

基于高通量测序技术的法医 DNA 分析方法研究

●徐梦莹



[摘要] 高通量测序技术,以其多位点并行检测、高灵敏度和深度解析能力,为法医 DNA 分析带来了革新性的解决方案。它有效突破了传统短串联重复序列分型方法在混合样本解析、微量与降解 DNA 检测,以及遗传标记分析中的局限性。高通量测序技术在传统短串联重复序列精细化分型、单核苷酸多态性标记扩展及痕迹来源推断等方面展现出明显优势。高通量测序技术通过优化目标区域捕获、文库制备和生物信息学分析流程,明显提升了个体识别的准确性和效率。本文综述了高通量测序技术在法医领域的应用原理、关键技术及特殊需求,并分析了其在复杂案件中的实际效能和未来发展趋势,旨在推动该技术在司法鉴定领域的全面应用。

[关键词] 高通量测序技术;法医 DNA 分析;STR 分型;SNP 标记

法医 DNA 分析是司法鉴定的重要技术手段,其广泛应用于个体识别、亲缘关系鉴定和复杂案件侦破中。传统短串联重复序列(STR)分型方法虽然成熟可靠,但在处理混合样本、微量及降解 DNA 检材,以及复杂遗传标记分析方面存在一定的局限性。近年来,高通量测序(HTS)技术的迅速发展为法医 DNA 分析提供了突破性解决方案。HTS 凭借强大的多位点并行检测能力、高分辨率特性以及深度解析优势,有效突破了传统技术的局限性。同时,HTS 技术凭借海量数据的生成能力和高灵敏度检测技术,在解析复杂样本、提升检出率以及丰富遗传数据库方面展现出明显优势。HTS 的多种应用潜力,包括 STR 精细化分型、单核苷酸多态性(SNP)标记扩展、混合样本解析及痕迹来源推断,为法医 DNA 分析提供了更精准和高效的技术支持。

高通量测序技术在法医 DNA 分析中的应用原理与需求

(一)高通量测序的基本原理

高通量测序技术是相对传统 Sanger 测序而言的第二代或第三代测序技术。高通量测序技术核心特征在于其能够同时对数量庞大的 DNA 分子进行并行测序,实现对海量数据的快速读取。以第二代测序平台(如 Illumina 平台)为例,待测 DNA 先经过适当的打断处理,然后进行文库制

备。接着,利用接头(Adapter)将其固定于测序芯片或微珠表面,通过边合成边测序技术(Sequencing-by-Synthesis, SBS)等先进的化学手段,确保在每轮碱基延伸过程中都能精准地采集到相应的荧光信号,从而准确记录碱基序列信息。相比于传统的毛细管电泳单次输出几百条碱基序列,高通量测序技术,如 Illumina 和 Ion Torrent 平台,能够在短时间内对多个样本进行深度检测,一次实验可读取数百万条序列,读取长度从 25bp 到 450bp 不等,数据产出量可达数十 G 到数十 Tb。

(二)法医 DNA 分析的特殊需求

法医 DNA 分析与医学或基础科研的基因组研究相比,具有明显的特殊性和应用导向,主要体现在以下方面。

(1)混合样本的分离与识别:在犯罪现场或司法鉴定中,常见到多人混合 DNA 样本,传统 STR 方法在混合比例较高时面临解析困难。高通量测序通过深度测序和多位点检测,能够更好地区分不同个体的等位基因。

(2)微量与降解样本的检测:法医样本往往含量比较低,且常受到环境、时间和外力等多重因素的破坏。高通量测序利用其高灵敏度和高覆盖度,可在相对较低起始模板量下实现稳定的扩增与测序,有助于提高对微量或降解样本的检出率。

(3)法医数据库与突变率考量:法医应用对于 STR 或 SNP 数据库的精确构建与比对至关重要。高通量测序带来

的多样性数据可丰富法医数据库，增强个体识别与亲缘鉴定的准确度。同时，通过深度测序可实时监测 DNA 区域的突变率，为法医遗传研究提供更多信息。

(4)操作简化与结果可解释性：法医鉴定需要相对较短的检测周期及较为稳健的自动化流程。随着高通量测序平台的持续更新迭代和生物信息分析软件的日益完善，法医应用中的自动化与流程简化逐渐成为现实。然而，这对仪器精度、试剂质量以及信息分析的专业能力提出了更为严格的要求。

(三)引入高通量测序技术的意义

将高通量测序技术引入法医 DNA 分析，不仅是对既有鉴定手段的补充，更是对法医基因分型体系的全面升级。高通量测序能够在一次实验中完成多个基因组的并行检测，大幅提高检测效率。高通量测序的高分辨率特性可捕捉到传统方法难以检测到的小结构变异或罕见突变，增强案件侦查的确定性。对于低质量、降解严重的样本，高通量测序也具备一定的适应能力，降低因 DNA 量或完整性不足而造成检验失败的风险。这些优势为侦查破案、排除嫌疑以及精准的亲子鉴定提供了更加可靠的技术支撑。

高通量测序在法医 DNA 分析中应用的关键技术

(一)目标区域捕获与文库制备

在法医 DNA 分析中，最常见的目标区域包括 STR 位点、SNP 位点、多态性 InDel 和微单倍型 (Microhaplotype, MH) 等。对于不同的目标区域，往往需要在测序前进行针对性的文库制备与富集。

(1)多重 PCR 扩增：针对若干 STR 或 SNP 位点设计多重引物，通过一次或少数几次 PCR 扩增对多个目标序列进行富集。该方法简单快速，适用于常规 STR 基因组的检测。

(2)捕获探针杂交：对基因组中的目标区域 (如特定 SNP 或罕见位点) 设计寡核苷酸探针，富集到目标序列后再进行测序。该方法灵活度较高，可根据案件需求定制探针方案。

(3)文库制备流程：在完成 PCR 扩增或探针捕获后，还需进行加接头、加索引 (Index)、纯化等一系列步骤，以便将富集后的目标片段转化为可在测序平台上进行测序的文库。文库制备的质量直接影响后续测序结果的准确性与一致性，是法医应用中的关键环节。

(二)测序平台的选择与性能比较

高通量测序平台包括 Illumina、Ion Torrent、PacBio 和 Oxford Nanopore 等。

(1)Illumina 平台：如 Miseq 和 NovaSeq 6000，采用 SBS (边合成边测序) 技术，提供高准确度和大数据量，读长

一般在 75~300bp 范围内，但通过特定型号如 Miseq，读长可达到 2×300bp。这些平台成本相对可控，适合法医研究中常见的 STR/SNP 分型。

(2)Ion Torrent 平台：测序原理基于半导体芯片检测氢离子的释放，读长与准确率与 Illumina 平台相近，但对 GC 含量极端不均匀区域可能存在偏好性。

(3)第三代测序平台：如 PacBio 与 Oxford Nanopore，读长可达到数千至数万碱基，可更完整地检测大片段 STR 或结构变异。然而，其单次读长准确度与成本尚有待进一步优化，尚未大规模应用于常规法医检测领域。

在法医应用中，应根据目标检测位点的长度、需要的深度覆盖以及经费预算选择合适的平台；对于微量、降解或高重复序列区域的检测，往往还需使用多种平台互补或采用不同测序策略以保证结果准确度。

(三)生物信息学分析流程

高通量测序在法医领域的应用离不开可靠的生物信息学分析，典型流程如下。

(1)数据预处理：去除测序文库中的接头序列与低质量碱基，对测序读长进行过滤、质控与聚类，以确保后续比对的准确性。

(2)比对与变异检测：将质控后的 Reads 比对到参考基因组或参考序列上，利用专门的算法识别 STR 基因组的重复数目变动或 SNP 等位基因分型信息。同时，对可能存在的插入/缺失变异进行统计和判定。

(3)混合样本解析：对多个人源 DNA 混合的样本，通过等位基因峰值深度、不同比例分型以及统计学方法区分不同个体的基因型。随着建模与机器学习技术的发展，高维度测序数据可支持更加精准的混合比例估计与个体分离。

(4)结果可视化与数据库比对：通过高通量测序技术，最终生成的分型信息可以进行可视化展示，并与法医数据库 (如 CODIS 数据库) 或本地数据库中的已知基因型进行比对，以确定个体身份或亲缘关系。例如，高通量测序技术能够一次检测 STR、SNP、mtDNA 等多种位点，获得更多的个体识别信息，给侦查人员提供更多有效线索，助力案件的及时侦破。

(5)质量控制与法庭科学证明：在实际司法应用中，生物信息学分析须遵循严格的标准操作流程 (SOP)，并提供详细的质量控制报告与法庭科学说明，以保证测序结果在法律层面的有效性和可采信性。

高通量测序技术在法医 DNA 分析中的应用与发展趋势

(一)在法医 STR 分型与 SNP 标记中的应用

(1)STR 分型的新方案：高通量测序为 STR 分型引入了序列信息层面的精细辨别，不仅能准确识别 STR 的重复次

数,还能深入揭示 STR 重复单元内部的碱基变异情况,进而明显提升个体识别的准确性;对于混合度较高的样本,深度测序可明显增加对稀有等位基因的检出能力,减少漏检。

(2)SNP 标记在法医中的优势:与 STR 相比,SNP 标记展现出突变率低、分型流程简便以及结果高度可重复性的优势,而且在降解严重的 DNA 样本中,SNP 标记的短片段检测尤为有效。高通量测序技术具备在一次实验中同时分析数百乃至上千个 SNP 位点的能力,为复杂案件的个体识别与族群溯源提供了前所未有的精确度。

(3)微单倍型:是一类由紧密连锁的 SNP 组成的短片段,可同时具有高多态性与短序列长度的优势,适合低质量 DNA 检测和混合样本分型。高通量测序提供了对微单倍型的高分辨率检测,为法医应用提供了新工具。

(二)在混合样本解析与痕迹鉴定中的应用

(1)混合比例与数量估算:传统法医分析方法在面对两人或多人高度混合时常难以给出明确的分型判断。高通量测序可借助深度分型方法,通过分析每个位点的测序深度与等位基因比例,较准确地评估混合比例,进而辅助筛查可能的嫌疑人数量。

(2)同卵多胎个体区分:传统 STR 方法在区分同卵双胞胎时面临较大挑战,而高通量测序技术则通过检测超高分辨率或低频突变,有望捕捉到罕见的体细胞突变或表现遗传差异,进而在一定程度上实现对同卵多胎个体的有效区分。虽然这一技术当前尚处于探索阶段,但为疑难案件鉴定提供了新的方向。

(3)痕迹来源鉴定:随着法医技术的发展,案件中常会遇到对具体生物学来源(如血液、唾液、皮肤组织等)进行识别的需求。借助 RNA 测序、甲基化测序等前沿技术,科学家能够推断出 DNA 的来源组织类型及个体差异,为法医鉴定中检材来源的判定提供更加丰富的参考依据。

(三)高通量测序技术在法医 DNA 分析中的未来发展趋势及其面临的挑战

(1)多平台整合与测序深度优化:为更精准地完成复杂样本分析,需要在不同测序平台之间进行策略整合,如利用短读长平台(Illumina)进行高精度分型,再辅助以长读长平台(PacBio 或 Nanopore)解析大型插入/缺失或复杂结构变异。此外,还需进一步探索在有限预算下如何平衡测序深度、目标区域选择与检出限需求。

(2)自动化与标准化:法医鉴定具有高时效与可重复性的要求。高通量测序流程中涉及 PCR 富集、文库制备及生物信息分析多个关键环节,需不断完善自动化设备以及标准操作流程,降低操作人员水平差异对结果的影响,并在法庭

科学层面建立行业认可的标准化规范。

(3)新兴技术与多组学融合:除了对基因组 DNA 本身进行测序分析,表观基因组学、转录组学和单细胞测序等新技术也可能在法医检验中发挥补充作用。例如,通过检测 DNA 甲基化,可以推断 DNA 来源组织类型以及个体的生理状态,如年龄推断、死亡时间鉴定、组织来源鉴定及个体识别等,为法医分析提供更为全面的信息支持。

(4)法律与伦理问题:随着高通量测序技术在法医领域的深入应用,确保数据隐私保护和伦理合规变得尤为重要。法医样本数据涉及个人敏感信息,必须严格遵守《世界医学协会赫尔辛基宣言》《涉及人的健康相关研究国际伦理指南》及《中华人民共和国生物安全法》等法律法规,以防范对个体基因信息的不当滥用。同时,跨地区或跨国的 DNA 数据库共享也需要法律和伦理层面的审慎考量。

Q 结束语

综上所述,高通量测序技术在法医 DNA 分析中展现出多位点并行检测、高灵敏度和深度解析的独特优势,有效解决了传统方法在混合样本解析、微量与降解 DNA 检材处理以及遗传标记分析中的一些问题。借助目标区域捕获技术和平台优化手段,并融合高效的生物信息学分析流程,可以明显增强个体识别的准确性及对复杂案件解析的效率。在法医实际应用中,随着技术自动化与流程标准化的推进,以及法律与伦理框架的完善,有助于提升高通量测序技术在法医 DNA 分析中的可操作性与社会接受度。未来,随着多组学技术的融合与新型测序平台的应用,法医 DNA 分析有望实现更精准的信息解读,从而为司法领域的技术革新注入持续动力。

参考文献

- [1]李妍.高通量测序技术的研究进展[J].中国医学工程,2019,27(3):32-37.
- [2]吴树恒,梁羽,吴薇.DNA 双链断裂高通量测序技术的研究进展[J].中国细胞生物学报,2024,46(8):1488-1496.
- [3]王鸿迪,于蛟,贾菲,等.应用高通量测序 DNA 甲基化差异体液识别研究[J].中国法医学杂志,2019,34(2):136-141.
- [4]谢全.基于高通量测序的法医 SNP 系谱检测技术研究[D].北京:中国人民公安大学,2023.

作者简介:

徐梦莹(1990—),女,汉族,江苏宿迁人,本科,宿迁子渊司法鉴定有限公司,研究方向:法医物证司法鉴定。