



一种基于多重PCR的人类线粒体基因组高通量测序方法

孙嘉仪, 杨超, 胡晓雯, 张娟, 周代占, 贺林*, 万春玲*

上海交通大学Bio-X研究院, 上海精神卫生中心, 上海 200030

* 联系人, E-mail: helin@sjtu.edu.cn; clwan@sjtu.edu.cn

收稿日期: 2016-12-08; 接受日期: 2017-02-14; 网络版发表日期: 2017-04-14

国家重点研发计划(批准号: 2016YFC1306802, 2016YFC1306900)和国家自然科学基金(批准号: 81421061, 81361120389)资助

摘要 人类线粒体基因组DNA(mtDNA)是一个16569 bp的双链闭合环状DNA分子,具有母系遗传、多拷贝、高异质性及高变异率等特点,是研究人类遗传和进化上广泛使用的分子标记.近几年,高通量测序技术的出现,使得在短时间内准确测定mtDNA序列成为可能;但目前常用的高通量测序建库方法操作复杂、研究费用相对较高.基于多重PCR扩增的测序方法具有高效率、高灵敏度、低成本的特点,因而适用于大规模线粒体基因组的变异检测分析.利用73个相互重叠的扩增子通过多重PCR方法来扩增中国人的线粒体全基因组,在扩增片段两端连接特定的接头序列,然后在IlluminaHiSeq X Ten平台上进行高通量测序.对获得的测序数据分析发现,mtDNA每个位点的测序深度均达到2000×以上;当测序深度为100×时,所有样本的序列覆盖度都达到100%;数据质量适用于后续的变异检测分析.利用本研究建立的基于多重PCR的二代测序方法无需片段化即可直接上机测序,在复杂遗传病的研究中有着广泛的应用前景.

关键词 人类线粒体基因组, 多重PCR, 二代测序

线粒体是真核细胞中重要的细胞器^[1],在氧化应激、代谢、炎症、自噬和凋亡等生物进程中有着至关重要的作用^[2,3].线粒体功能异常可迅速导致细胞内能量代谢障碍,ATP供应短缺,从而导致一系列疾病的发生,如精神分裂症、帕金森综合征、糖尿病、阿尔兹海默综合征^[4-7],还有研究发现一些癌症发生发展与mtDNA突变密切相关^[8].

线粒体基因组位于线粒体中,是独立于核染色体之外的双链闭合环状遗传物质.人类线粒体基因组全长16569 bp^[9],分为编码区和非编码区两个部分.其中

编码区编码13个蛋白、22个tRNA和2个rRNA,具有较高的保守性;非编码区包含一个复制起始区和一些转录调节因子^[10],该区域变异率较高.人类每个体细胞中包含成百上千个线粒体,每一个线粒体内含有2~10个mtDNA的拷贝^[11].

相较于核DNA,mtDNA具有许多特点,如母系遗传、多拷贝、高变异率和高异质性等^[12].由于这些特点,mtDNA在人类进化和种族遗传等相关领域中已成为非常重要的研究材料^[13,14],而在辅助法医学鉴定领域,特别是在无法得到完整分型的核DNA的降解检材

引用格式: 孙嘉仪, 杨超, 胡晓雯, 等. 一种基于多重PCR的人类线粒体基因组高通量测序方法. 中国科学: 生命科学, 2017, 47: 396-402
Sun J Y, Yang C, Hu X W, et al. A cost-efficient method for sequencing the human whole mitochondrial genome utilizing multiplex PCR-based next generation sequencing. Sci Sin Vitae, 2017, 47: 396-402, doi: 396-40210.1360/N052017-00063

中, mtDNA信息便显示出重要的利用价值^[15].

传统的mtDNA测序技术主要是利用基于克隆文库或PCR扩增产物的桑格(Sanger)测序方法. 虽然桑格测序能够检测出许多突变, 但该方法通量低, 且无法给出定量信息. 高通量测序等技术的兴起, 使得在短时间内对线粒体基因组序列进行快速准确的测序成为可能, 也使人们对mtDNA的特点有了更为深入的认识. 到目前为止, 已经建立了一些mtDNA二代测序方法, 例如, 基于长片段PCR的、基于滚环复制的、基于杂交捕获的方法. 基于长片段PCR扩增的测序方法比较便捷, 自己就可以设计出合适的引物, 无需商业化定制, 但该方法对样本质量要求较高^[16]; 基于滚环复制的测序方法操作简单, 所需的样本量较少, 但该方法对mtDNA的纯度要求较高^[17]; 基于杂交捕获的测序方法对中长片段DNA非常有优势, 且特异性较高, 但该方法对样本的需求量大, 且研究费用很高^[18].

应用传统的建库方法进行线粒体高通量测序很难在大规模样本量的研究中进行推广, 而基于多重PCR的高通量测序方法则表现出许多优势. 多重PCR技术(multiplex PCR)是指在一个反应体系中同时扩增不同区域的多个目的片段, 具有高效、快捷、高特异性、高灵敏度, 以及低成本等突出优势, 已成为生命科学和医学研究的重要手段. Levin等人^[19]在1999年针对人类mtDNA设计了58对可覆盖mtDNA全部区域的特异性引物, 经PCR扩增后成功通过Sanger法完成了测序, 对于线粒体测序在法医鉴定、人类遗传进化和医疗诊断等领域中的应用推广具有十分重要的意义.

本研究采用多重PCR技术, 通过73对相互重叠的扩增子对人类线粒体基因组全序列进行扩增, 在扩增片段两端加入特定接头序列使之适合后续的上机测序及数据分析; 对扩增体系及反应条件进行优化以提高扩增效率, 通过多重PCR扩增后, 产物无需片段化即可直接上机进行高通量测序(Illumina测序平台). 本研究探索的方法灵敏度高、快捷、准确且经济可行, 可应用于大规模样本的mtDNA测序研究.

1 材料与方法

1.1 DNA样本

本研究使用的DNA样本是10年前使用QIAamp DNA Mini Kit试剂盒(Qiagen, 德国)提取的, 并保存在

上海交通大学Bio-X研究院样本库-80℃冰箱中. 从中随机选取96例中国人群正常人的DNA样本, 经Nano Drop 2000分光光度计(Thermo Fisher Scientific, 美国)重新定量后将所有样本稀释到10 ng/μL备用.

1.2 多重PCR扩增

(1) 多重PCR引物的设计. 应用引物设计在线软件Prime-blast(<https://www.ncbi.nlm.nih.gov/tools/>)及Multiplex PCR Primer Designing Tool设计多重PCR引物: 根据数据库中mtDNA修正后的剑桥序列(revised cambridge reference sequence, rCRS, NC_012920.1), 以及二代测序读长(Illumina HiSeq X Ten为150 bp×2)设定扩增序列的长度范围为200~300 bp, 退火温度范围为58℃~62℃, 确保相邻扩增片段之间有重叠区域来选择引物合适的起始及终止位置, 从输出结果中优先选择扩增序列接近300 bp的引物. 最终共设计73对引物, 可覆盖线粒体全基因组序列(网络版附表1); 所有引物均由上海捷瑞生物工程有限公司合成.

(2) 多重PCR引物的分组. 将73对引物按图1所示分成3组, 每组内的扩增产物在mtDNA基因组上的位置均不相邻; 第一组和第二组各包含24对引物, 第三组含25对引物. 在少量样本中进行了二代测序的预实

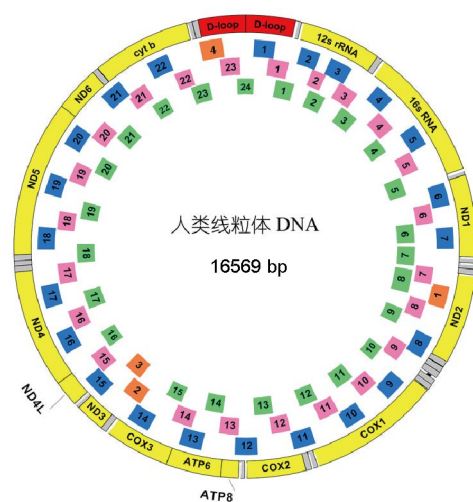


图1 PCR扩增子在mtDNA上分布示意图

图中外部所示为mtDNA双链环状结构: 包括D-loop区(红色方框), 13个编码蛋白的基因和2个编码rRNA的基因(黄色方框), 22个编码tRNA的基因(灰色方框); 内部73个方框示73个扩增子: 蓝色为第1组, 粉色为第2组, 绿色为第3组, 橙色为第4组

验. 结果显示, 有4对引物对应的片段扩增效率很低, 将这4对引物挑出来归为第4组. 最终各组PCR扩增结果经质量控制后显示均适于进行后续实验.

(3) 多重PCR反应体系和扩增条件. PCR反应在96孔板中进行, 6 μ L反应体系中含基因组DNA模板10 ng, 每对引物在体系中的终浓度为50~70 nmol/L(网络版附表1), 1 $\times$ KAPA2G Robust HotStartReadyMix(Kapa Biosystems, 美国); 反应在9700型PCR仪(Applied Biosystems, 美国)中进行.

PCR扩增条件: 95 $^{\circ}$ C预变性2 min; 95 $^{\circ}$ C变性30 s, 56 $^{\circ}$ C退火30 s, 72 $^{\circ}$ C延伸60 s, 反应进行45个循环; 72 $^{\circ}$ C延伸5 min, 4 $^{\circ}$ C保存.

1.3 mtDNA高通量测序

(1) 多重PCR产物的酶切纯化. 多重PCR产物用核酸外切酶 I (exonuclease I, New England Biolabs, 美国)和虾碱性磷酸酶(shrimp alkaline phosphatase, SAP, Affymetrix, 美国)处理. Exonuclease I 是单链特异性3'→5'核酸外切酶, 可将ssDNA 3'-OH末端分解生成5'-单核苷酸, 去除PCR反应体系中残留的单链引物; SAP可将脱氧核苷三磷酸的磷酸基团分解, 去除PCR反应中残留的dNTP, SAP-Exonuclease I 混合使用达到纯化PCR产物的目的. 酶切纯化反应体系: 6 μ L的PCR产物, 0.5 U的SAP, 2 U的Exonuclease I, 1 $\times$ SAP缓冲液, 1 $\times$ Exonuclease I缓冲液. 反应条件: 37 $^{\circ}$ C孵育60 min, 80 $^{\circ}$ C反应20 min.

(2) 高通量测序前处理-加接头. 将上述酶切纯化后的产物稀释100倍, 连接Illumina二代测序平台所需的接头Access Array™ Barcode Library for Illumina Sequencers(Fluidigm, 美国). 接头上含有与测序引物互补结合的片段, 通过该片段和测序引物结合来对目的片段进行测序. 此外不同接头含有不同的标签序列, 用于后续数据分析时区分不同的样本. 连接反应: 95 $^{\circ}$ C预变性10 min; 95 $^{\circ}$ C变性15 s, 60 $^{\circ}$ C退火30 s, 72 $^{\circ}$ C延伸30 s, 反应进行12个循环; 72 $^{\circ}$ C延伸3 min, 4 $^{\circ}$ C保存.

(3) 高通量测序前处理-磁珠纯化. 将连接不同接头的样本等体积混合, 用AgencourtAMPure XP磁珠(Beckman Coulter, 美国)进行纯化. 对磁珠纯化后的产物用预制的E-Gel®EX 2%琼脂糖凝胶(Invitrogen, 美国)电泳, 因PCR产物连接接头后长度在300~500 bp范围内, 为确保回收所有产物, 故割胶回收300~500 bp的

条带用于后续实验.

(4) 高通量测序. 用Agilent 2100 Bioanalyzer(Agilent Technologies, 美国)检测样本DNA片段的大小及浓度, 质检合格后, 在IlluminaHiSeq X 10平台上进行二代测序.

1.4 数据分析

IlluminaHiSeq X 10高通量双末端测序后得到FASTQ文件, 将每个样本的测序数据根据不同的接头序列信息进行分离. 用BWA软件将样本测序得到的序列和数据库中的人类线粒体基因组参考序列(rCRS, NC_012920.1)进行比对, 输出SAM格式的文件, 然后用SAMtools软件将SAM文件转化成BAM格式的文件, 接着将BAM文件进行排序、索引. 使用GATK(genome analysis toolkit)分析软件统计测序片段的读数、数据总产量, 评估测序深度、覆盖度和均一性.

2 结果与分析

2.1 高通量测序数据的产量

96个样本共得到7008个扩增片段, 经HiSeq X Ten高通量测序, 得到大小为28.2 GB的数据. 每个样本检测到的碱基数为154.39~574.24 Mbp不等(图2), 平均值为304.16 Mbp(网络版附表2).

2.2 高通量测序数据的总读数和匹配率

所有样本的测序片段总读数(total reads)在 $1.07\times 10^6\sim 3.95\times 10^6$ 范围内(图3A). 平均读数为 2.11×10^6 , 其中平均 2.09×10^6 个读数能够匹配到线粒体参考序列上, 平均匹配率为99.19%(网络版附表3). 79个样本的匹配率达99%~100%, 14个样本的匹配率为98%~99%, 3个样本的匹配率为96%~98%(图3B).

2.3 高通量测序的深度和覆盖度

样本的平均测序深度范围为 $9318\times\sim 34660\times$ (网络版附表4), 72%的样本平均测序深度集中在 $14000\times\sim 20000\times$ (图4A). 统计结果表明, 测序深度为 $100\times$ 时, 所有样本的序列覆盖度均为100%; 进一步统计测序深度达到 $6000\times$ 的片段, 结果有94个(97.9%)样本的序列覆盖度为100%, 仅两个(2.1%)样本的序列覆盖度为99.9%(图4B和C, 网络版附表4).

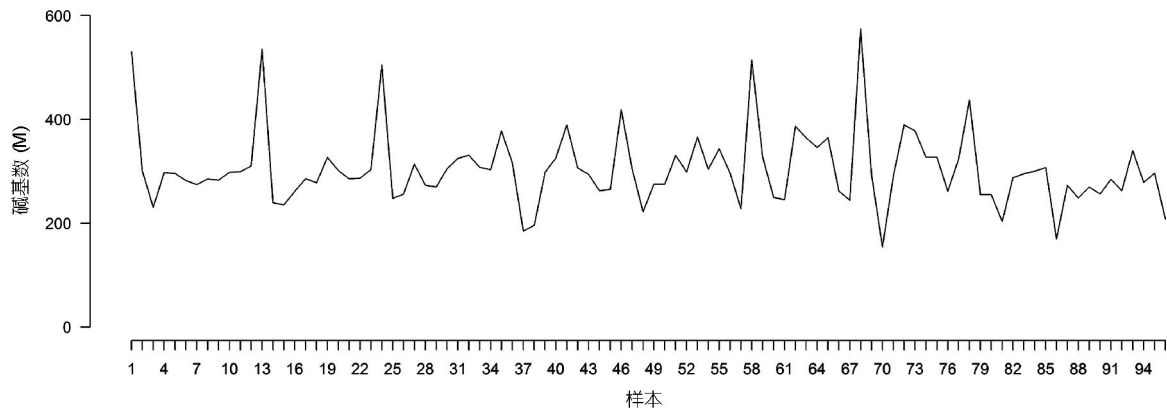


图2 样本检测得到的碱基数
横坐标为样本编号, 纵坐标为碱基数

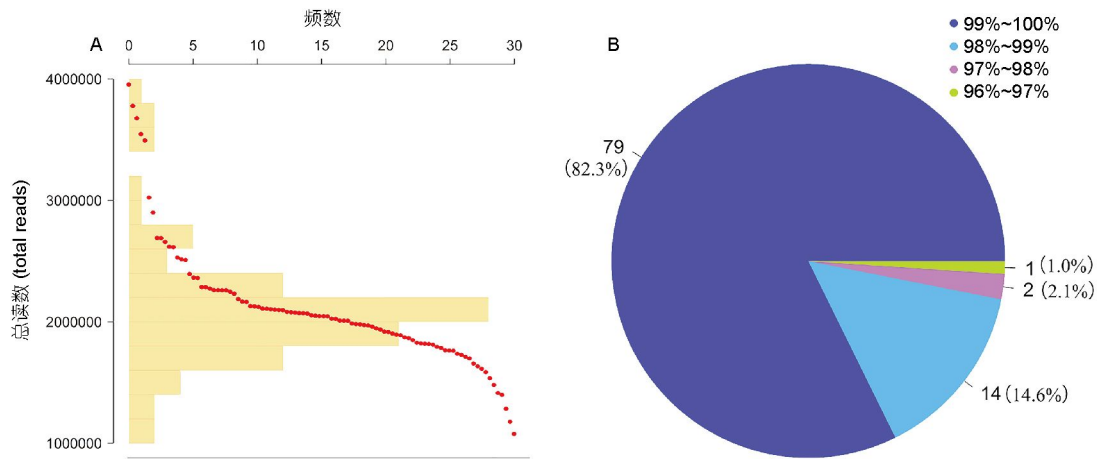


图3 样本的总读数(total reads)和匹配率

A: 96个样本总读数的分布情况, 红色散点表示每个样本的总读数值; B: 96个样本匹配率的分布情况

2.4 碱基位点和扩增子的测序深度

线粒体全基因组序列(16569 bp)上每个位点的测序深度范围为 $2571\times\sim 82380\times$ (图5A); 每个扩增子的平均测序深度为 $5806\times\sim 62910\times$ 不等(图5B, 网络版附表5). 71%的扩增子的平均测序深度集中在 $10000\times\sim 30000\times$ 范围内.

3 讨论

人类线粒体DNA是位于细胞核以外的唯一基因组DNA, 与体细胞核DNA只有两个拷贝相比, mtDNA有成百上千个拷贝, 其没有组蛋白的保护, 更易受活性氧自由基(reactive oxygen species, ROS)的攻击, 容易发生突变^[20]. mtDNA突变会导致线粒体功能异常, 进而

迅速导致细胞内能量代谢障碍, ATP供应短缺. 线粒体是细胞内ROS的主要来源, 伴随电子转移过程中产生的ROS会改变线粒体膜的通透性, 引起线粒体的结构异常, 如线粒体肿胀、线粒体膜完整性破坏, 进而引起细胞色素C释放进入细胞质, 引起细胞死亡^[21]. 许多研究证据都表明, mtDNA突变后更易引起细胞的结构和功能异常, 且与多种疾病的发生发展密切相关^[22]. 目前常规的mtDNA检测方法大多是对高突变区进行一代测序. 一代测序能够准确地检测出大多数突变, 但无法给出异质性位点不同基因型比例的定量信息, 且对于异质性低于15%的突变体很难检出^[23], 而且该方法步骤繁琐、费时费力、测序区段有限, 极大地限制了mtDNA在复杂疾病中的研究和应用.

高通量测序技术的出现极大地推动了生物学和

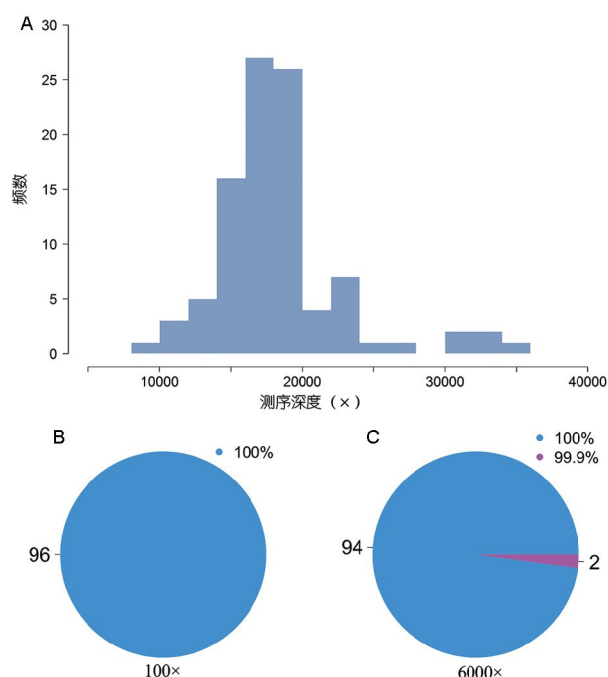


图4 每个样本的平均测序深度及覆盖度

A: 所有样本的平均测序深度分布; B: 统计测序深度为100×的片段, 所有样本的序列覆盖度均为100%; C: 统计测序深度为6000×的片段, 94个样本的序列覆盖度为100%, 2个样本为99.9%

医学的发展. 高通量测序技术能在单次试验中对几十万至几百万条DNA分子同时进行序列测定, 使得对一个物种(或细胞器)的基因组在短期内进行细致全貌的分析成为可能. 此外, 高通量测序技术还有完美的定量功能, 因为样品中某种DNA被测序的次数反映了样品中这种DNA的丰度, 这一点在拷贝数较多的线粒体基因组测序中则能提供更多的线索. 因此, 通过高

通量测序对线粒体基因组序列进行分析, 能够使研究者们有机会对其SNP位点进行精细定位及遗传特征分析^[24], 为从DNA分子水平上阐释线粒体基因突变导致的疾病提供了机制方面的指导信息, 也为线粒体疾病患者实现个体化精准医疗提供了理论基础.

目前常用的线粒体高通量测序的建库方法成本很高, 很难在大规模样本的研究中进行推广. 基于多重PCR的高通量测序方法可以实现适于复杂疾病的大规模样本分析. 该方法早在1988年就被用来快速检测人杜氏肌营养不良相关基因外显子的缺失情况^[25]. 2010年He等人^[26]通过多重PCR测序的方法对结直肠癌病人的线粒体进行了研究, 发现了结直肠癌病人线粒体突变的新位点和母系遗传等规律, 但该研究中采用的650 bp片段的PCR扩增产物仍需片段化处理后才能进行上机测序, 测序成本高. 而本研究采用的PCR扩增产物片段较小, 无需片段化处理可直接上机. Gunnarsdottir等人^[27]在2011年运用罗氏454 Life Sciences平台进行mtDNA测序, 结果在平均测序深度为55×时, 仍有<1%的序列未覆盖; Chaitanya等人^[28]在2015年利用多重PCR技术在Ion Torrent平台上对mtDNA进行测序, 其结果在获得100% mtDNA覆盖度时的测序深度只能达到50×. 本研究利用73对相互重叠的扩增子对线粒体序列全覆盖扩增, 所有样本100%序列覆盖时的测序深度达100×. 与传统的Sanger测序相比, 二代高通量测序具有碱基识别错误率高的弱点, 将会影响变异检测分析结果的准确性, 但当测序深度达到30×时, 则可以通过不同reads进行矫正以提高其准确率, 并且随着测序深度的增加, 准确率则进一步提升, 完全可以和Sanger测序媲

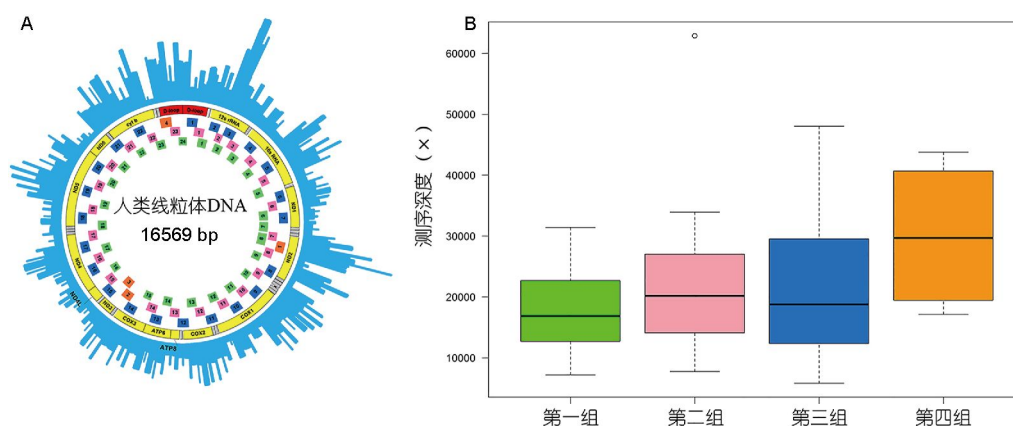


图5 每个位点的测序深度和每对扩增子的测序深度

A: mtDNA每个位点的测序深度(96个样本的中位数); B: 每对扩增子的测序深度(96个样本的平均值)

美. 由于mtDNA具有高度异质性, 更高的测序深度则意味着能够对异质性进行更好的分析. 本研究结果表明, 利用建立的基于多重PCR的方法对mtDNA进行高通量测序所产生的数据质量适用于后续的变异检测分析.

本研究在多重PCR实验之前, 预先进行了单重PCR检测, 以确保每对引物都能够扩增出正确的、单一的目的条带. 为了提高扩增效率, 并使多重PCR反应中各扩增子的扩增效率相对平衡, 通过对扩增体系及反应条件进行了优化以提高扩增效率. 最初在多重PCR反应中使用了相等物质的量的引物(每对引物的终浓度为50 nmol/L), 但是扩增效率并非均一, 在优化了各种反应条件后, 有些引物对应的产物仍不明显; 最终通过改变反应中各种引物的比例, 增加扩增效率低的引物的量, 降低扩增效率高的引物的量, 使各个扩增

子获得较均一的扩增效率. 本研究采用多重PCR技术扩增人类线粒体全基因组, 扩增产物不需要片段化处理即可直接上机进行高通量测序(Illumina测序平台). 本研究提出的测序方法实验操作过程简单, 需要的样本DNA仅需10 ng, 且对样本的完整性要求不高, 可以对降解检材进行检测.

迄今为止, 已经有很多mtDNA的高通量测序方法, 但几乎没有方法是针对亚洲人群的, 本研究建立的方法所设计的引物是根据中国人mtDNA序列进行修改调整过的, 而且亚洲人群之间的遗传差异相对较小, 所以该方法可以特异性地对亚洲人群进行线粒体全基因组测序. 本研究探索出的基于多重PCR扩增的mtDNA高通量测序方法灵敏度高、快捷、准确且经济可行, 适用于线粒体疾病的精准医疗研究, 也为其他复杂遗传疾病的研究提供了新思路.

参考文献

- 1 Frey T G, Mannella C A. The internal structure of mitochondria. *Trends Biochem Sci*, 2000, 25: 319–324
- 2 Kujoth G C, Hiona A, Pugh T D, et al. Mitochondrial DNA mutations, oxidative stress, and apoptosis in mammalian aging. *Science*, 2005, 309: 481–484
- 3 Ott M, Gogvadze V, Orrenius S, et al. Mitochondria, oxidative stress and cell death. *Apoptosis*, 2007, 12: 913–922
- 4 Martorell L, Segué S, Folch G, et al. New variants in the mitochondrial genomes of schizophrenic patients. *Eur J Hum Genet*, 2006, 14: 520–528
- 5 Bender A, Krishnan K J, Morris C M, et al. High levels of mitochondrial DNA deletions in substantia nigra neurons in aging and Parkinson disease. *Nat Genet*, 2006, 38: 515–517
- 6 Charoute H, Kefi R, Bounaceur S, et al. Novel variants of mitochondrial DNA associated with Type 2 diabetes mellitus in Moroccan population. *Mitochondrial DNA A DNA Mapp Seq Anal*, 2016, 7: 1–5
- 7 Tillement L, Lecanu L, Papadopoulos V. Alzheimer's disease: effects of β -amyloid on mitochondria. *Mitochondrion*, 2011, 11: 13–21
- 8 Verschoor M L, Ungard R, Harbottle A, et al. Mitochondria and cancer: past, present, and future. *BioMed Res Int*, 2013, 2013: 1–10
- 9 李彩霞, 胡亚冲, 龙建纲, 等. 线粒体DNA甲基化研究进展. *中国科学: 生命科学*, 2016, 46: 880–885
- 10 Sigurðardóttir S, Helgason A, Gulcher J R, et al. The mutation rate in the human mtDNA control region. *Am J Hum Genet*, 2000, 66: 1599–1609
- 11 Ding J, Sidore C, Butler T J, et al. Assessing mitochondrial DNA variation and copy number in lymphocytes of ~2,000 sardinians using tailored sequencing analysis tools. *PLoS Genet*, 2015, 11: e1005306
- 12 Brown A, Amunts A, Bai X, et al. Structure of the large ribosomal subunit from human mitochondria. *Science*, 2014, 346: 718–722
- 13 Brown W M, George M, Jr., Wilson A C. Rapid evolution of animal mitochondrial DNA. *Proc Natl Acad Sci USA*, 1979, 76: 1967–1971
- 14 Yao Y, Zhang Y. Pitfalls in the analysis of ancient human mtDNA. *Chin Sci Bull*, 2003, 48: 826–830
- 15 Schneider P M. Scientific standards for studies in forensic genetics. *Forensic Sci Int*, 2007, 165: 238–243
- 16 Maricic T, Whitten M, Pääbo S. Multiplexed DNA sequence capture of mitochondrial genomes using PCR products. *PLoS ONE*, 2010, 5: e14004
- 17 Simison W B, Lindberg D R, Boore J L. Rolling circle amplification of metazoan mitochondrial genomes. *Mol Phylogenet Evol*, 2006, 39: 562–567
- 18 Mohandesan E, Speller C F, Peters J, et al. Combined hybridization capture and shotgun sequencing for ancient DNA analysis of extinct wild and domestic dromedary camel. *Mol Ecol Resour*, 2017, 17: 300–313
- 19 Levin B C, Cheng H, Reeder D J. A human mitochondrial DNA standard reference material for quality control in forensic identification, medical diagnosis, and mutation detection. *Genomics*, 1999, 55: 135–146
- 20 Bolisetty S, Jaimes E A. Mitochondria and reactive oxygen species: physiology and pathophysiology. *Int J Mol Sci*, 2013, 14: 6306–6344

- 21 Wang C, Youle R J. The role of mitochondria in apoptosis. *Annu Rev Genet*, 2009, 43: 95–118
- 22 Neiman M, Taylor D R. The causes of mutation accumulation in mitochondrial genomes. *Proc R Soc B-Biol Sci*, 2009, 276: 1201–1209
- 23 周晓光, 任鲁风, 李运涛, 等. 下一代测序技术: 技术回顾与展望. *中国科学: 生命科学*, 2010, 40: 23–37
- 24 Lee E Y, Lee H Y, Oh S Y, et al. Massively parallel sequencing of the entire control region and targeted coding region SNPs of degraded mtDNA using a simplified library preparation method. *Forensic Sci Int-Genets*, 2016, 22: 37–43
- 25 Chamberlain J S, Gibbs R A, Rainer J E, et al. Deletion screening of the Duchenne muscular dystrophy locus via multiplex DNA amplification. *Nucleic Acids Res*, 1988, 16: 11141–11156
- 26 He Y, Wu J, Dressman D C, et al. Heteroplasmic mitochondrial DNA mutations in normal and tumour cells. *Nature*, 2010, 464: 610–614
- 27 Gunnarsdottir E D, Li M, Bauchet M, et al. High-throughput sequencing of complete human mtDNA genomes from the Philippines. *Genome Res*, 2011, 21: 1–11
- 28 Chaitanya L, Ralf A, van Oven M, et al. Simultaneous whole mitochondrial genome sequencing with short overlapping amplicons suitable for degraded DNA using the ion torrent personal genome machine. *Hum Mutat*, 2015, 36: 1236–1247

A cost-efficient method for sequencing the human whole mitochondrial genome utilizing multiplex PCR-based next generation sequencing

SUN JiaYi, YANG Chao, HU XiaoWen, ZHANG Juan, ZHOU DaiZhan, HE Lin & WAN ChunLing

Shanghai Mental Health Center, Bio-X Institutes, Shanghai Jiao Tong University, Shanghai 200030, China

The mitochondrial genome of human is organized as a circular loop of double-stranded DNA with 16569 bp in length. Mitochondrial DNA (mtDNA) is extensively used as a biomarker in human evolution and population genetics due to the characteristics of maternal inheritance, multiple copies in cells, high mutational rates, high heteroplasmy levels and so on. Next-generation sequencing (NGS) has become a rapid and efficient approach to sequence mtDNA recently, while library-building is complex and expensive. Multiplex PCR has the advantage of high sensitivity and saving time and effort. Therefore, this technique was used to amplify whole complete mitochondrial genomes in Chinese in the present study. The whole mitochondrial genome was covered by 73 overlapped short amplicons, which was subsequently sequenced on an Illumina HiSeq X Ten instrument. The mean per-mtDNA-base read depth was higher than 20000×. 100% mitochondrial genome coverage was obtained across all the samples with a depth of coverage threshold of 100×. The data of high throughput sequencing was qualified for mutation analysis. The present sequencing method based on multiplex PCR leaves out the process of library-building. This sequencing methodology that we have developed is efficiency, economic and has a broad application prospect in the research of complex genetic disorders.

human mitochondrial DNA, multiplex PCR, next-generation sequencing

doi: [10.1360/N052017-00063](https://doi.org/10.1360/N052017-00063)