

权 利 要 求 书

1. 一种构建大库容的基因突变文库的方法，所述方法包括：

5 (1) 根据需要构建文库的氨基酸序列的编码核苷酸序列设计并合成具有突变核苷酸和限制性核酸内切酶的酶切位点的两个或多个寡聚物池，其中设计的相邻的两个寡聚物池酶切后产生相同的粘性末端；

(2) 扩增所述寡聚物池；

(3) 在反应体系中组装寡聚物池，得到组装的寡聚物池；

(4) 扩增组装后的寡聚物池，得到大库容的基因突变文库。

10 2. 根据权利要求 1 所述的方法，所述步骤(2)包括：分别以所述各寡聚物池作为模板，以根据所述各寡聚物池的序列设计的正向引物和反向引物作为引物对，使用高保真 DNA 聚合酶对所述各寡聚物池进行 PCR 扩增，得到扩增的各寡聚物池。

15 3. 根据权利要求 1 或 2 所述的方法，所述步骤(3)包括：加入所述扩增的各寡聚物池，同时加入限制性核酸内切酶和 DNA 连接酶，使用边切边连的方式对所述各扩增的寡聚物池进行组装，得到组装后的寡聚物池。

20 4. 根据权利要求 1-3 中任一项所述的方法，所述步骤(4)包括：以组装后的寡聚物池作为模板，以所设计的第一个寡聚物池的正向引物和最后一个寡聚物池的反向引物作为引物对，使用高保真 DNA 聚合酶进行 PCR，对组装后的寡聚物池进行 PCR 扩增，从而得到大库容的基因突变文库。

5. 根据权利要求 1-4 中任一项所述的方法，其中所述基因突变文库的库容为多达 10^5 ，优选多达 10^6 ，更优选多达 10^7 ，更优选多达 10^8 ，更优选多达 10^9 ，更优选多达 10^{10} 。

25 6. 根据权利要求 1-5 中任一项所述的方法，其中所述限制性核酸内切酶是 IIS 型限制性核酸内切酶。

7. 根据权利要求 6 所述的方法，其中所述 IIS 型限制性核酸内切酶选自下组中一种或多种：*AclI*、*AlwI*、*BbsI*、*BbvI*、*BccI*、*BceAI*、*BciVI*、*BfuAI*、*BmrI*、*BpmI*、*BpuEI*、*BsaI*、*BseRI*、*BsgI*、*BsmAI*、*BsmBI*、*BsmFI*、*BspMI*、*BspQI*、*BsrDI*、*BtgZI*、*BtsCI*、*BtsI*、*EarI*、*EciI*、*EcoP15I*、*FauI*、*FokI*、*HgaI*、*HphI*、*HpyAV*、*MboII*、*MmeI*、*MnII*、*PleI*、*SapI* 和 *SfaNI*。

8. 根据权利要求 7 所述的方法，其中所述 IIS 型限制性核酸内切酶是

BsaI。

9. 根据权利要求 1-8 中任一项所述的方法，其中步骤(1)包括如下：

(i) 在所述编码核苷酸序列中确定粘性末端，以所述粘性末端 3'端将所述序列分成对应于所述两个或多个寡聚物池的两段或多段序列，

5 (ii) 如果将所述序列分成两段，则：

在第一段粘性末端后的 3'端依次引入限制性核酸内切酶的识别序列的反向互补序列和特异性序列 1，从而得到寡聚物池 1，且在第二段的 5'端依次引入粘性末端、限制性核酸内切酶的识别序列和特异性序列 2，从而得到寡聚物池 2；

10 如果将所述序列分成 n 段， n 为大于等于 3 的正整数，则：

在第一段粘性末端后的 3'端依次引入限制性核酸内切酶的识别序列的反向互补序列和特异性序列 1，从而得到寡聚物池 1，且在第二段的 5'端依次引入粘性末端、限制性核酸内切酶的识别序列和特异性序列 2，在第二段粘性末端后的 3'端依次引入限制性核酸内切酶的识别序列的反向互补序列和特异性序列 3，从而得到寡聚物池 2，以此类推，且在第 n 段的 5'端依次引入粘性末端、限制性核酸内切酶的识别序列和特异性序列 $2n-2$ ，在第二段粘性末端后的 3'端依次引入限制性核酸内切酶的识别序列的反向互补序列和特异性序列 $2n-1$ ，从而得到寡聚物池 n 。

15 10. 根据权利要求 9 所述的方法，其中所述特异性序列 1-特异性序列 $2n-1$ 是与原编码核苷酸序列不同源的随机序列。

11. 根据权利要求 6-10 中任一项所述的方法，其中所述粘性末端为单粘性末端或简并粘性末端。

12. 根据权利要求 11 所述的方法，其中所述粘性末端为单粘性末端，且所述寡聚物池的个数为 2-6 个，优选 2 个。

25 13. 根据权利要求 11 所述的方法，其中所述粘性末端为简并粘性末端，且所述寡聚物池的个数为 2 个。

14. 根据权利要求 6-13 中任一项所述的方法，其中所述粘性末端的 GC 含量为 50-75%。

30 15. 根据权利要求 6-14 中任一项所述的方法，其中所述粘性末端不含回文结构。

16. 根据权利要求 2-15 中任一项所述的方法，其中，步骤(3)中所述边

切边连的方式是 Golden Gate 方式。

17. 根据权利要求 2-16 中任一项所述的方法，其中，步骤(2)中所述的 PCR 扩增体系还包含牛血清白蛋白。

18. 根据权利要求 1-17 中任一项所述的方法，其进一步包括：

5 (5) 对步骤(4)得到的基因突变文库产物进行回收和/或纯化，以得到最终文库产物。

19. 根据权利要求 18 所述的方法，其中通过凝胶电泳对步骤(4)得到的基因突变文库产物进行回收和/或纯化。

20. 根据权利要求 18 或 19 所述的方法，其进一步包括：

10 (6) 对步骤(5)中得到的最终文库产物进行测序，以验证序列分布和/或检测氨基酸分布。

21. 根据权利要求 20 所述的方法，其中所述测序为 Sanger 测序和/或 NGS 测序。

15 22. 根据权利要求 1-21 中任一项所述的方法，其中在步骤(1)中合成的每个寡聚物池中的突变核苷酸的个数为 1-108 个，优选 1-21 个。

23. 根据权利要求 1-22 中任一项所述的方法，其中在步骤(1)中合成的每个寡聚物池中的突变核苷酸编码 1-36 个，优选 1-7 个氨基酸的突变。

24. 根据权利要求 1-23 中任一项所述的方法，其中在步骤(1)中合成的每个寡聚物池中的突变核苷酸所编码的突变氨基酸是相邻或不相邻的。

20 25. 根据权利要求 2-24 中任一项所述的方法，其中所述高保真 DNA 聚合酶选自下组中的一种或多种：Phusion DNA 聚合酶、Q5 聚合酶和 primerSTAR 聚合酶。

26. 使用根据权利要求 1-25 中任一项所述的方法构建的基因突变文库。

25 27. 权利要求 1-25 中任一项所述方法构建的基因突变文库在蛋白或多肽筛选中的用途。

28. 一种解析蛋白质中的氨基酸突变与蛋白质的特性、调控和/或功能之间的关系的方法，其包括以下步骤：

(1) 使用根据权利要求 1-25 中任一项所述的方法构建基因突变文库；

30 (2) 将所构建的基因突变文库中的突变基因编码的蛋白质的特征、调控和/或功能与未突变的蛋白质的情况进行比较；和

(3) 解析蛋白质中的氨基酸突变与蛋白质的特性、调控和/或功能之间的

关系。