

[19] 中华人民共和国国家知识产权局

[51] Int. Cl.
C12Q 1/68 (2006.01)



[12] 发明专利申请公布说明书

[21] 申请号 200610089526.7

[43] 公开日 2007年1月17日

[11] 公开号 CN 1896284A

[22] 申请日 2006.6.30

[21] 申请号 200610089526.7

[71] 申请人 博奥生物有限公司

地址 102206 北京市昌平区生命科学园路 18 号

共同申请人 清华大学

[72] 发明人 李彩霞 高华方 刘湘 蔡斌
张地 程京

[74] 专利代理机构 北京纪凯知识产权代理有限公司
代理人 关畅 任风华

权利要求书 3 页 说明书 20 页 附图 4 页

[54] 发明名称

一种鉴别等位基因类型的方法

[57] 摘要

本发明公开了一种鉴别等位基因的方法。该方法是以待测样品的 DNA 为模板，用待检测目的基因的一个或一个以上突变位点对应的等位基因的引物组和不具有 5'→3'外切核酸酶活性的 DNA 聚合酶，进行多重 PCR 扩增包括突变位点在内的待测目的基因的等位基因片段，将得到的多重 PCR 扩增产物与通用芯片杂交，根据杂交结果确定是哪一种等位基因；所述引物组包括一种通用引物，和用于扩增每种目的基因的每个突变位点对应的两种等位基因特异性引物和一种共用引物；所有的等位基因特异性引物中每种引物的 5'端均连接有不同的 Tag 序列；所述通用芯片包括若干种 Tag 探针，所述 Tag 探针与所述 Tag 序列相同，其中，每一种 Tag 探针只与多重 PCR 产生的互补序列杂交结合。

1、一种鉴别等位基因类型的方法，是以待测样品的基因组 DNA 为模板，用待检测目的基因的一个或一个以上突变位点对应的等位基因的引物组和不具有 5'→3' 外切核酸酶活性的 DNA 聚合酶，进行多重 PCR 扩增包括突变位点在内的待测目的基因的等位基因片段，将得到的多重 PCR 扩增产物与通用芯片杂交，根据杂交结果确定是哪种等位基因；

所述引物组包括一种通用引物，和用于扩增每种目的基因的每个突变位点对应的两种等位基因特异性引物和一种共用引物；每种等位基因特异性引物的 5' 端均连接有不同的 Tag 序列；

所述 Tag 序列可以和所述多重 PCR 产物具有的互补 Tag 序列杂交，所有的 Tag 序列的 T_m 值之差小于等于 5°C，相互之间以及与所有的引物之间没有交叉杂交，并且无发夹结构，与含有所述目的基因的物种的基因组的同源性较低；

所述通用芯片包括若干种 Tag 探针，所述 Tag 探针与所述 Tag 序列一一对应，一种 Tag 探针含有只与其相对应的 Tag 序列相同的寡核苷酸序列；其中，每一种 Tag 探针可以与多重 PCR 反应产生的互补的序列杂交。

2、根据权利要求 1 所述的方法，其特征在于：所述待检测目的基因至少有一种。

3、根据权利要求 1 所述的方法，其特征在于：所述共用引物和通用引物带有荧光分子或可通过化学发光或固体微粒进行标记检测的分子。

4、根据权利要求 1 所述的方法，其特征在于：所述多重 PCR 扩增在一管中进行，或分为多管进行多管多重 PCR 反应。

5、根据权利要求 1 所述的方法，其特征在于：所述目的基因的两个等位基因是由于单核苷酸的置换、插入或缺失产生，或多个碱基的插入或缺失产生；

所述用于扩增目的基因的两个等位基因的两种等位基因特异性引物中的一种的 3' 末端碱基与目的基因突变位点的野生型碱基相同或互补，另一种的 3' 末端碱基与目的基因突变位点的突变型碱基相同或互补。

6、根据权利要求 5 所述的方法，其特征在于：所述用于扩增目的基因的两个等位基因的两种等位基因特异性引物中引入人工错配碱基；所述人工错配碱基是天然的 A、T、C、G 四种碱基或其类似物。

7、根据权利要求1所述的方法，其特征在于：所述共用引物的5'末端还连接有通用引物序列；所述通用引物序列与所述等位基因特异性引物的 T_m 值之差小于等于 5°C ，无发夹结构和二聚体形成，与含有所述目的基因的物种的基因组的同源性较低，与所有的Tag序列无交叉杂交。

8、根据权利要求1所述的方法，其特征在于：所述Tag序列为单链寡核苷酸序列或肽核酸序列。

9、权利要求1至8中任一所述的方法，其特征在于：所述等位基因是由基因突变产生的。

10、根据权利要求9所述的方法，其特征在于：所述基因突变为遗传性耳聋基因突变；

所述待检测目的基因为 GenBank Accession Number 为 NC_000013 的 GJB2 (cox26)、GenBank Accession Number 为 NC_000001 的 GJB3(cox31)、GenBank Accession Number 为 NC_000007 的 SLC26A4(PDS)和 GenBank Accession Number 为 NC_001807 的 12S rRNA 基因；

所述引物组包括用于扩增 GenBank Accession Number 为 NC_000013 的 GJB2 (cox26) 基因的 35delG、167delT、176del16、235delC 和 299delAT 突变位点分别对应的两种等位基因特异性引物和一种共用引物，用于扩增 GenBank Accession Number 为 NC_000001 的 GJB3(cox31)基因的 538 C>T 和 547 G>A 突变位点分别对应的两种等位基因特异性引物和一种共用引物，用于扩增 GenBank Accession Number 为 NC_000007 的 SLC26A4 (PDS) 基因的 707 T>C 和 2168 A>G 和 IVS7-2 A>G 突变位点分别对应的两种等位基因特异性引物和一种共用引物，GenBank Accession Number 为 NC_001807 的 12S rRNA 基因 1555A>G 突变位点对应的两种等位基因特异性引物和一种共用引物；

用于扩增 GenBank Accession Number 为 NC_000013 的 GJB2 基因的 35delG 突变位点的两个等位基因特异性引物的 5' 端分别连接有具有序列表中序列 1 和序列 2 的核苷酸序列的 Tag1 和 Tag2；

用于扩增 GenBank Accession Number 为 NC_000013 的 GJB2 基因的 167delT 突变位点的两个等位基因特异性引物的 5' 端分别连接有具有序列表中序列 3 和序列 4 的核苷酸序列的 Tag3 和 Tag4；

用于扩增 GenBank Accession Number 为 NC_000013 的 GJB2 基因的 176del16 突变位点的两个等位基因特异性引物的 5' 端分别连接有具有序列表中序列 5 和序

列 6 的核苷酸序列的 Tag5 和 Tag6;

用于扩增 GenBank Accession Number 为 NC_000013 的 GJB2 基因的 235delC 突变位点的两个等位基因特异性引物的 5' 端分别连接有具有序列列表中序列 7 和序列 8 的核苷酸序列的 Tag7 和 Tag8;

用于扩增 GenBank Accession Number 为 NC_000013 的 GJB2 基因的 299delAT 突变位点的两个等位基因特异性引物的 5' 端分别连接有具有序列列表中序列 9 和序列 10 的核苷酸序列的 Tag9 和 Tag10;

用于扩增 GenBank Accession Number 为 NC_000001 的 GJB3 基因的 538 C>T 突变位点的两个等位基因特异性引物的 5' 端分别连接有具有序列列表中序列 11 和序列 12 的核苷酸序列的 Tag11 和 Tag12;

用于扩增 GenBank Accession Number 为 NC_000001 的 GJB3 基因的 547 G>A 突变位点的两个等位基因特异性引物的 5' 端分别连接有具有序列列表中序列 13 和序列 14 的核苷酸序列的 Tag13 和 Tag14;

用于扩增 GenBank Accession Number 为 NC_000007 的 SLC26A4 基因的 707 T>C 突变位点的两个等位基因特异性引物的 5' 端分别连接有具有序列列表中序列 15 和序列 16 的核苷酸序列的 Tag15 和 Tag16;

用于扩增 GenBank Accession Number 为 NC_000007 的 SLC26A4 基因的 2168 A>G 突变位点的两个等位基因特异性引物的 5' 端分别连接有具有序列列表中序列 19 和序列 20 的核苷酸序列的 Tag19 和 Tag20;

用于扩增 GenBank Accession Number 为 NC_000007 的 SLC26A4 基因的 IVS7-2 A>G 突变位点的两个等位基因特异性引物的 5' 端分别连接有具有序列列表中序列 21 和序列 22 的核苷酸序列的 Tag21 和 Tag22;

用于扩增 GenBank Accession Number 为 NC_001807 的 12S rRNA 基因 1555A>G 突变位点的两个等位基因特异性引物的 5' 端分别连接有具有序列列表中序列 17 和序列 18 的核苷酸序列的 Tag17 和 Tag18。

11、根据权利要求 10 所述的方法，其特征在于：所述通用芯片包括分别含有序列列表中序列 1 至序列 22 的 22 种 Tag 序列的 22 种 Tag 探针。

12、一种鉴别遗传性耳聋等位基因类型的芯片，其特征在于：所述芯片包括分别含有序列列表中序列 1 至序列 22 的核苷酸序列的 22 种 Tag 探针。

13、包括权利要求 12 所述的鉴别遗传性耳聋等位基因类型的芯片的试剂盒。

一种鉴别等位基因类型的方法

技术领域

本发明涉及基因分析领域中一种鉴别等位基因类型的方法。

背景技术

鉴于基因序列分析在生命科学研究和医学临床实践中具有重要的应用价值，人们已经建立了很多用于基因序列分析的方法。比较经典的方法有：限制性酶切片段长度多态性分析（Restriction fragment length polymorphism, RFLP）、单链构象多态性分析（Single Strand Conformation Polymorphism, SSCP）、核酸测序（Sequence Based Typing, SBT）、变性高效液相色谱分析（Denaturing High Performance Liquid Chromatography, DHPLC）、等位基因特性 PCR 技术（Allele-specific PCR, ASPCR）、序列特异性探针杂交法（Sequence Specific Oligonucleotide Probe, SSOP）等。但这些技术都存在一定的缺点，如成本太高、准确性不够或操作过程复杂等，而且它们还有一个共同的不足，即不能高通量、大规模的进行基因多态性分析。为了应对后基因组时代对高通量基因序列分析技术的要求，国际上正在开发一些新的技术。如：MALDI-TOF MS 质谱分析法（Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry）、连接酶检测反应技术（Ligase detection reaction, LDR）、微测序技术即单碱基延伸（SBE, single base chain extension）、BeadArray 技术、通用芯片（Universal array）技术和高密度基因芯片（High density microarray）技术等。

在上述方法中，等位基因特性 PCR（allele-specific PCR, ASPCR）技术是一种简便易行的基因序列分析方法，它是由 Newton 等于 1989 年建立，它是 PCR 技术应用的发展，也称为扩增阻碍突变系统（amplification refractory mutation system, ARMS），或 PCR-序列特异性引物（PCR-sequence specific primer, PCR-SSP）等。用于对基因序列中的已知突变或多态性进行检测，该法使用缺失 3'-5' 外切活性的 DNA 聚合酶，当引物的 3' 末端与模板不匹配时，该引物不能延伸，PCR 反应被阻滞。该方法简便实用，但通量较低，尤其是对多个基因位点进行同时检测时比较麻烦。为了提高检测通量，研究者们采取了多种策略，Rebecca Robert 等借鉴多重 PCR 原理进行检测，将多个位点的野生型和突变型引物分两管多重扩

增，最后凝胶电泳检测，两条泳道检测，一条为野生型产物，另一条为突变型产物，不同位点间通过扩增产物的长度进行区分。Gomez-Llorente C 等采用单管多重扩增结合自动化毛细管电泳检测的方法，不同位点间通过片段长度进行区分，同一位点的两条等位基因特异性引物标记两种荧光，通过颜色区分突变型和野生型。Michele Boniotta 等将单管多重扩增与溶解曲线分析结合进行多位点检测，在等位基因特异性引物 5' 末端连接 GC 尾巴从而区分两等位基因的 T_m 值，结合 SYBR Green I 进行荧光定量分析。Shannon Eaker 等将 ARMS 与芯片检测结合，首先进行多重 ARMS 扩增，产物标记后与等位基因特异性寡核苷酸芯片杂交分型。

通用芯片 (Universal array) 技术是近年来发展起来的一种用于基因序列分析的高通量技术，是由 Barany 等 (U.S. Pat. No. 6506594) 首先提出的，将液相连接酶检测反应 (ligase detection reaction, LDR) 与微阵列技术相结合，该方法可对低丰度的基因点突变进行非常灵敏的检测。将连接酶反应 (LDR) 下游共用探针的 3' 末端连接荧光素分子，在上游等位基因特异探针的 5' 末端连接不同的 cZip-code 序列，这种 cZip-code 序列是人工设计，并经过严格的筛选，能够与通用芯片上的 Zip-code 序列完全互补，每一种 cZip-code 和 Zip-code 序列的组合对应一种靶基因分子中可能出现的突变点或 SNP；当上游等位基因特异性探针与靶标互补时，连接酶将上下游探针连接起来，然后进行芯片杂交，通过不同 Zip-code 位置和荧光信号进行核酸序列差异检测。这种方法的灵敏度很高，可以准确检测出野生序列中混杂的 1% 甚至更低量的单碱基突变。如果将 cZip-code 连接在其它的特异探针上，就可以用设计好的同一种芯片检测其它任意的靶基因序列，从而使芯片具有了通用性。通用芯片与液相酶促反应结合检测基因多态性 (或突变) 大大克服了等位基因特异性寡核苷酸芯片特异性低的缺点。

在我国，耳聋人群是最大的残疾人群体，约有 2 千万患者。其中，50% 与遗传有关。由于导致耳聋的基因突变点位于多个基因上，目前临床主要用基因测序的方法进行基因序列分析，不但操作繁琐、通量低，而且价格较高。

发明内容

本发明的目的是提供一种鉴别等位基因类型的方法。

本发明所提供的鉴别等位基因类型的方法，是以待测样品的基因组 DNA 为模板，用待检测目的基因的一个或一个以上突变位点对应的等位基因的引物组和不具有 5'→3' 外切核酸酶活性的 DNA 聚合酶，进行多重 PCR 扩增包括突变位点在内的待测目的基因的等位基因片段，将得到的多重 PCR 扩增产物与通用芯片杂交，

根据杂交结果确定是哪一种等位基因；

所述引物组包括一种通用引物，和用于扩增每种目的基因的每个突变位点对应的两种等位基因特异性引物和一种共用引物；每种等位基因特异性引物的 5' 端均连接有不同的 Tag 序列；

所述 Tag 序列可以和所述多重 PCR 产物具有的互补 Tag 序列杂交，所有的 Tag 序列的 T_m 值之差小于等于 5°C （有相似的 T_m 值），相互之间以及与所有的引物之间没有交叉杂交，并且无发夹结构，与含有所述目的基因的物种的基因组同源性较低（连续匹配或相同的碱基数小于 10 个）；

所述通用芯片包括若干种 Tag 探针，所述 Tag 探针与所述 Tag 序列一一对应，一种 Tag 探针含有只与其相对应的 Tag 序列相同的寡核苷酸序列；其中，每一种 Tag 探针可以与多重 PCR 反应产生的互补的序列杂交。

其中，所述 Tag 序列具体可以通过生物信息学的方法人工设计 20—24bp 序列，或者取自不同于含有所述目的基因的物种的基因组 DNA 序列，如目的基因来源于人，可取自细菌等其它物种的基因组 DNA 序列，如结核分支杆菌等。

所述待检测目的基因至少有一种。

其中，所述共用引物和通用引物 5' 末端带有荧光分子或可通过化学发光或固体微粒进行标记检测的分子，如生物素。

所述多重 PCR 扩增在一管中进行，或分为多管进行多管多重 PCR 反应。

所述目的基因的两个等位基因是由于单核苷酸的置换、插入或缺失产生，或多个碱基的插入或缺失产生；

所述用于扩增目的基因的两个等位基因的一种等位基因特异性引物中的一种的 3' 末端碱基与目的基因突变位点的野生型碱基相同或互补，另一种的 3' 末端碱基与目的基因突变位点的突变型碱基相同或互补。

为提高检测的特异性，所述用于扩增目的基因的两个等位基因的一种等位基因特异性引物中引入人工错配碱基；所述人工错配碱基是天然的 A、T、C、G 四种碱基或其类似物。当这种新引入的人工突变遇到引物中原有的天然的碱基突变时将产生协同作用，使得引物和目的基因序列的结合更加不稳定，而当引物中仅有人工突变时则可以和目的基因序列结合，这样使得匹配引物（仅有人工突变的引物）的扩增产量与不匹配引物（既有人工突变，又有天然突变的引物）的扩增产量差异变大，从而特异的区分靶分子中的单碱基差异。

使用多重不对称 PCR 扩增获得用于杂交的靶标，每个位点的荧光标记的共用

引物与等位基因特异性引物的量比例为 10~25: 1, 为了提高单链 PCR 产物的量, 共用引物的 5' 末端还连接有通用引物序列, 多重扩增时, 另外加入一种荧光标记的通用引物, 浓度高于等位基因特异性引物, 从而通过引物浓度的差异获得大量的单链产物, 用于杂交。所述通用引物序列与所述等位基因特异性引物的 T_m 值之差小于等于 5°C, 无发夹结构和二聚体形成; 与含有所述目的基因的物种的基因组同源性较低(连续匹配或相同的碱基数小于 10 个); 与所有的 Tag 序列无交叉杂交。

所述 Tag 序列可为单链寡核苷酸序列或肽核酸序列。

上述方法中, 所述等位基因是由基因突变产生的。

基因突变是指由于 DNA 碱基的置换、插入或缺失而引起的基因结构碱基序列差异。

所述基因突变具体可为遗传性耳聋基因突变; 其中, 所述待检测目的基因为 GenBank Accession Number 为 NC_000013 的 GJB2 (cox26), GenBank Accession Number 为 NC_000001 的 GJB3 (cox31), GenBank Accession Number 为 NC_000007 的 SLC26A4 (PDS) 和 GenBank Accession Number 为 NC_001807 的 12S rRNA 基因;

所述引物组包括用于扩增 GenBank Accession Number 为 NC_000013 的 GJB2 (cox26) 基因的 35delG、167delT、176del16、235delC 和 299delAT 突变位点分别对应的两种等位基因特异性引物和一种共用引物, 用于扩增 GenBank Accession Number 为 NC_000001 的 GJB3 (cox31) 基因的 538 C>T 和 547 G>A 突变位点分别对应的两种等位基因特异性引物和一种共用引物, 用于扩增 GenBank Accession Number 为 NC_000007 的 SLC26A4 (PDS) 基因的 707 T>C 和 2168 A>G 和 IVS7-2 A>G 突变位点分别对应的两种等位基因特异性引物和一种共用引物, GenBank Accession Number 为 NC_001807 的 12S rRNA 基因 1555A>G 突变位点对应的两种等位基因特异性引物和一种共用引物;

用于扩增 GenBank Accession Number 为 NC_000013 的 GJB2 基因的 35delG 突变位点的两个等位基因特异性引物的 5' 端分别连接有具有序列表中序列 1 和序列 2 的核苷酸序列的 Tag1 和 Tag2;

用于扩增 GenBank Accession Number 为 NC_000013 的 GJB2 基因的 167delT 突变位点的两个等位基因特异性引物的 5' 端分别连接有具有序列表中序列 3 和序列 4 的核苷酸序列的 Tag3 和 Tag4;

用于扩增 GenBank Accession Number 为 NC_000013 的 GJB2 基因的 176del16

突变位点的两个等位基因特异性引物的 5' 端分别连接有具有序列表中序列 5 和序列 6 的核苷酸序列的 Tag5 和 Tag6;

用于扩增 GenBank Accession Number 为 NC_000013 的 GJB2 基因的 235delC 突变位点的两个等位基因特异性引物的 5' 端分别连接有具有序列表中序列 7 和序列 8 的核苷酸序列的 Tag7 和 Tag8;

用于扩增 GenBank Accession Number 为 NC_000013 的 GJB2 基因的 299delAT 突变位点的两个等位基因特异性引物的 5' 端分别连接有具有序列表中序列 9 和序列 10 的核苷酸序列的 Tag9 和 Tag10;

用于扩增 GenBank Accession Number 为 NC_000001 的 GJB3 基因的 538 C>T 突变位点的两个等位基因特异性引物的 5' 端分别连接有具有序列表中序列 11 和序列 12 的核苷酸序列的 Tag11 和 Tag12;

用于扩增 GenBank Accession Number 为 NC_000001 的 GJB3 基因的 547 G>A 突变位点的两个等位基因特异性引物的 5' 端分别连接有具有序列表中序列 13 和序列 14 的核苷酸序列的 Tag13 和 Tag14;

用于扩增 GenBank Accession Number 为 NC_000007 的 SLC26A4 基因的 707 T>C 突变位点的两个等位基因特异性引物的 5' 端分别连接有具有序列表中序列 15 和序列 16 的核苷酸序列的 Tag15 和 Tag16;

用于扩增 GenBank Accession Number 为 NC_000007 的 SLC26A4 基因的 2168 A>G 突变位点的两个等位基因特异性引物的 5' 端分别连接有具有序列表中序列 19 和序列 20 的核苷酸序列的 Tag19 和 Tag20;

用于扩增 GenBank Accession Number 为 NC_000007 的 SLC26A4 基因的 IVS7-2 A>G 突变位点的两个等位基因特异性引物的 5' 端分别连接有具有序列表中序列 21 和序列 22 的核苷酸序列的 Tag21 和 Tag22;

用于扩增 GenBank Accession Number 为 NC_001807 的 12S rRNA 基因 1555A>G 突变位点的两个等位基因特异性引物的 5' 端分别连接有具有序列表中序列 17 和序列 18 的核苷酸序列的 Tag17 和 Tag18。

所述通用芯片包括分别含有序列表中序列 1 至序列 22 的 22 种 Tag 序列的 22 种 Tag 探针。

所述 22 种 Tag 探针的 5' 端进行了氨基化, 所述 Tag 序列的 5' 端连接有 15 个 T。

上述方法可以结合微型全分析系统制成自动化检测装置。

本发明还提供了一种鉴别遗传性耳聋等位基因类型的芯片。

本发明所提供的鉴别遗传性耳聋等位基因类型的芯片，分别含有序列表中序列 1 至序列 22 的核苷酸序列的 22 种 Tag 探针。

所述 22 种 Tag 探针的 5' 端进行了氨基化，所述 Tag 序列的 5' 端连接有 15 个 T。

包括该鉴别遗传性耳聋等位基因类型的芯片的试剂盒也属于本发明的保护范围。

本发明将多重等位基因特异性 PCR 技术与通用芯片技术相结合，建立了一种高通量、低成本、易于操作的基因序列分析方法，其反应原理见图 1 所示：首先，在一只反应管中用基因位点特异性引物（包括等位基因特异性引物和共用引物）进行多重等位基因特异性 PCR 反应，然后把 PCR 产物与通用芯片进行杂交反应，最后通过芯片扫描进行结果检测。本发明的方法在每个位点的共用引物 5' 末端荧光标记，并且连接通用引物以提高多重 PCR 的单链产量，两种等位基因特异性引物 5' 末端连接不同的标签（Tag）序列，扩增产物与通用芯片上对应的标签（Tag）序列杂交，最后通过芯片上 Tag 序列的位置和杂交信号值进行基因序列分析。

本发明中的多重等位基因特异性 PCR 不同于一般意义上的简单多重 PCR。首先，其等位基因特异性引物的 3' 末端碱基有两种形式，一种是末端碱基和待检测的基因位点野生型（SNP 或突变）匹配，一种是末端碱基和待检测的基因位点突变型（SNP 或突变）匹配，二者相互对照，这增加了反应的特异性和准确性；其次，在引物的末端或者连接 Tag 标签序列用于对待检测的基因位点进行编码，或者连接有标记分子用于最终的检测；另外，为了增强反应的特异性可以在引物序列中引入人工突变点，为了提高检测敏感性可以在共用引物序列末端引入通用引物，在 PCR 体系中加入荧光标记的通用引物（与共用引物末端的通用引物序列相同），待前两轮 PCR 反应结束后，产生通用引物的互补序列，荧光标记的通用引物与之结合继续扩增，增加了 PCR 产物中单链 DNA 产量，从而增强杂交信号。

本发明中的通用芯片也不同于一般意义上的固定有序列特异性探针的基因芯片，在普通芯片上固定的探针是针对基因的特定位置进行检测的基因片段，检测不同的目标基因需要设计不同的探针，制备不同的芯片。而本发明中所指的通用芯片是一种用来对预先筛选出的一套序列特异性基因片段（Tag 序列）进行识别的基因芯片，芯片上固定的探针不是针对特定的基因序列，而是针对特定的 Tag 序列。Tag 序列可以用来对不同的基因位点、对不同基因的不同位点、对不同物种的

不同基因位点进行编码（相当于一个条码），对于不同的检测目的可以用相同的一套 Tag 序列，这样对于不同检测目的，只要在芯片上固定一套可以识别这套 Tag 序列的基因片段就可完成检测任务。因此该通用芯片是一种真正意义上的“通用”芯片，尽管检测目的不同，但都可以用同样的一张芯片来进行检测，检测过程是一个对编码在不同基因位点上的 Tag 序列进行解码的过程。这种芯片是与普通意义上的核酸序列特异性基因芯片完全不同的一种方便实用的生物芯片。

本发明的方法快速、简便、更具可操作性、易于实现自动化操作。可用于基因突变检测、基因多态性分析等方面，适用于临床疾病突变检测、药物基因组学分析、法医学鉴定等基因分析领域。

附图说明

图 1 为本发明方法的反应原理示意图

图 2 为通用芯片验证实验结果

图 3 为通用芯片点阵排布示意图

图 4 为用图 3 所示的通用芯片对耳聋病人的临床样品进行实际检测的结果

图 5 为分别用磁珠和荧光标记的通用芯片检测结果

具体实施方式

下述实验方法，如无特别说明，均为常规方法。

实施例 1、用本发明的通用芯片和方法检测已知遗传性耳聋基因突变分型的病例样品

1. 临床样品来源及其 DNA 提取

已知遗传性耳聋基因突变分型的病例样品由中国人民解放军总医院耳鼻喉科提供。使用 Wizard® Genomic DNA Purification Kit (Promega, Madison, USA) 提取全血中的基因组 DNA。

2. 制备多重 PCR 引物及通用芯片的探针

检测如表 1 所示的四个目的基因共 11 个突变位点的多重 PCR 引物及探针如下：

(1) 引物

多重 PCR 引物序列见表 1。

突变形式中的“del”表示突变位点为缺失突变，如 35delG 表示 GenBank Accession Number 为 NC_000013 的 GJB2 基因的 35 位碱基 G 缺失；突变形式中的“>”表示突变位点为置换突变，如 538C>T 表示 GenBank Accession Number 为 NC_000001 的 GJB3 的 538 位碱基由 C 突变为 T。

引物名称中后缀为 WT 的引物代表扩增包括突变位点在内的野生型等位基因特异性引物，引物名称中后缀为 MU 的引物代表扩增包括突变位点在内的突变型等位基因特异性引物，引物名称中后缀为 C 的引物代表扩增包括该突变位点在内的野生型和突变型等位基因片段的共用引物，每个突变位点的两种等位基因特异性引物分别与该突变位点的共用引物配对，如 35delG-WT 和 35delG-C 配对，35delG-MU 和 35delG-C 配对。

引物序列中的 Tag1 至 Tag22 分别具有序列表中序列 1 至序列 22 的核苷酸序列；共用引物序列中的 UP 表示序列为 GCACGCTATCACGTTAGAC 的通用引物序列。共用引物的 5' 端修饰有荧光标记 TAMRA。

为提高检测特异性，部分等位基因特异性引物中引入人工错配碱基（下划线示）。

等位基因特异性引物 538C>T-WT、547G>A-WT、707T>C-WT 和 2168A>G-WT 的 3' 末端碱基（粗体示）与目的基因突变位点的野生型碱基相同，等位基因特异性引物 538C>T-MU、547G>A-MU、707T>C-MU 和 2168A>G-MU 的 3' 末端碱基（粗体示）与目的基因突变位点的突变型碱基相同，等位基因特异性引物 IVS7-2A>G-WT 和 1555A>G-WT 的 3' 末端碱基（粗体示）与目的基因突变位点的野生型碱基互补，等位基因特异性引物 1555A>G-MU 和 IVS7-2A>G-MU 的 3' 末端碱基（粗体示）与目的基因突变位点的突变型碱基互补。

(2) 探针

通用芯片是由与多重 PCR 产物杂交的 22 种 Tag 探针、QC(点样的阳性质控)、BC(点样的阴性质控)、NC(杂交的阴性质控)和 PC(杂交的阳性质控)组成的阵列。QC 是一端带有 Hex 标记，另一端氨基修饰的寡核苷酸探针，用于观察芯片点样和固定的效率，其序列为 NH₂-TTTTTTTTTTTTTTAGAGTGCTTGGTGCCATAAC-HEX；BC 为 50%的 DMSO，点于 QC 之后，用于质控点样过程中有无样品残留污染；NC 是一段氨基修饰的寡核苷酸探针，与杂交液中的所有待检测序列不会杂交，用于观察有无非特异杂交，其序列为 NH₂-TTTTTTTTTTTTTTGCAACCACCACCGGAGG；PC 是一段氨基修饰的寡核苷酸探针，可以与杂交液中添加的荧光标记的互补序列(c-PC)杂交，用于杂交过程的质控，其序列为 NH₂-TTTTTTTTTTTTTTGGTATCGCGACCGCATCCCAATCT。

通用芯片上的 Tag 探针序列结构通式为：NH₂-TTTTTTTTTTTTTT-TagX，其中 X 为 1 至 22 中的任何一个自然数。如 Tag1 探针的序列结构为

NH₂-TTTTTTTTTTTTTT-Tag1，Tag22 探针的序列结构为 NH₂-TTTTTTTTTTTTTT-Tag22。即探针的 5' 端氨基修饰，氨基旁边连接 poly-T15，然后是分别具有序列中序列 1 至 22 的核苷酸序列的 Tag1 至 Tag22。其中，Tag1 至 Tag22 的核苷酸序列分别与引物序列中的 Tag1 至 Tag22 的核苷酸序列相同。

表 1. 检测的突变位点、引物及 Tag 序列

基因及基因 Bank 编号	突变形式	引物名称	引物序列 (5' →3')	Tag 探针中的 Tag 序列(5' →3')
GJB2 NC_000013	35delG	35delG-WT	Tag1-TGTTTGTTCACACCCCGCAG	GTTACTGCTACGCGTGCTACGT
		35delG-MU	Tag2-TGTTTGTTCACACCCCGCAG	CATGAGCAAGCTGTCTAAGGCG
		35delG-C	TAMRA-UP-GCATGCTTGCTTACCCAGAC	
	167delT	167delT-WT	Tag3-CGACTTTGTCTGCAACACCCTG	CGACGAGCTGCCGCGCAAGAT
		167delT-MU	Tag4-ACTTTGTCTGCAACACCCG	TATCGCGACCGCATCCAATCT
	176del16	176del16-WT	Tag5-CCAGGCTGCAAGAACGTGTG	GCTCGAAGAGGGCTACAGATC
		176del16-MU	Tag6-ACCCTGCAGCCAGCTACG	TTCCCGTCCGTCATCGTCAAG
	235delC	235delC-WT	Tag7-ATCCGGCTATGGGCCTG	GATCGGCGGTGAAGCGAAAGG
		235delC-MU	Tag8-ATCCGGCTATGGGCCTG	GATGGTGATCTCGCGGTGCG
	299delAT	299delAT-WT	Tag9-TGGCTACCGGAGACATGA	TGTGCGCCCGAGTTCGGTATC
		299delAT-MU	Tag10-CGTGGCCTACCGGAGACGA	TTGATCCCATCGAAGGACGATG
		GJB2-C	TAMRA-UP-GAGCCTTCGATGCGGACC	
GJB3 NC_000001	538C>T	538C>T-WT	Tag11-GTGGACTGTACATTGGCC	TGATGCGTCTGGGACGTGCCTG
		538C>T-MU	Tag12-GTGGACTGTACATTGGCT	CAGAGCATCAACGACGCAGGA
	547G>A	547G>A-WT	Tag13-ACATTGCCAGACCTACCG	ACGATCAACGC GGAGACACAG
		547G>A-MU	Tag14-TACATTGCCAGACCTACCA	ACGAGACACGCAACGAGACAG
		GJB3-C	TAMRA-UP-TCGAGGCTTGTCTTGTGC	
SLC26A4 NC_000007	707T>C	707T>C-WT	Tag15-CAAGTGTGGTCTCACAGCT	TTGAAAGCCTACACGCGAGCG
		707T>C-MU	Tag16-AAGTGTGGTCTCACAGCC	CAAGCAGAGCTATGGTTCGCTG
		707T>C-C	TAMRA-UP-GAGGTCTCAGTCTCAAATCC	
	2168A>G	2168A>G-WT	Tag19-GACACATCTTTTGTCTGCTCA	GTTAGGGTCGGCCAACTCTCC
		2168A>G-MU	Tag20-ACATCTTTTGTCTGCTCA	GACAAAGGTCTGCCAGCACCA
		2168A>G-C	TAMRA-UP-CAAGGTTTTCCAGATTGCTGAG	
	IVS7-2A>G	IVS7-2A>G-WT	Tag21-AATGGCAGTAGCAATTATCGACT	TGCAACACGCTAGGATCTCCTC
		IVS7-2A>G-MU	Tag22-TGGCAGTAGCAATTATCGTCC	TGCACTTCTCGGTAGGACGCGA
	IVS7-2A>G-C	TAMRA-UP-CGTGTAGCAGCAGGAAGTAT		
12S rRNA NC_001807	1555A>G	1555A>G-WT	Tag17-ACTTACCATGTTACGACTAGT	GTCAGTATCGGTCGCTTACG
		1555A>G-MU	Tag18-CACTTACCATGTTACGACTTGC	CCATACTACGCAACTGTGCA

		1555A>G-C	TAMRA-UP-CCCTGATGAAGGCTACAAAG	
		通用引物(UP)	TAMRA- GCACGCTATCACGTTAGAC	

所有的引物与探针由上海 Invitrogen 公司 (Invitrogen Co., Shanghai, China) 合成与纯化。

将探针固定到醛基化修饰的玻璃片上。所有的 Tag 探针用 50% DMSO 溶解, 终浓度为 15 μ M, 每个点重复五次点制到玻片上。图 3 为通用芯片点阵排布示意图, 实际点阵中每个点延水平方向连续重复 5 个点。其中包括检测 GJB2 基因上的 35delG、167delT、176del116、235delC、299delAT 突变位点的 Tag1 至 Tag10 探针; 检测 GJB3 基因上的 538C>T、547G>A 突变位点的 Tag11 至 Tag14 探针; 检测 SLC26A4 基因上的 707T>C、2168A>G、IVS7-2A>G 突变位点的 Tag15、Tag16、Tag19、Tag20、Tag21 和 Tag22 探针; 检测线粒体 12S rRNA 上的 1555A>G 突变点的 Tag17 和 Tag18 探针。图 3 中, 35W 代表 Tag1 探针, 35M 代表 Tag2 探针, 167W 代表 Tag3 探针, 167M 代表 Tag4 探针, 176W 代表 Tag5 探针, 176M 代表 Tag6 探针, 235W 代表 Tag7 探针, 235M 代表 Tag8 探针, 299W 代表 Tag9 探针, 299M 代表 Tag10 探针, 538W 代表 Tag11 探针, 538M 代表 Tag12 探针, 547W 代表 Tag13 探针, 547M 代表 Tag14 探针, 707W 代表 Tag15 探针, 707M 代表 Tag16 探针, 1555W 代表 Tag17 探针, 1555M 代表 Tag18 探针, 2168W 代表 Tag19 探针, 2168M 代表 Tag20 探针, IVS7-2W 代表 Tag21 探针, IVS7-2M 代表 Tag22 探针。

3. 多重等位基因特异性 PCR

该多重 PCR 扩增中用到了通用引物 UP, 共用引物序列 5' 末端连接通用引物, 在 PCR 体系中加入荧光标记的通用引物 (与共用引物末端的通用引物序列相同), 待前两轮 PCR 反应结束后, 产生通用引物的互补序列, 荧光标记的通用引物与之结合继续扩增, 增加了 PCR 产物中单链 DNA 产量, 从而增强杂交信号

分别以各个病例样品的全血基因组 DNA 为模板, 进行多重 PCR。为防止某些引物间的相互作用, 多重 PCR 在两管中进行。其中, 引物 547G>A-WT, 547G>A-MU, GJB3-C, 2168A>G-WT, 2168A>G-MU 和 2168A>G-C, IVS7-2 A>G-WT, IVS7-2 A>G-MU 和 IVS7-2 A>G-C 在同一管中进行, 表 1 中的其它引物在另一管扩增。25 μ l 扩增体系, 包括 0.2 mM dNTPs, 1 \times Qiagen PCR buffer, 添加 MgCl₂ 至 2mM, pH 8.7, 1 单位 HotStarTaq DNA Polymerase (Qiagen, Hilden, Germany), 和 100ng 基因组 DNA, 各位点的检测引物, 其中共用引物浓度高于等位基因特异性引物, 通用引物分别加入两管中。25 μ l 扩增体系中, 35delG-WT 0.02 μ M, 35delG-MU 0.02 μ M,

35delG-C 0.4 μ M, 167delT-WT 0.03 μ M, 167delT-MU 0.03 μ M, 176del16-WT 0.03 μ M, 176del16-MU 0.03 μ M, 235delC-WT 0.02 μ M, 235delC-MU 0.02 μ M, 299delAT-WT 0.02 μ M, 299delAT-MU 0.03 μ M, GJB2-C 0.6 μ M, 538C>T-WT 0.02 μ M, 538C>T-MU 0.02 μ M, 547G>A-WT 0.02 μ M, 547G>A-MU 0.02 μ M, GJB3-C 0.4 μ M, 707T>C-WT 0.02 μ M, 707T>C-MU 0.02 μ M, 707T>C-C 0.3 μ M, 2168A>G-WT 0.03 μ M, 2168A>G-MU 0.03 μ M, 2168A>G-C 0.4 μ M, IVS7-2A>G-W 0.03 μ M, IVS7-2A>G-MU 0.03 μ M, IVS7-2A>G-C 0.4 μ M, 1555A>G-WT 0.008 μ M, 1555A>G-MU 0.008 μ M, 1555A>G-C 0.2 μ M, 通用引物 1 μ M。

扩增参数为:

先 95 $^{\circ}$ C 15 分钟;

然后 94 $^{\circ}$ C 30 秒, 0.5 $^{\circ}$ C/秒降至 56 $^{\circ}$ C, 56 $^{\circ}$ C 30 秒, 0.2 $^{\circ}$ C/秒升至 70 $^{\circ}$ C, 70 $^{\circ}$ C 45 秒, 10 个循环;

再 90 $^{\circ}$ C 30 秒, 0.5 $^{\circ}$ C/秒降至 56 $^{\circ}$ C, 56 $^{\circ}$ C 30 秒, 0.2 $^{\circ}$ C/秒升至 70 $^{\circ}$ C, 70 $^{\circ}$ C 45 秒, 22 个循环;

最后 60 $^{\circ}$ C 10 分钟;

4 $^{\circ}$ C 保存。

4. 通用芯片杂交

两管 PCR 产物混合, 取 10 μ l 混合物加到 20 μ l 杂交缓冲液中 (6 \times SSC, 5 \times Denhardt's reagent, 25% 甲酰胺, 0.1%SDS, 5nM c-PC (与芯片 PC 互补的序列, 5'端 TAMRA 标记))。98 $^{\circ}$ C 变性 5 分钟, 冰浴, 杂交混合物加入到两个相邻的点阵中作为重复实验。然后将芯片放置 50 $^{\circ}$ C 水浴中杂交 1 小时。取出玻片分别在两种洗液中 42 $^{\circ}$ C 各洗 2 min, 洗液 I: 0.3 \times SSC/0.1%SDS, 洗液 II: 0.06 \times SSC。最后, 玻片稍离心甩干。

为验证通用芯片的特异性, 人工合成荧光标记的 Tag1 至 Tag22 的互补核苷酸序列 (cTag1 至 cTag22), 用杂交缓冲液溶解至 5nM, 分别与芯片进行杂交, 杂交与洗片条件同上。

5. 数据分析

使用 ScanArrayTM express microarray scanner (PerkinElmer, MA, USA) 扫描玻片, 扫描图见图 4 所示。Laser power 和 photomultiplier tube (PMT) power 分别为 90% 和 70%。使用 GenePix 4.0 software (Axon Instruments, CA, USA) 提取数据。取每个位点的 AMSI 值 (absolute median signal intensities), 即

荧光信号中位值减去背景值。每个位点的 AMSI 值应该不低于 1000。为了排除引物二聚体导致的假阳性信号，每个位点的 AMSI 值与 PCR 阴性对照（不加模板）相应位点 AMSI 的比值应该大于 10。满足这两个条件后，该位点的信号才可以算作阳性。对于通用芯片验证实验，ratio 值被用来判断 Tag 探针的特异性。特异性的标准为：ratio 应该大于 10；该 Tag 探针的 AMSI 值至少为 1000。

$$\text{Ratio} = (\text{AMSI})_{\text{期望阳性信号}} / (\text{AMSI})_{\text{阴性信号最大值}}$$

通用芯片验证实验结果见图 2，T 代表 Tag，结果表明 Tag1 至 Tag22 探针的 AMSI 值均大于 1000，ratio 值均大于 10。说明该通用芯片具有较高的特异性。图 2 中，T1 至 T22 分别表示 Tag1 至 Tag22 探针。

图 4 显示的是用图 3 所示的通用芯片对耳聋病人的临床样品进行实际检测的结果。芯片中左边的一列探针均为野生型检测探针，右边一列探针为突变型检测探针，每种探针延水平方向重复点样五次。因此，对于已知野生型（WT）样品基因组 DNA（gDNA）的杂交结果左边的野生型检测探针全部出现杂交信号，而右边的突变型检测探针没有杂交信号（工作正常）；gDNA 235delC MU、gDNA 299delAT MU、gDNA 1555A>G MU 和 gDNA IVS7-2A>G MU 分别为已知遗传性耳聋基因突变类型 235delC MU、299delAT MU、1555A>G MU 和 IVS7-2A>G MU 的病例样品的基因组 DNA，代表了四种突变情况，杂交图中相应的突变检测探针已经给出了相应的正确检测信号。由图 4 结果可知，所有位点检测结果都正确无误，特异性很好。

实施例 2、用本发明的通用芯片和方法检测已知遗传性耳聋基因突变分型的病例样品

为了简化检测流程，降低该方法对仪器的依赖程度，设计了用磁珠进行最终标记的实验。这样检测结果就可以用 CCD 照相机拍摄或用低倍显微镜直接目测，这使得复杂的基因检测变的非常容易和简便。

选择 235delC 和 299delAT 两个位点进行此实验。将相关引物和探针上的荧光标记换为 biotin 标记。PCR 条件、杂交和清洗过程同实施例 1 所述。使用 Streptavidin 包被的 MyOne™ Dynalbeads (Dynal invitrogen, Oslo, Norway) 显示杂交结果。先将 Streptavidin 包被的 Dynalbeads 按照说明书方法进行预处理。磁珠混合液加到玻片上，常温结合 10 分钟后，使用磁铁去除非特异结合的磁珠，离心甩干。最后使用光学显微镜观察结果，用 CCD 照相机拍摄图像。反应结果如图 5 所示，点阵形式为 4×10 格式（见图中左上角排布格局），每种探针重复点 5 个点。所用的引物和探针序列见表 1 所示。可见结果显示非常清楚，与图

中右边的荧光检测结果一致。图 5 中，A 为通用芯片点阵排布示意图，实际点阵中每个点延水平方向连续重复 5 个点。其中包括检测 GJB2 基因上的 235delC 和 299delAT 突变位点的 Tag7 至 Tag10 探针。图 4 A 中，235W 代表 Tag7 探针，235M 代表 Tag8 探针，299W 代表 Tag9 探针，299M 代表 Tag10 探针，QC（点样阳性质控），BC（点样阴性质控），NC（杂交阴性质控）和 PC（杂交阳性质控）。图 5 中 B 是以磁珠显示的通用芯片检测结果，图 5 中 C 是以荧光信号显示的通用芯片检测结果。WT 为已知野生型基因组 DNA 样品，235delC MU、299delAT MU 分别为已知遗传性耳聋基因突变类型 235delC MU 和 299delAT MU 的病例样品的基因组 DNA。

序列表

<160>22

<210>1

<211>22

<212>DNA

<213>人工序列

<220>

<223>

<400>1

gttactgcta cgcgtgctac gt

22

<210>2

<211>22

<212>DNA

<213>人工序列

<220>

<223>

<400>2

catgagcaag ctgtctaagg cg

22

<210>3

<211>21

<212>DNA

<213>人工序列

<220>

<223>

<400>3

cgacgagctg ccgcgcaaga t

21

<210>4

<211>21

<212>DNA

<213>人工序列

<220>

<223>

<400>4

tatcgcgacc gcatccaatc t

21

<210>5

<211>21

<212>DNA

<213>人工序列

<220>

<223>

<400>5

gctcgaagag ggctacagat c

21

<210>6

<211>22

<212>DNA

<213>人工序列

<220>

<223>

<400>6

ttcccgtccg tcatcgctca ag

22

<210>7

<211>21

<212>DNA

<213>人工序列

<220>

<223>

<400>7

gatcggcggg gaagcgaag g 21

<210>8

<211>21

<212>DNA

<213>人工序列

<220>

<223>

<400>8

gatggtgatc tcgcggtgc g 21

<210>9

<211>22

<212>DNA

<213>人工序列

<220>

<223>

<400>9

tgtgcgcccg agttcggta tc 22

<210>10

<211>22

<212>DNA

<213>人工序列

<220>

<223>

<400>10

ttgatcccat cgaaggacga tg 22

<210>11

<211>22

<212>DNA

<213>人工序列

<220>

<223>

<400>11

tgatgcgtct gggacgtgcc tg

22

<210>12

<211>21

<212>DNA

<213>人工序列

<220>

<223>

<400>12

cagagcatca acgacgcagg a

21

<210>13

<211>21

<212>DNA

<213>人工序列

<220>

<223>

<400>13

acgatcaacg cggagacaca g

21

<210>14

<211>21

<212>DNA

<213>人工序列

<220>

<223>

<400>14
acgagacacg caacgagaca g 21

<210>15
<211>21
<212>DNA
<213>人工序列

<220>
<223>

<400>15
ttgaaagcct acacgcgagc g 21

<210>16
<211>22
<212>DNA
<213>人工序列

<220>
<223>

<400>16
caagcagagc tatggttcgc tg 22

<210>17
<211>22
<212>DNA
<213>人工序列

<220>
<223>

<400>17
gtcagtatcg cgttcgctta cg 22

<210>18	
<211>21	
<212>DNA	
<213>人工序列	
<220>	
<223>	
<400>18	
ccatactcac gcaactgtgc a	21
<210>19	
<211>22	
<212>DNA	
<213>人工序列	
<220>	
<223>	
<400>19	
gttagggtcg gccaaactct cc	22
<210>20	
<211>22	
<212>DNA	
<213>人工序列	
<220>	
<223>	
<400>20	
gacaaaggtc tgcccagcac ca	22
<210>21	
<211>22	
<212>DNA	
<213>人工序列	

<220>

<223>

<400>21

tgcaacacgc taggatctcc tc

22

<210>22

<211>21

<212>DNA

<213>人工序列

<220>

<223>

<400>22

tgcaacttctc ggtaggcagc ga

22

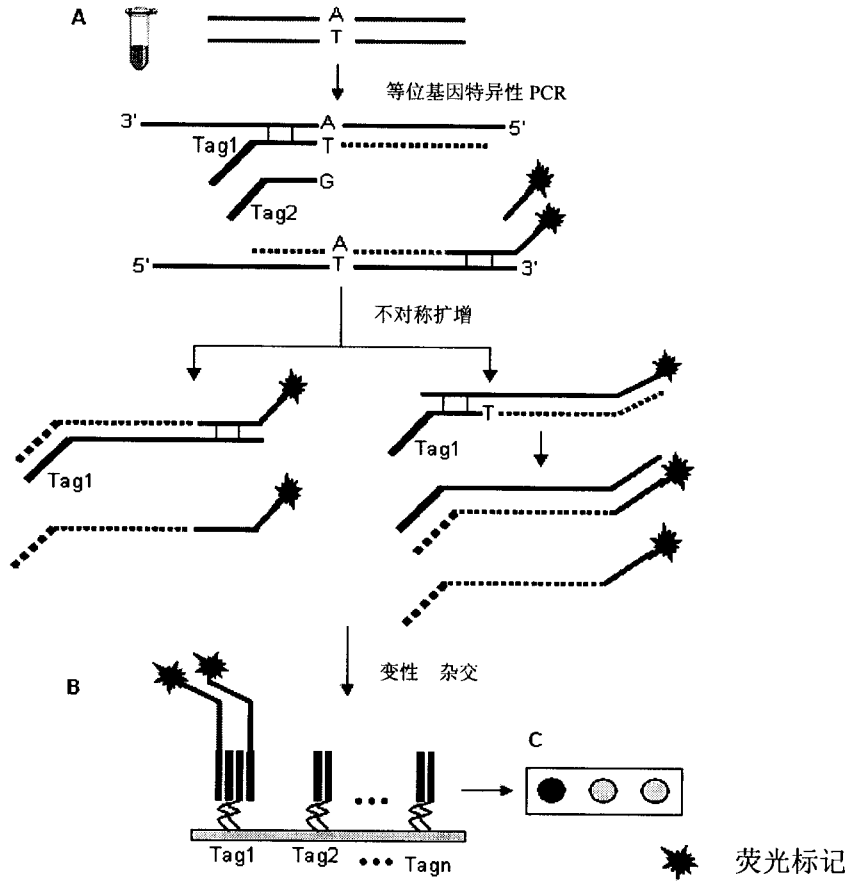


图 1

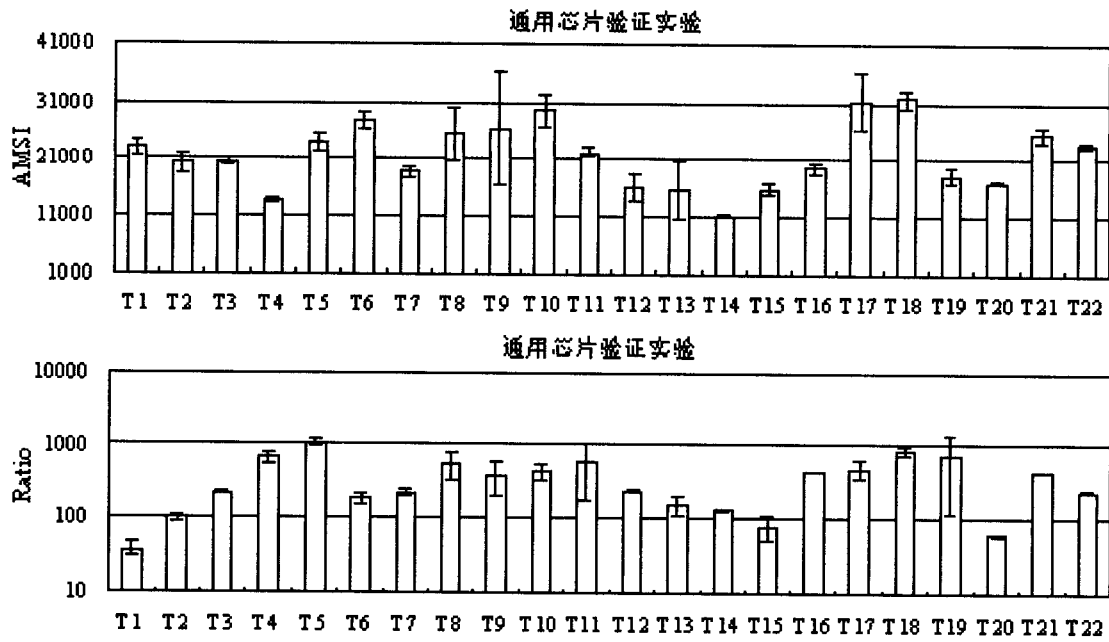


图 2



图 3

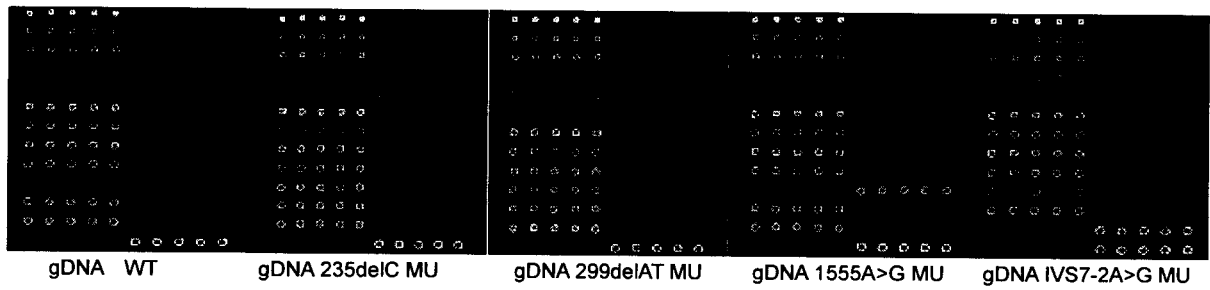


图 4

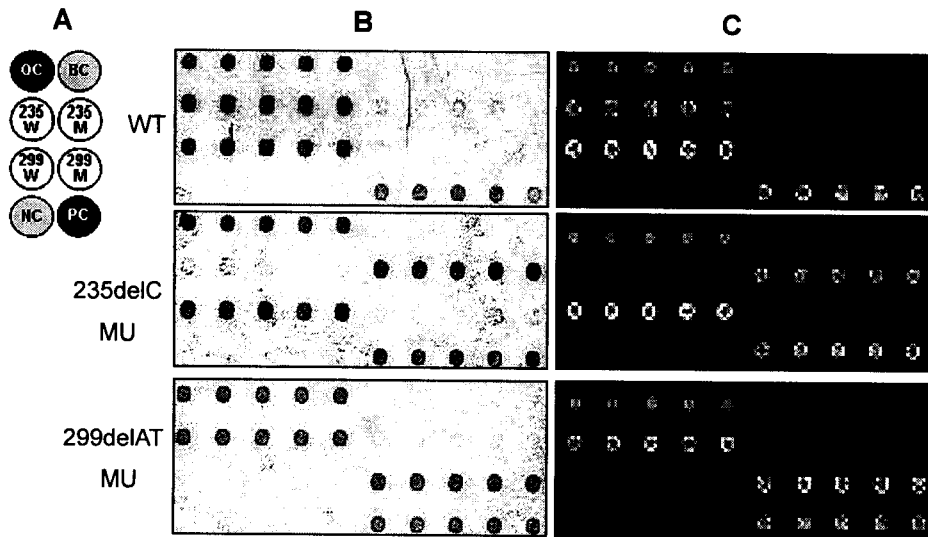


图 5