

巢式多重 PCR 高通量测序文库制备方法及其试剂盒

技术领域

本发明涉及文库制备技术领域，具体涉及一种巢式多重 PCR 高通量测序文库制备方法及其试剂盒。

背景技术

随着测序技术的发展，对基因组候选区段的重测序需求日益增加，人们对序列的关注超过了少数的 SNP，候选区段的范围可能在 5kb-10M 之间。使用传统 sanger 法或目标区域杂交捕获测序价格昂贵，使用目标区域捕获测序很好地解决了这一难题。目标区域捕获技术可大致分为两种：一种是基于杂交的捕获测序技术，另一种是基于多重 PCR 的捕获技术。两者通过多重探针或者引物对感兴趣的基因区域一次性捕获，结合高通量测序技术，多样本同时测序，得到目标区域的序列信息。相对于全基因组测序，目标捕获测序技术在大样本量筛查的同时极大地降低成本。但前者的实验流程繁琐，探针成本较高，限制了其在临床上的应用，后者实验操作简单，灵活性强，适用于孟德尔遗传性疾病的筛查和诊断、GWAS 候选区段重测序、QTL 定位区段重测序、精准医疗研究与应用等领域。

高通量 SNP 检测服务结合多重 PCR 和高通量测序技术，对需要检测的位点设计特异性引物，在单管内进行多重 PCR 扩增，不同的样本以不同的标签（barcode）引物区分。混合样本后，在测序平台上，对扩增子进行测序，测序结果使用生物信息学方法，区分不同样本，最终获得每个位点的 SNP 信息。这种方法适用于不同目的遗传学研究，例如疾病基因组研究、肿瘤基因组研究、疾病与基因的关联研究、临床分子诊断等，在植物基因组研究中，可用于 QTL 定位及分子育种，非常适合大规模样本的 SNP 分析。

虽然多重 PCR 实验操作简单、单个检测成本很低，但其在实验前期需要对多对引物进行反复测试优化，费时费力。特别是在超高重的 PCR 中，引物序列的复杂性使得引物很容易形成引物二聚体。引物二聚体的形成会急剧消耗 PCR 反应体系中的原料，导致 PCR 很快达到平台期；形成的引物二聚体在后续的测序中也会被测序，形成无效数据，影响数据的利用效率。最严重的是那些容易形成引物二聚体的引物，会严重影响该引物对应的目标扩增区域的扩增效率，导致该目标测序深度低，最终影响整个扩增体系的均一性。此外，引物的特异性也极大地影响了多重扩增的性能。

在多重扩增中随着引物对数的增加，引物二聚体的形成不可避免，很多公司通过额外的酶处理对过多的引物二聚体进行消除，如 Ampliseq 公司，先采用一步特异性扩增，然后通过酶消化掉引物二聚体和扩增子中的特异性引物序列，最后对产物进行文库制备，整个过程繁琐。后来 Paragon 公司改良了这个方法，发明了 Clean Plex 二聚体消除技术，采用特定的酶只消除多重 PCR 过程中的引物二聚体，然后采用通用引物对消化后的产物进行通用扩增，但是对于引物的非特异性扩增没有更好的解决方法。

概而言之，现有技术的缺陷体现在：（1）现有方法本身无法减少引物二聚体的形成，其方法是在引物二聚体产生之后采用酶消化的方式进行，采用事后去除引物二聚体的方式，其多重 PCR 过程中产生的引物二聚体会极大地影响 PCR 扩增的效率和均一性；（2）现有方法需额外采用一步消化步骤，操作比较繁琐；（3）采用常规的两条引物扩增，对于一些特异性不好的区域很难设计特异的引物；（4）现有方法对于连续区域扩增需要在多个 PCR 管中分别进行。

发明内容

本申请提供一种巢式多重 PCR 高通量测序文库制备方法及其试剂盒，有效避免多重 PCR 扩增过程中产生的二聚体，并且采用巢式扩增方法大大增加引物的特异性，提高数据的利用率，还能有效抑制重叠扩增子引物之间的互相扩增。

根据第一方面，一种实施例中提供一种巢式多重 PCR 高通量测序文库制备方法，该方法包括：

第一轮 PCR 扩增：使用正向引物和反向引物对目标区域进行扩增，其中上述正向引物是与上述目标区域结合的正向特异性序列，上述反向引物包括 3'端的与上述目标区域结合的反向特异性序列和 5'端的第一通用测序序列；

扩增产物的纯化：对上述第一轮 PCR 扩增的产物进行纯化；

第二轮 PCR 扩增：使用巢式引物、第一标签引物和第二通用引物对纯化后的产物进行扩增，其中上述巢式引物位于上述正向引物的下游，上述巢式引物包括 5'端的第二通用测序序列和 3'端的特异性序列，上述第一标签引物 3'端序列和上述反向引物 5'端的第一通用测序序列部分或全部序列相同，且上述第一标签引物还包括第一标签序列，上述第二通用引物 3'端序列和上述巢式引物 5'端的第二通用测序序列部分或者全部序列相同。

在优选实施例中，上述第二通用引物是第二标签引物，该第二标签引物 3'端序

列和上述巢式引物 5'端的第二通用测序序列部分或者全部序列相同，且上述第二标签引物还包括第二标签序列。

根据第一方面，另一种实施例中提供一种巢式多重 PCR 高通量测序文库制备方法，该方法包括：

第一轮 PCR 扩增：使用正向引物、反向引物和第一标签引物对目标区域进行扩增，其中上述正向引物是与上述目标区域结合的正向特异性序列，上述反向引物包括 3'端的与上述目标区域结合的反向特异性序列和 5'端的第一通用测序序列，上述第一标签引物 3'端序列和上述反向引物 5'端的第一通用测序序列部分或全部序列相同，且上述第一标签引物还包括第一标签序列；

扩增产物的纯化：对上述第一轮 PCR 扩增的产物进行纯化；

第二轮 PCR 扩增：使用巢式引物、第一通用引物和第二通用引物对纯化后的产物进行扩增，其中上述巢式引物位于上述正向引物的下游，上述巢式引物包括 5'端的第二通用测序序列和 3'端的特异性序列，上述第一通用引物 3'端序列和上述第一标签引物 5'端部分或全部序列相同，上述第二通用引物 3'端序列和上述巢式引物 5'端的第二通用测序序列部分或者全部序列相同。

在优选实施例中，上述第二通用引物是第二标签引物，该第二标签引物 3'端序列和上述巢式引物 5'端的第二通用测序序列部分或者全部序列相同，且上述第二标签引物还包括第二标签序列。

在优选实施例中，上述反向引物上还包括分子标签序列。

在优选实施例中，上述分子标签序列是 8-24bp 的随机序列。

在优选实施例中，上述方法对多个连续的目标区域进行扩增，每个目标区域都有对应的正向引物、反向引物和巢式引物，并且相邻扩增子区域之间形成扩增重叠。

在优选实施例中，上述正向引物、反向引物和巢式引物均是由多条分别靶向不同的目标区域的引物组成的引物池。

在优选实施例中，每个目标区域的正向引物和相邻区域的正向引物、巢式引物和相邻区域的巢式引物、反向引物和相邻区域的反向引物扩增方向相反。

在优选实施例中，上述第一轮 PCR 扩增和/或第二轮 PCR 扩增的循环数均是 2-30 个循环。

在优选实施例中，上述第一标签序列和/或第二标签序列的长度均是 8-15bp。

根据第二方面，一种实施例中提供一种巢式多重 PCR 高通量测序文库制备试剂

盒，该试剂盒包括：

第一轮 PCR 扩增引物，其包括正向引物和反向引物，其中上述正向引物是与目标区域结合的正向特异性序列，上述反向引物包括 3'端的与上述目标区域结合的反向特异性序列和 5'端的第一通用测序序列，上述第一轮 PCR 扩增引物用于对目标区域进行第一轮 PCR 扩增；

第二轮 PCR 扩增引物，其包括巢式引物、第一标签引物和第二通用引物，其中上述巢式引物位于上述正向引物的下游，上述巢式引物包括 5'端的第二通用测序序列和 3'端的特异性序列，上述第一标签引物 3'端序列和上述反向引物 5'端的第一通用测序序列部分或全部序列相同，且上述第一标签引物还包括第一标签序列，上述第二通用引物 3'端序列和上述巢式引物 5'端的第二通用测序序列部分或者全部序列相同，上述第二轮 PCR 扩增引物用于对纯化后的上述第一轮 PCR 扩增的产物进行第二轮 PCR 扩增。

在优选实施例中，上述第二通用引物是第二标签引物，该第二标签引物 3'端序列和上述巢式引物 5'端的第二通用测序序列部分或者全部序列相同，且上述第二标签引物还包括第二标签序列。

根据第二方面，另一种实施例中提供一种巢式多重 PCR 高通量测序文库制备试剂盒，该试剂盒包括：

第一轮 PCR 扩增引物，其包括正向引物、反向引物和第一标签引物，其中上述正向引物是与上述目标区域结合的正向特异性序列，上述反向引物包括 3'端的与上述目标区域结合的反向特异性序列和 5'端的第一通用测序序列，上述第一标签引物 3'端序列和上述反向引物 5'端的第一通用测序序列部分或全部序列相同，且上述第一标签引物还包括第一标签序列，上述第一轮 PCR 扩增引物用于对目标区域进行第一轮 PCR 扩增；

第二轮 PCR 扩增引物，其包括巢式引物、第一通用引物和第二通用引物，其中上述巢式引物位于上述正向引物的下游，上述巢式引物包括 5'端的第二通用测序序列和 3'端的特异性序列，上述第一通用引物 3'端序列和上述第一标签引物 5'端部分或全部序列相同，上述第二通用引物 3'端序列和上述巢式引物 5'端的第二通用测序序列部分或者全部序列相同，上述第二轮 PCR 扩增引物用于对纯化后的上述第一轮 PCR 扩增的产物进行第二轮 PCR 扩增。

在优选实施例中，上述第二通用引物是第二标签引物，该第二标签引物 3'端序

列和上述巢式引物 5'端的第二通用测序序列部分或者全部序列相同，且上述第二标签引物还包括第二标签序列。

在优选实施例中，上述反向引物上还包括分子标签序列。

在优选实施例中，上述分子标签序列是 8-24bp 的随机序列。

在优选实施例中，上述正向引物、反向引物和巢式引物均是由多条分别靶向多个连续的目标区域的引物组成的引物池，并且相邻扩增子区域之间形成扩增重叠。

在优选实施例中，每个目标区域的正向引物和相邻区域的正向引物、巢式引物和相邻区域的巢式引物、反向引物和相邻区域的反向引物扩增方向相反。

在优选实施例中，上述第一标签序列和/或第二标签序列的长度均是 8-15bp。

本发明的巢式多重 PCR 高通量测序文库制备方法及其试剂盒，能够有效降低多重 PCR 扩增过程中产生的引物二聚体，三条引物的扩增方式，有效增加产物的特异性。此外，优选实施例能够有效抑制连续区域扩增中产生的无效扩增片段；采用分子标签标记特异性引物，能够对原始模板进行唯一标记，可以对原始模板进行溯源和错误矫正。

附图说明

图 1 为现有技术中两步 PCR 扩增流程示意图；

图 2 为本发明实施例中巢式多重 PCR 扩增流程示意图；

图 3 为本发明实施例中巢式多重 PCR 连续区域扩增示意图；

图 4 为本发明实施例中常规巢式 PCR 引物设计示意图 (a)，引入分子标签的巢式 PCR 引物设计示意图 (b)，连续区域巢式 PCR 引物设计示意图 (c)；

图 5 为本发明实施例中不同样本标签引入方式示意图；

图 6 为本发明实施例中不同样本标签引入方式示意图；

图 7 为本发明实施例中不同扩增子均一性实验结果图；

图 8 为本发明实施例中巢式多重 PCR 得到的文库电泳图。

具体实施方式

下面通过具体实施方式结合附图对本发明作进一步详细说明。在以下的实施方式中，很多细节描述是为了使得本申请能被更好的理解。然而，本领域技术人员可以毫不费力的认识到，其中部分特征在不同情况下是可以省略的，或者可以由其他材料、方法所替代。

另外，说明书中所描述的特点、操作或者特征可以以任意适当的方式结合形成

各种实施方式。同时，方法描述中的各步骤或者动作也可以按照本领域技术人员所能显而易见的方式进行顺序调换或调整。因此，说明书和附图中的各种顺序只是为了清楚描述某一个实施例，并不意味着是必须的顺序，除非另有说明其中某个顺序是必须遵循的。

本文中为特征所编序号本身，例如“第一”、“第二”等，仅用于区分所描述的对象，不具有任何顺序或技术含义。

如图 1 所示，现有技术中两步 PCR 扩增流程显示，在特异性扩增的过程中，两条特异性引物（图 1 中特异性引物 1 和特异性引物 2）分别带上两个测序接头的部分序列，这两种引物在第一轮 PCR 扩增过程中形成引物二聚体，这种引物二聚体在后续的第二轮 PCR 通用扩增中会被放大，得到大量的二聚体产物，影响测序数据。

本发明的方法，解决了特异性 PCR 扩增过程中产生的引物二聚体产物在随后的通用 PCR 扩增过程中放大的问题。

如图 2 所示，本发明的方法，在特异性扩增中，一条特异性引物（下称反向引物）带有第一通用测序序列，另一条特异性引物（下称正向引物）不带测序序列。反向引物和反向引物形成引物二聚体（图 2 中引物二聚体 3），由于通用测序序列的存在，这类引物在后续的扩增中会形成颈环结构，而不会被扩增放大，正向引物和正向引物、正向引物和反向引物之间形成的引物二聚体（图 2 中引物二聚体 1 和引物二聚体 2）无法被后续的通用引物利用而无法进行扩增，可以有效降低第一轮 PCR 扩增中引物二聚体在最终产物中的比例。

如图 2 所示，采用半巢式扩增，能够提高目标区域扩增的特异性。具体而言，在第二轮 PCR 扩增中，在正向引物下游设计巢式引物，该巢式引物带有第二通用测序序列，用这条引物和第一标签引物对得到的产物进行第二轮 PCR 扩增，得到最终的测序文库。该方法采用巢式扩增，在第一轮 PCR 扩增产物的基础上再进行一步特异性选择，提高最终产物的特异性。该步包括加入巢式引物和第二通用引物进行 PCR 扩增，扩增过程中只有巢式引物之间容易形成引物二聚体（图 2 中引物二聚体 4），巢式引物之间形成的引物二聚体在后续的扩增中会形成颈环结构，而不会被扩增放大，因此也能明显抑制引物二聚体的形成。

如图 3 和图 4 所示，对于连续区域，特殊的引物扩增策略，避免重叠扩增子之间重叠引物互相扩增。具体而言，每个区域有三条引物组成，正向引物、巢式引物、反向引物，设计正向引物和下一个相邻的正向引物之间形成扩增重叠，巢式引物和

下一个相邻的巢式引物之间形成扩增重叠，反向引物和下一个反向引物之间形成扩增重叠（图 4）。正向引物和正向引物会产生产物（图 3 中产物 4），该产物会被第二步巢式扩增中的巢式引物利用，进行扩增，但是由于巢式引物都带有第二通用测序序列，在后续的扩增中会形成颈环结构而被抑制；反向引物和反向引物会产生产物（图 3 中产物 2），由于反向引物都带有第一通用测序序列，该产物在后续扩增中会形成颈环结构而被抑制扩增（图 3），因此这些重叠的产物都不会被扩增放大。

如图 4a 所示，对于每一个扩增子，设计三条引物：正向引物、巢式引物、反向引物进行两步 PCR 扩增。巢式引物位于正向引物的下游引物，两者可以重叠或相隔，巢式引物 5'端有一段测序通用序列，3'端为特异性序列，3'端的序列设计在需要扩增的目标区域上游；正向引物设计在巢式引物的上游，其 5'端可以加或者不加（优选不加）测序通用序列；反向引物 5'端有一段测序通用序列，3'端为特异性序列，3'端的序列设计在需要扩增的目标区域的下游。

如图 4b 所示，如需要对低频突变进行检测，需要在反向引物中的通用序列和特异序列之间加上分子标签序列，该分子标签可以是 8-24bp 的随机序列，可以产生 4^{8-24} 种随机标签分子，用来标记原始的 DNA 分子。当分子标签的数量远远大于模板的数量时，每一个模板都会带上一个唯一的分子标签（Unique molecular index），通过这个分子标签可以对原始的 DNA 模板进行追溯，用来去掉扩增的重复和测序错误。引入分子标签到引物上会急剧增加引物之间二聚体生成的可能性，但是我们的方法形成的引物二聚体在后续的过程中会被有效抑制。

如图 4c 所示，对于连续区域的扩增，设计多对引物对连续区域进行覆盖，相邻扩增子区域有重叠（除去扩增子中引物序列外，上一个扩增的 3'端在下一个扩增子的 5'端的下游）。每一个区域由三条引物扩增，正向引物、巢式引物、反向引物。每个区域的正向引物和相邻区域的正向引物、巢式引物和相邻区域的巢式引物、反向引物和相邻区域的反向引物相向设计（扩增方向相反）。例如，第一个扩增子从 5'至 3'方向依次为正向引物 1、巢式引物 1 和反向引物 1，那么第二个扩增子从 5'至 3'方向依次为反向引物 2、巢式引物 2 和正向引物 2，第三个扩增子从 5'至 3'方向依次为正向引物 3、巢式引物 3 和方向引物 3，以此类推，直到所有扩增子能够覆盖整个连续区域，第二个扩增子的反向引物 2 的 3'端在反向引物 1 的 3'端的上游，第三个扩增子的巢式引物 3 的 3'端在巢式引物 2 的 3'端的上游，通过该种设计将连续区域的所有序列都能够扩增到相应的扩增子中。

如图 5a 所示,在第一轮 PCR 扩增中,采用正向引物和反向引物进行 2-30 轮 PCR 扩增。得到的产物进行纯化,去掉多余的引物和其他离子,然后进行第二轮 PCR 扩增。第二轮 PCR 扩增中用巢式引物、第一标签引物和第二通用引物进行进行扩增 2-30 个循环,其中第一标签引物的 3'端序列和反向引物 5'端的部分或全部序列相同,标签引物中间有 8-15bp 的标签序列,该标签序列在不同测序平台有不同的固定序列,用于区分不同样本以便后续的多样本混合测序;第二通用引物 3'端序列和巢式引物 5'端的部分或者全部序列相同。在扩增的第一个循环,巢式引物和第一标签引物进行扩增,在后续的循环中巢式引物、第一标签引物和第二通用引物一起对目标文库进行扩增,最终得到扩增子文库。

如图 5b 所示,对于双样本标签序列测序,第二轮 PCR 扩增中用巢式引物、第一标签引物和第二标签引物扩增 2-30 个循环,其中第一标签引物 3'端的序列和反向引物 5'端的部分或全部序列相同,第二标签引物 3'端的序列和巢式引物的 5'端部分或者全部序列相同,标签引物中间有 8-15bp 的标签序列,该标签序列在不同测序平台有不同的固定序列,用于区分不同样本以便后续的多样本混合测序。在扩增的第一个循环,巢式引物和第一标签引物进行扩增,在后续的循环中巢式引物、第一标签引物和第二标签引物一起对目标文库进行扩增,最终得到双样本标签的扩增子文库。

如图 6a 所示,对于第一轮 PCR 扩增,除了加正向引物和反向引物,还可以加入第一标签引物,第一标签引物中间有 8-15nt 的标签序列,在第一轮 PCR 扩增就可以对样本进行区分,用于区分不同样本以便后续的多样本混合测序,该第一标签引物的 3'端和反向引物的 5'端部分或者全部重叠。在扩增的第一个循环,正向引物和反向引物进行扩增,在后续的循环正向引物、反向引物 and 第一标签引物一起进行扩增。第二轮 PCR 扩增中用巢式引物、第一通用引物和第二通用引物进行进行扩增 2-30 个循环,第一通用引物的 3'端和第一标签引物的 5'端部分或者全部相同,在扩增的第一个循环,第一通用引物和巢式引物先进行扩增,在后续的循环中,第一通用引物、第二通用引物和巢式引物一起进行扩增,得到测序文库。

如图 6b 所示,对于双样本标签序列测序,第二轮 PCR 扩增中用巢式引物、第二标签引物和第一通用引物扩增 2-30 个循环,其中第二标签的 3'端序列和巢式引物 5'端的部分或全部序列相同,标签引物中间有 8-15bp 的标签序列,该标签序列在不同测序平台有不同的固定序列,用于区分不同样本以便后续的多样本混合测序;

第一通用引物的 3'端序列和第一标签引物的 5'端部分或者全部序列相同。在扩增的第一个循环，巢式引物和第一通用引物进行扩增，在后续的循环中巢式引物、第二标签引物和第一通用引物一起对目标文库进行扩增，最终得到双样本标签的扩增子文库。

以下通过具体实施例详细说明本发明的技术方案，应当理解，实施例仅是示例性的，不能理解为对本发明保护范围的限制。

实施例 1

引物设计：针对肿瘤用药相关位点设计巢式多重 PCR 引物，每个扩增子由正向引物、巢式引物和反向引物组成，在反向引物上引入 15nt 的随机序列，用于模板的溯源和错误矫正，采用这些引物对标准品 HD701（horizon 公司）进行多重 PCR 文库制备并完成高通量测序，得到的数据进行分析后检测每个目标位点的突变。

1、第一轮 PCR 扩增，采用 QIAGEN Multiplex PCR Kit(货号 Cat No./ID:206143)，进行 PCR 扩增。

在 200μL PCR 管子配置如下表 1 所示的试剂体系：

表 1	
组分	用量
上一步反应物（标准品HD701）	20 μL
2X PCR反应酶	25 μL
正向引物池（10μM）	2.5 μL
反向引物池（10μM）	2.5 μL
总量	50 μL

#正向引物池、反向引物池如表2和表3所示。

表2 正向引物池

引物	序列
正向引物 01	AATGACATAACAGTTATGATTTTGCAG（SEQ ID NO: 1）
正向引物 02	TGAGTCCTGGCGCTGTGT（SEQ ID NO: 2）
正向引物 03	TGTTGGATCATATTCGTCCACAA（SEQ ID NO: 3）
正向引物 04	CCAGCAGGATGAACCGGG（SEQ ID NO: 4）
正向引物 05	AGTGGAGAAGCTCCCAACC（SEQ ID NO: 5）
正向引物 06	GACAACCCCCACGTGTGC（SEQ ID NO: 6）
正向引物 07	GCAGCCAGGAACGTACTG（SEQ ID NO: 7）

正向引物 08	CCTTACTCATGGTCGGATCACA (SEQ ID NO: 8)
正向引物 09	AGGGACTAGGCGTGGGAT (SEQ ID NO: 9)
正向引物 10	GACAAACTCTACGTCTCCTCC (SEQ ID NO: 10)
正向引物 11	GCCTCAATTCTTACCATCCACAA (SEQ ID NO: 11)
	GAGACAATGAATTAAGGGAAAATGACAAAG (SEQ ID
正向引物 12	NO: 12)
正向引物 13	CATTCGAAAGACYCTAGCCTTAGATA (SEQ ID NO: 13)

正向引物池由上述引物10 μ M等摩尔混合形成，Y为兼并碱基。

表3 反向引物池

引物	序列
反向引物 01	CGCTTGGCCTCCGACTTNNNNNNNNNNNNNNNNNAGGACTTAGCAA GAAGTTATGGAA (SEQ ID NO: 27)
反向引物 02	CGCTTGGCCTCCGACTTNNNNNNNNNNNNNNNNNCATGATGGATGTC ACGTTCTCAAA (SEQ ID NO: 27)
反向引物 03	CGCTTGGCCTCCGACTTNNNNNNNNNNNNNNNNNGAATATAAACTTG TGGTAGTTGGAGCT (SEQ ID NO: 29)
反向引物 04	CGCTTGGCCTCCGACTTNNNNNNNNNNNNNNNNNTGGTTACATCCCT CTCTGCT (SEQ ID NO: 30)
反向引物 05	CGCTTGGCCTCCGACTTNNNNNNNNNNNNNNNNNTTATACACCGTGC CGAACGC (SEQ ID NO: 31)
反向引物 06	CGCTTGGCCTCCGACTTNNNNNNNNNNNNNNNNNTGTTCCCGGACA TAGTCCAG (SEQ ID NO: 32)
反向引物 07	CGCTTGGCCTCCGACTTNNNNNNNNNNNNNNNNNGCCTCCTTCTGCA TGGTATTCTTT (SEQ ID NO: 33)
反向引物 08	CGCTTGGCCTCCGACTTNNNNNNNNNNNNNNNNNGAGAATGGGTAC TCACGTTTCCTT (SEQ ID NO: 34)
反向引物 09	CGCTTGGCCTCCGACTTNNNNNNNNNNNNNNNNNTGTGGAAAAGTC CCAATGGAATATC (SEQ ID NO: 35)
反向引物 10	CGCTTGGCCTCCGACTTNNNNNNNNNNNNNNNNCTCTCTCCACCA GAGCGA (SEQ ID NO: 36)

	CGCTTGGCCTCCGACTTNNNNNNNNNNNNNNNNNAATAGGTGATTTT
反向引物 11	GGTCTAGCTACA (SEQ ID NO: 37)
	CGCTTGGCCTCCGACTTNNNNNNNNNNNNNNNNNATTTTAGCACTTA
反向引物 12	CCTGTGACTCCA (SEQ ID NO: 38)
	CGCTTGGCCTCCGACTTNNNNNNNNNNNNNNNNNTGTGTGGAAGAT
反向引物 13	CCAATCCATT (SEQ ID NO: 39)

反向引物池由上述引物10μM等摩尔混合形成。

按照以下表4所示的程序进行第一轮PCR扩增：

表4

温度	时间	循环数
95℃	2min	1个循环
95℃	10s	2个循环
62℃	2min	
72℃	30s	
72℃	5min	1个循环

得到的PCR产物中加入80μL的XP磁珠 (beads) 进行纯化(beckman公司Agencourt AMPure XP磁珠，货号A63881)，得到的产物溶解于20μL TE溶液中。

2、第二轮PCR扩增，采用QIAGEN Multiplex PCR Kit(货号Cat No./ID: 206143)，进行PCR扩增。

在 200μL PCR 管子配置如下表 5 所示的试剂体系：

表 5

组分	用量
上一步反应物	20 μL
2X PCR反应酶	25 μL
巢式引物池	2.5 μL
第二通用引物	2.5 μL
第一标签引物	2.5 μL
总量	50 μL

#巢式引物池如表6。

表6 巢式引物池

引物	序列
----	----

	ACATGGCTACGATCCGACTTTCAGTGTTACTTACCTGTCTTGTC
巢式引物 01	(SEQ ID NO: 14)
	ACATGGCTACGATCCGACTTTCAGGATGGTGGATGTGGG (SEQ
巢式引物 02	ID NO: 15)
	ACATGGCTACGATCCGACTTGCTGTATCGTCAAGGCACTC (SEQ
巢式引物 03	ID NO: 16)
	ACATGGCTACGATCCGACTTACCCCAATGCAGCGAACAA (SEQ
巢式引物 04	ID NO: 17)
	ACATGGCTACGATCCGACTTAGCTCTCTTGAGGATCTTGAAGG
巢式引物 05	(SEQ ID NO: 18)
	ACATGGCTACGATCCGACTTCTGCCTCACCTCCACCGT (SEQ ID
巢式引物 06	NO: 19)
	ACATGGCTACGATCCGACTTAACACCGCAGCATGTCAA (SEQ ID
巢式引物 07	NO: 20)
	ACATGGCTACGATCCGACTTTGTGATTTTGGTCTAGCCAGAG
巢式引物 08	(SEQ ID NO: 21)
	ACATGGCTACGATCCGACTTGATGATGGGCTCCCGGAA (SEQ ID
巢式引物 09	NO: 22)
	ACATGGCTACGATCCGACTTCGTCTCCTCCGACCACTGT (SEQ ID
巢式引物 10	NO: 23)
	ACATGGCTACGATCCGACTTGATCCAGACAACTGTTCAAACCTG
巢式引物 11	(SEQ ID NO: 24)
	ACATGGCTACGATCCGACTTGCAATTTCTACACGAGATCCTCT
巢式引物 12	(SEQ ID NO: 25)
	ACATGGCTACGATCCGACTTGCAAGAGGCTTTGGAGTATTTTCAT
巢式引物 13	G (SEQ ID NO: 26)

巢式引物池由上述引物10 μ M等摩尔混合形成。

第一标签引物为：

TGTGAGCCAAGGAGTTNNNNNNNNNN[#]TTGTCTTCCTAAGACCGCTTGGC
CTCCGACTT (SEQ ID NO: 40)，#其中N为标签序列，是随机碱基。

第二通用引物为：/5Phos/#GAACGACATGGCTACGATCCGACTT (SEQ ID NO: 41)，引物5'端磷酸化，用于MGI平台的环化。

按照以下表7的程序进行第一轮PCR扩增。

表7

温度	时间	循环数
95℃	2min	1个循环
95℃	10s	24个循环
62℃	2min	
72℃	30s	
72℃	5min	1个循环

得到的PCR产物中加入80μL的XP磁珠 (beads)进行纯化(beckman公司Agencourt AMPure XP磁珠，货号A63881)，得到的产物溶解于20μL TE溶液中。

3、文库质检

对文库进行检测，条带范围在150-200bp之间，结果如图8所示。

4、上机测序

得到的所有产物进行标准化，进行等量混合，混合得到的文库进行平行测序，测序平台MGISEQ-2000，测序类型PE100。

5、数据分析

分析步骤包括过滤接头引物序列、比对等基本的步骤，得到的基本信息如表8，然后采用GATK对目标位点进行突变检测（表9）。

表8 下机数据统计

编号	原始数据	比对率	捕获率	平均深度	均一性 0.2X
样本 1	8349037	99.85%	99.72%	1277933	100%
样本 2	11265885	99.85%	99.68%	1711470	100%
样本 3	10184135	99.88%	99.73%	1542768	100%
样本 4	15748586	99.89%	99.75%	2389774	100%

表9 突变位点检测

突变位点	理论检测	实际检	实际检	实际检	实际检
	值	测值 1	测值 2	测值 3	测值 4
E545K	9.0%	9.7%	9.3%	9.6%	8.6%

H1047R	17.5%	18.7%	18.4%	17.2%	15.8%
D816V	10.0%	9.2%	11.0%	9.2%	10.2%
G719S	24.5%	26.7%	23.5%	24.5%	26.5%
T790M	1.0%	0.9%	1.0%	1.0%	1.0%
L858R	3.0%	3.1%	2.9%	3.3%	3.1%
V600E	10.5%	9.6%	11.2%	9.8%	10.3%
G13D	15.0%	14.7%	14.7%	16.5%	13.8%
G12D	6.0%	6.5%	6.3%	6.4%	6.2%

由图8可以看到,采用巢式多重PCR得到的产物没有引物二聚体和非特异性产物。表8显示下机数据中比对率、捕获率都能达到99%, 0.2X平均深度的均一性可以达到100%, 其中各个扩增子深度在7倍以内(图7), 表明扩增性能优良。对目标位点进行突变检测, 所有样本实际检测到的突变频率和标准品突变检测频率接近, 实际值大约等于 1 ± 0.1 倍的理论值(表9)。

以上应用了具体个例对本发明进行阐述, 只是用于帮助理解本发明, 并不用以限制本发明。对于本发明所属技术领域的技术人员, 依据本发明的思想, 还可以做出若干简单推演、变形或替换。

权 利 要 求

1. 一种巢式多重 PCR 高通量测序文库制备方法, 其特征在于, 所述方法包括:

第一轮 PCR 扩增: 使用正向引物和反向引物对目标区域进行扩增, 其中所述正向引物是与所述目标区域结合的正向特异性序列, 所述反向引物包括 3'端的与所述目标区域结合的反向特异性序列和 5'端的第一通用测序序列;

扩增产物的纯化: 对所述第一轮 PCR 扩增的产物进行纯化;

第二轮 PCR 扩增: 使用巢式引物、第一标签引物和第二通用引物对纯化后的产物进行扩增, 其中所述巢式引物位于所述正向引物的下游, 所述巢式引物包括 5'端的第二通用测序序列和 3'端的特异性序列, 所述第一标签引物 3'端序列和所述反向引物 5'端的第一通用测序序列部分或全部序列相同, 且所述第一标签引物还包括第一标签序列, 所述第二通用引物 3'端序列和所述巢式引物 5'端的第二通用测序序列部分或者全部序列相同。

2. 根据权利要求 1 所述的方法, 其特征在于, 所述第二通用引物是第二标签引物, 该第二标签引物 3'端序列和所述巢式引物 5'端的第二通用测序序列部分或者全部序列相同, 且所述第二标签引物还包括第二标签序列。

3. 一种巢式多重 PCR 高通量测序文库制备方法, 其特征在于, 所述方法包括:

第一轮 PCR 扩增: 使用正向引物、反向引物和第一标签引物对目标区域进行扩增, 其中所述正向引物是与所述目标区域结合的正向特异性序列, 所述反向引物包括 3'端的与所述目标区域结合的反向特异性序列和 5'端的第一通用测序序列, 所述第一标签引物 3'端序列和所述反向引物 5'端的第一通用测序序列部分或全部序列相同, 且所述第一标签引物还包括第一标签序列;

扩增产物的纯化: 对所述第一轮 PCR 扩增的产物进行纯化;

第二轮 PCR 扩增: 使用巢式引物、第一通用引物和第二通用引物对纯化后的产物进行扩增, 其中所述巢式引物位于所述正向引物的下游, 所述巢式引物包括 5'端的第二通用测序序列和 3'端的特异性序列, 所述第一通用引物 3'端序列和所述第一标签引物 5'端部分或全部序列相同, 所述第二通用引物 3'端序列和所述巢式引物 5'端的第二通用测序序列部分或者全部序列相同。

4. 根据权利要求 3 所述的方法, 其特征在于, 所述第二通用引物是第二标签引物, 该第二标签引物 3'端序列和所述巢式引物 5'端的第二通用测序序列部分或者全

部序列相同，且所述第二标签引物还包括第二标签序列。

5. 根据权利要求 1 至 4 任一项所述的方法，其特征在于，所述反向引物上还包括分子标签序列。

6. 根据权利要求 5 所述的方法，其特征在于，所述分子标签序列是 8-24bp 的随机序列。

7. 根据权利要求 1 至 4 任一项所述的方法，其特征在于，所述方法对多个连续的目标区域进行扩增，每个目标区域都有对应的正向引物、反向引物和巢式引物，并且相邻扩增子区域之间形成扩增重叠。

8. 根据权利要求 7 所述的方法，其特征在于，所述正向引物、反向引物和巢式引物均是由多条分别靶向不同的目标区域的引物组成的引物池。

9. 根据权利要求 7 所述的方法，其特征在于，每个目标区域的正向引物和相邻区域的正向引物、巢式引物和相邻区域的巢式引物、反向引物和相邻区域的反向引物扩增方向相反。

10. 根据权利要求 1 至 4 任一项所述的方法，其特征在于，所述第一轮 PCR 扩增和/或第二轮 PCR 扩增的循环数均是 2-30 个循环。

11. 根据权利要求 1 至 4 任一项所述的方法，其特征在于，所述第一标签序列和/或第二标签序列的长度均是 8-15bp。

12. 一种巢式多重 PCR 高通量测序文库制备试剂盒，其特征在于，所述试剂盒包括：

第一轮 PCR 扩增引物，其包括正向引物和反向引物，其中所述正向引物是与目标区域结合的正向特异性序列，所述反向引物包括 3'端的与所述目标区域结合的反向特异性序列和 5'端的第一通用测序序列，所述第一轮 PCR 扩增引物用于对目标区域进行第一轮 PCR 扩增；

第二轮 PCR 扩增引物，其包括巢式引物、第一标签引物和第二通用引物，其中所述巢式引物位于所述正向引物的下游，所述巢式引物包括 5'端的第二通用测序序列和 3'端的特异性序列，所述第一标签引物 3'端序列和所述反向引物 5'端的第一通用测序序列部分或全部序列相同，且所述第一标签引物还包括第一标签序列，所述第二通用引物 3'端序列和所述巢式引物 5'端的第二通用测序序列部分或者全部序列相同，所述第二轮 PCR 扩增引物用于对纯化后的所述第一轮 PCR 扩增的产物进行第二轮 PCR 扩增。

13. 根据权利要求 12 所述的试剂盒，其特征在于，所述第二通用引物是第二标签引物，该第二标签引物 3'端序列和所述巢式引物 5'端的第二通用测序序列部分或者全部序列相同，且所述第二标签引物还包括第二标签序列。

14. 一种巢式多重 PCR 高通量测序文库制备试剂盒，其特征在于，所述试剂盒包括：

第一轮 PCR 扩增引物，其包括正向引物、反向引物和第一标签引物，其中所述正向引物是与所述目标区域结合的正向特异性序列，所述反向引物包括 3'端的与所述目标区域结合的反向特异性序列和 5'端的第一通用测序序列，所述第一标签引物 3'端序列和所述反向引物 5'端的第一通用测序序列部分或全部序列相同，且所述第一标签引物还包括第一标签序列，所述第一轮 PCR 扩增引物用于对目标区域进行第一轮 PCR 扩增；

第二轮 PCR 扩增引物，其包括巢式引物、第一通用引物和第二通用引物，其中所述巢式引物位于所述正向引物的下游，所述巢式引物包括 5'端的第二通用测序序列和 3'端的特异性序列，所述第一通用引物 3'端序列和所述第一标签引物 5'端部分或全部序列相同，所述第二通用引物 3'端序列和所述巢式引物 5'端的第二通用测序序列部分或者全部序列相同，所述第二轮 PCR 扩增引物用于对纯化后的所述第一轮 PCR 扩增的产物进行第二轮 PCR 扩增。

15. 根据权利要求 14 所述的试剂盒，其特征在于，所述第二通用引物是第二标签引物，该第二标签引物 3'端序列和所述巢式引物 5'端的第二通用测序序列部分或者全部序列相同，且所述第二标签引物还包括第二标签序列。

16. 根据权利要求 12 至 15 任一项所述的试剂盒，其特征在于，所述反向引物上还包括分子标签序列。

17. 根据权利要求 16 所述的试剂盒，其特征在于，所述分子标签序列是 8-24bp 的随机序列。

18. 根据权利要求 12 至 15 任一项所述的试剂盒，其特征在于，所述正向引物、反向引物和巢式引物均是由多条分别靶向多个连续的目标区域的引物组成的引物池，并且相邻扩增子区域之间形成扩增重叠。

19. 根据权利要求 18 所述的试剂盒，其特征在于，每个目标区域的正向引物和相邻区域的正向引物、巢式引物和相邻区域的巢式引物、反向引物和相邻区域的反向引物扩增方向相反。

20. 根据权利要求 12 至 15 任一项所述的试剂盒，其特征在于，所述第一标签序列和/或第二标签序列的长度均是 8-15bp。