范围内可进行各种改良。术语“一个”或“一种”表示一种或多种其修饰的元素(例如,“一种试剂”可表示一种或多种试剂),除非上下文清楚表示所描述的是元素之一或是一种以上的元素。本文所使用的术语“约”表示在基础参数的10%范围内的数值(即 $\pm 10\%$),在一列数值的开头处使用术语“约”表示修饰该列数值中的每个数值(即,“约1、2和3”是约1、约2和约3)。例如,“约100克”的重量能包含90克-110克的重量。因此,应理解,尽管通过代表性实施方式和任选的特征具体公开了本技术,但是本领域技术人员能对本文所公开内容进行改良和变化,应认为此类改良和变化落在本技术的范围内。

[0404] 本技术的实施方式在所附的权利要求中列出。

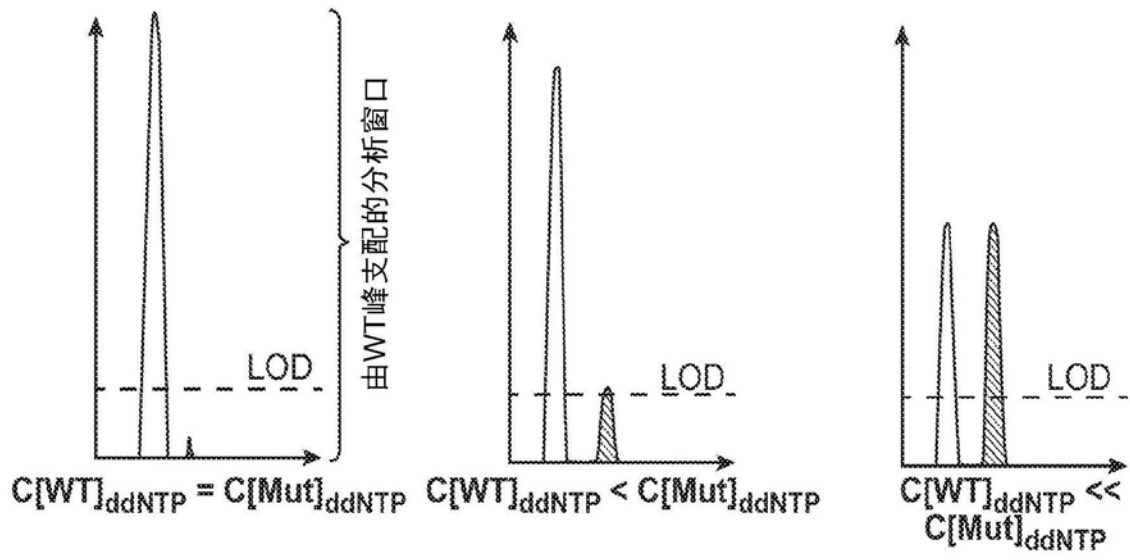


图1

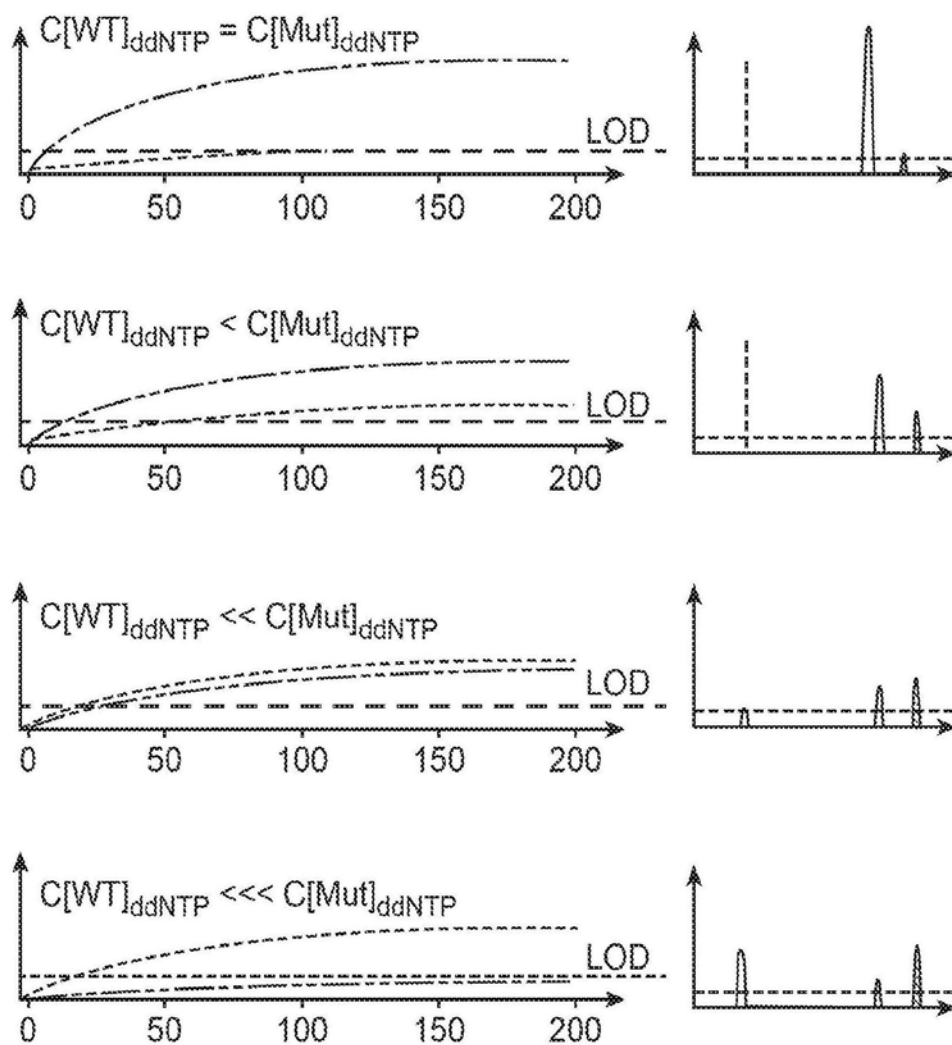


图2

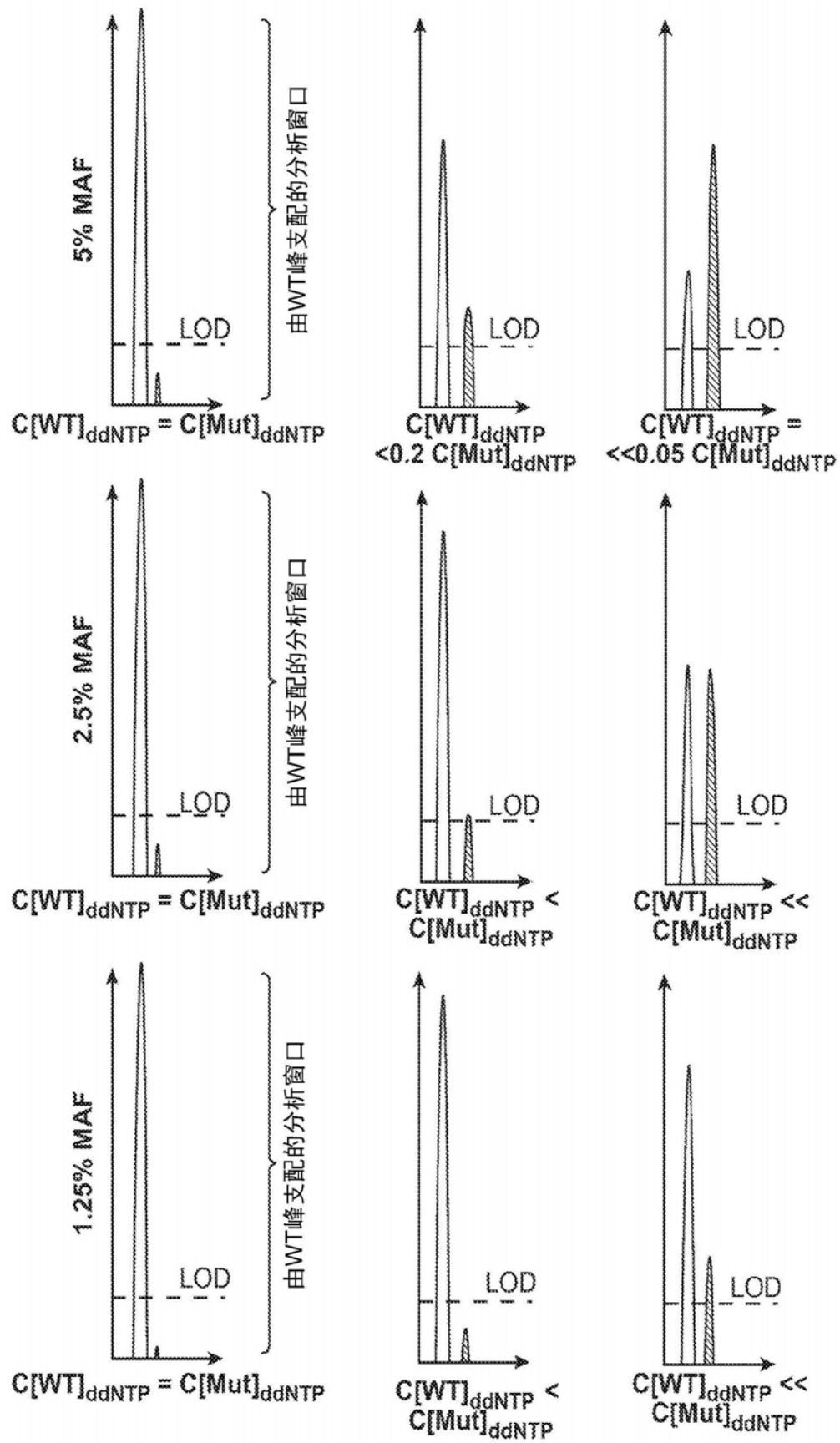


图3

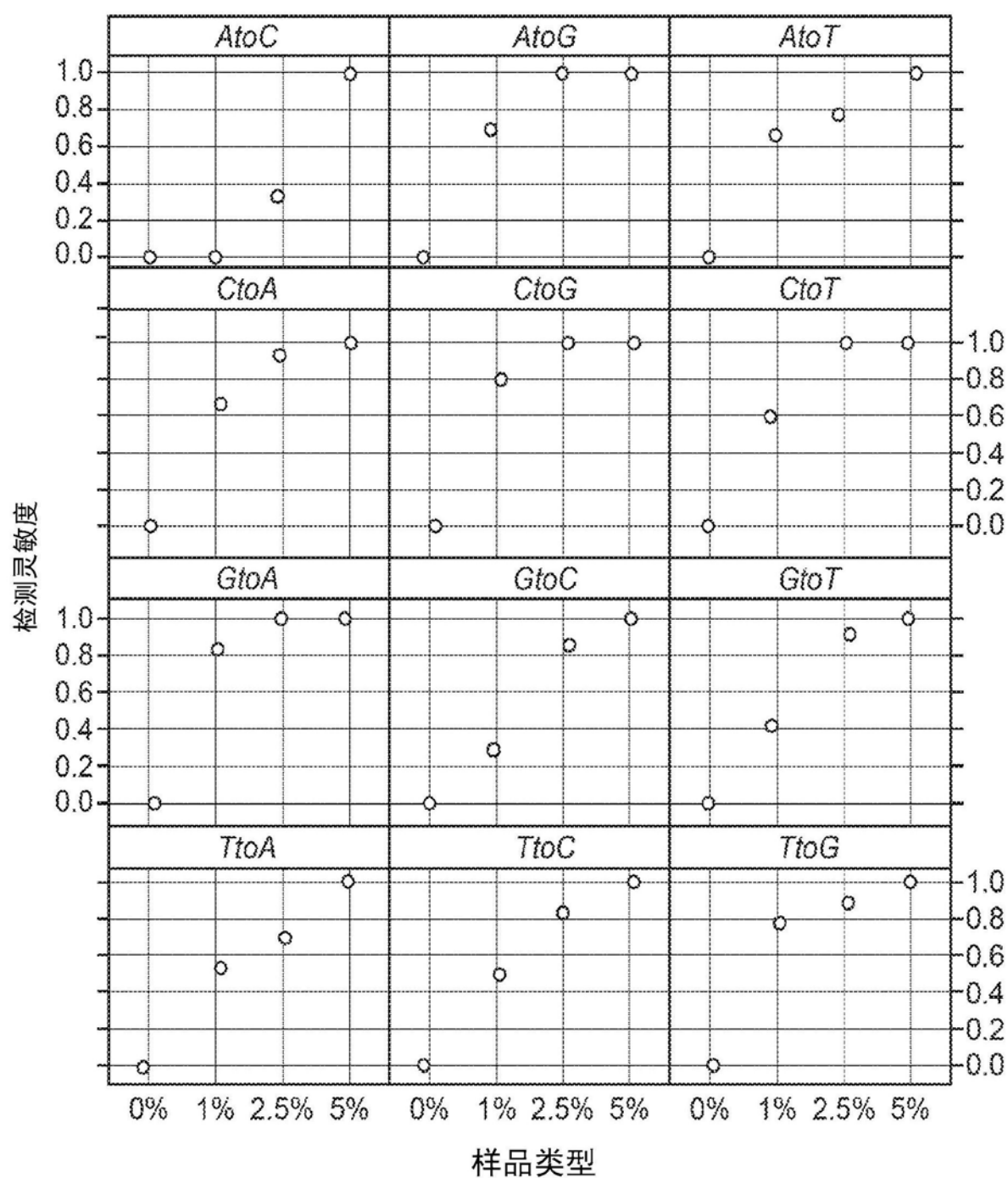


图4

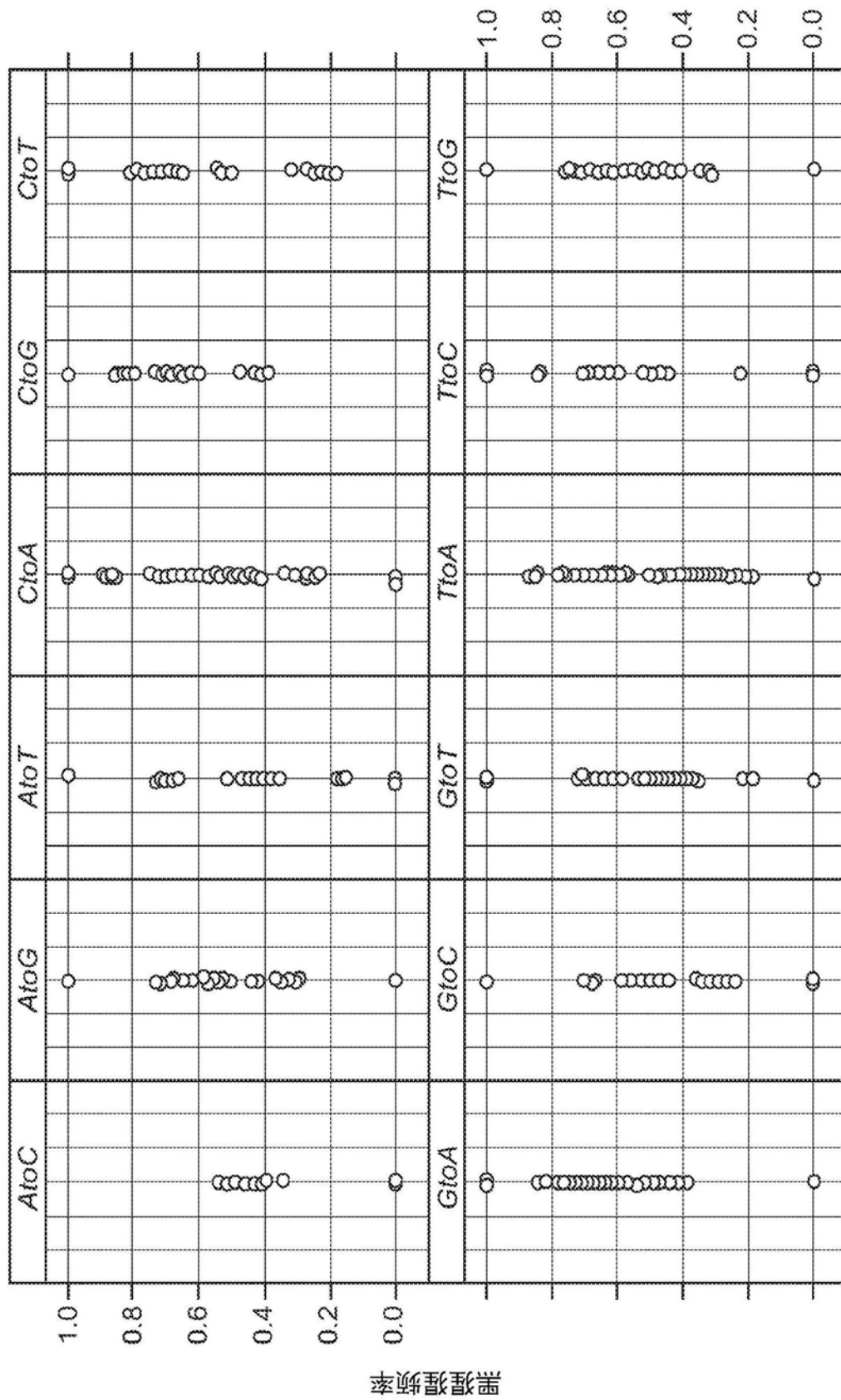


图5

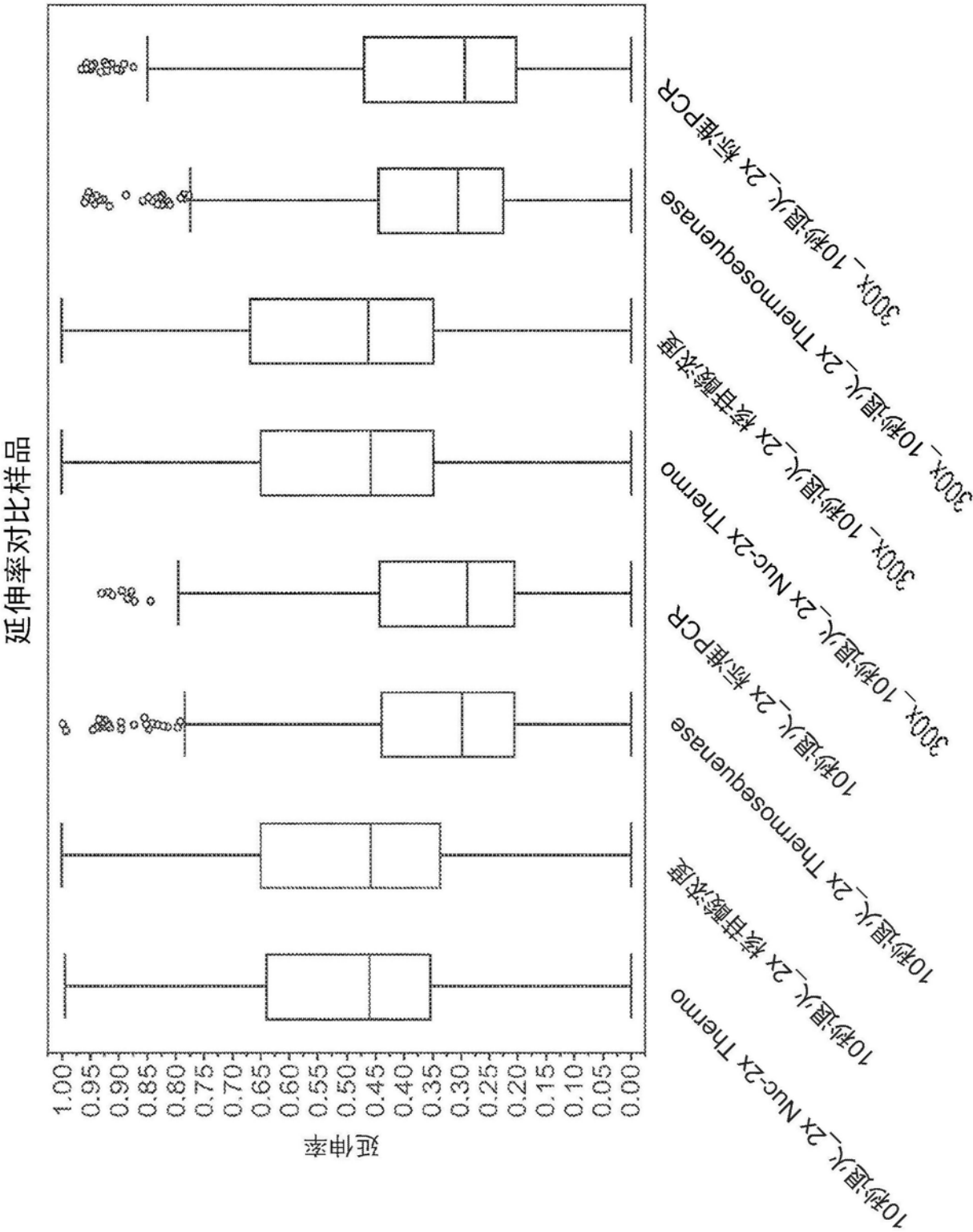


图6