



混合生物检材的拆分策略及其法医学应用*

高 静^{1,2,3)} 臧丽丽⁴⁾ 王春年^{1,2)} 贾 镇^{1,2)} 刘 京²⁾ 江 丽^{2)**} 李彩霞^{1,2)**}⁽¹⁾ 中国人民公安大学侦查学院, 北京 100038;⁽²⁾ 公安部鉴定中心, 法医遗传学公安部重点实验室, 北京市现场物证检验工程技术研究中心,
现场物证溯源技术国家工程实验室, 北京 100038;⁽³⁾ 金昌市公安局金川分局, 金昌 737100; ⁽⁴⁾ 甘肃省公安厅刑事侦查总队, 兰州 730046)

摘要 混合生物检材中的不同贡献者拆分是法医遗传学领域长期面临的核心技术挑战之一, 拆分准确性直接影响案件侦破效率与司法公正性。本文以是否需要依托拆分软件为界限, 系统梳理了现有的两类混合生物检材拆分策略: 一是基于生物实验的检材前处理策略, 通过显微可视化分选进行不同来源细胞的物理分离与独立检测, 或通过特异性物理化学处理在检测前实现区分; 二是基于计算生物学的数据后处理策略, 借助生物信息学算法对混合DNA检测数据进行拆分, 从而实现贡献者组分的精准区分。在此基础上总结了两类策略的学术争议及优劣对比, 评估其在法医学实践中的适用场景及局限性, 并结合基层实战中检材条件多样、检验资源有限等特点, 提出了基于样本质量评估的分级应用策略及多策略联合应用路径。最后, 对未来混合检材快检设备的发展路径和人工智能在混合拆分中的应用前景进行展望, 以为混合生物检材的精准解析提供系统的方法学参考。

关键词 混合拆分, 单细胞分离, 单细胞测序, 概率分型, 法医遗传学

中图分类号 R89, D919.2

DOI: 10.3724/j.pibb.2026.0335

CSTR: 32369.14.pibb.20260335

法医现场混合生物检材是指在案件现场提取的、包含多位贡献者的生物检材, 如何精准拆分并获得每位贡献者的单一基因型(即分型), 长期以来都是司法实践中的难点和痛点。不同于单一贡献者来源的生物检材可直接检测获得单一分型, 混合检材的检测结果需要对包含的不同贡献者进行推断拆分, 从而确定贡献者的数量与各自的分型。但混合拆分存在不确定性, 其原因是不同贡献者的基因座检测信号间常常重叠, 同时受到等位基因丢失、插入、峰高不平衡和扩增变异性等随机效应的影响^[1]。这不仅限制了生物检材在案件侦破中的线索支撑作用, 更难以满足诉讼过程中对证据明确性、排他性的严格要求。加之近年来仪器灵敏度不断提高, 从生物检材中检出混合分型的占比显著提高, 而随着贡献者人数的增加, 拆分的复杂性和难度呈指数级上升^[2]。为了突破这一困境, 研究者依托多种遗传标记, 包括短串联重复序列(short tandem repeat, STR)、单核苷酸多态性(single

nucleotide polymorphism, SNP)、微单倍型(microhaplotype, MH)、插入/缺失多态性(insertion-deletion polymorphism, DIP)、编码区SNP(coding region SNP, cSNP)、线粒体DNA(mitochondrial DNA, mtDNA)以及DNA甲基化(DNA methylation, DNAm)等, 提出了多种混合拆分方法。本文根据是否需要依托拆分软件为界限, 系统梳理了现有两类策略(图1)^[3]: 其中基于生物实验的检材前处理策略可通过显微可视化分选对不同来源的细胞进行物理分离与独立检测, 或特异性物理化学处理在检测前实现区分; 而基于计算生物学的数据后处理策略则是借助生物信息学算法对混合DNA检测数据进行拆分, 包括概率基因

* 中央级公益性科研院所基本科研业务费专项资金(2025JB031, 2026JB025)资助项目。

** 通讯联系人。

江丽 Tel: 010-83752707, E-mail: jl@mail.bnu.edu.cn

李彩霞 Tel: 010-83752706, E-mail: licaixia@tsinghua.org.cn

收稿日期: 2026-06-17, 接受日期: 2026-10-08

分型软件 (probabilistic genotyping software, PGS) 及单细胞测序混合拆分软件, 本文现对这两类拆分

策略及其在法医学实践中的应用进行综述。

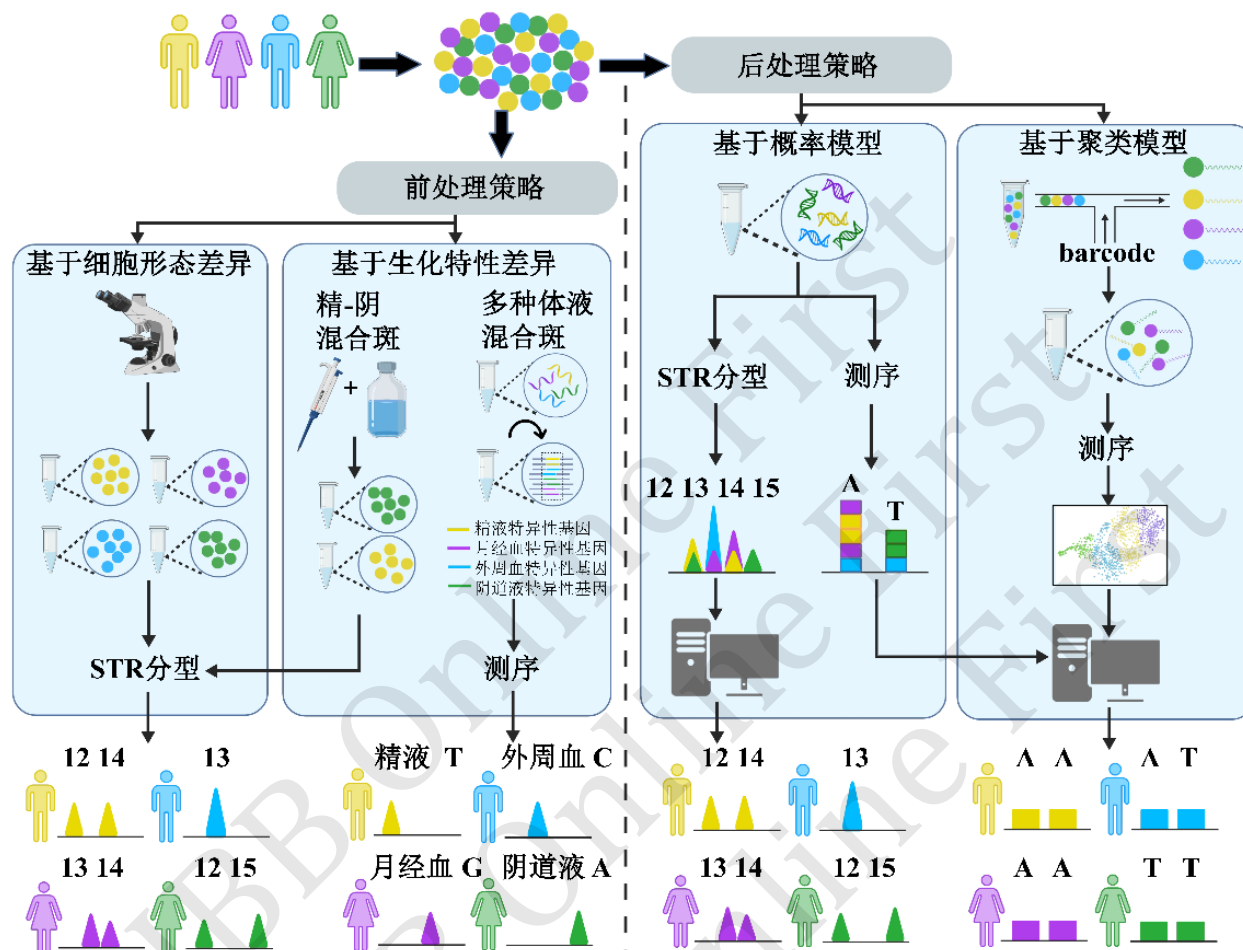


Fig. 1 Schematic diagram of two deconvolution strategies

图1 两种拆分策略示意图

1 前处理策略及法医学应用

在混合生物检材的前处理策略中, 如何从不同来源的细胞或核酸中准确获取单一组分的遗传信息, 是决定下游检测成败的关键。目前主要有两种解决方案, 一种是利用细胞自身的物理形态将其逐一分离; 另一种是在检测前通过化学试剂或特异性引物进行实验设计, 使得不同来源的遗传信息在后续检测结果中得到区分。

1.1 基于细胞形态差异的物理拆分

这种拆分方式是对混合检材进行不同形态细胞的物理分离与选择, 再对分选后的细胞进行检测,

适用于能够辨认出完整细胞形态的生物检材, 即“细胞分选+检测”。为了保证后续分析的准确性和可靠性, 通常采用中低通量而精准度高的细胞捕获技术, 借助显微镜等可视化手段, 通过显微吸针操作、激光捕获显微切割技术 (laser capture microdissection, LCM)、光镊技术和介电泳分离 (dielectrophoresis, DEP) 技术等进行细胞的分选。

其中显微吸针操作技术, 依靠玻璃吸针装置在显微镜下完成细胞的捕获与转移, 操作过程高度依赖技术人员的经验与熟练度。通过这种方式, 曾成功在性侵案件中从受害人皮肤与嫌疑人口腔上皮细胞混合物中分离出嫌疑人细胞^[4-5]。Huffman等^[6]

提出了基于立体显微镜和钨丝针进行细胞分选的细胞亚采样 (direct single cell subsampling, DSCS) 概念, 将混合细胞拆分为单个或少量细胞的子样本进行检测以获得单一结果。用机械臂替代人工操作的显微操作技术, 也可成功捕获口腔上皮细胞、精子细胞和指纹中的遗留细胞^[7]。这种方式的局限性在于高度依赖人工、速度较慢, 虽可通过自动化操作系统提高效率, 但需要自动化设备的支持。

LCM 能够从组织切片或新鲜冰冻样本中精准切取目标细胞, 并保留其原位空间信息^[8-9], 用于分离精液-阴道分泌物混合斑^[10-12] (以下简称精-阴混合斑)。光镊技术则是借助颗粒或细胞之间在形状、尺寸和折射率等方面的差异, 利用所受光力的不同来实现分选的目的^[13], 同样用于精-阴混合斑中精子细胞的分离^[14-15]。从案例来看, LCM 仅在精-阴混合斑这类高浓度的检材中有所应用, 而尚未扩展到其他类型的检材。

采用“显微成像+介电泳”技术的 DEPArray™ 平台^[16], 利用微电极阵列产生的非均匀电场形成电笼并捕获细胞, 并结合高清荧光显微镜将细胞的形态特征可视化从而进行分选, 因此适合分选具有不同外观形态的细胞^[16-17], 例如血液和口腔上皮细胞^[18-19]、精子和阴道上皮细胞^[20]。但该平台耗材成本高, 单次只能做一份检材, 不适合日常案件的常规检验。

此外, 荧光激活细胞分选 (fluorescence activated cell sorting, FACS) 是流式细胞术的一种特殊应用, 根据细胞特征和表面的荧光信号将特定细胞从群体中分离出来, 虽有 FACS 结合免疫荧光从多个贡献者的精子细胞混合物中分离出不同贡献者的案例^[21], 但通常需要大于 1 万个细胞作为起始输入, 因此很难用于微量样品^[22]。

1.2 基于生化特性差异的化学拆分

化学分离策略是通过添加化学试剂或设计特异性分子引物, 在检测前对不同贡献者来源的细胞或核酸进行差异化处理, 经过后续的扩增和检测, 最终从检测结果中直接区分不同来源的遗传信息。

长期以来在性侵案件中采用的差异裂解法^[23]是一种经典的化学拆分, 利用女性上皮细胞核和男性精子细胞核的生化差异, 依靠多种试剂进行先后

裂解和检测, 实现精-阴混合斑中男女成分的分别检出。虽然仅适用于精液与非精液的二元混合场景, 但因试剂价格低廉、实验操作简便, 且不需要引入额外的设备, 目前在基层检案中仍是主流方法。该方法的不足之处在于男女比例失衡、不同组分消化时间差异等因素易造成检出男女混合分型或未检出有效分型。

另一种化学前处理方法是直接作用于分子层面, 通过特异性识别并捕获特定体液的分子标记, 实现对混合斑中不同贡献者的拆分。该技术的核心原理是基于人体基因表达的组织特异性, 不同体液来源的细胞具有特异性信使 RNA (messenger RNA, mRNA) 转录本, 可作为识别混合斑中相应体液成分的分子标记^[24], 且这些 mRNA 分子上携带的 cSNP 能够真实反映贡献者的基因型^[25]。因此, 提取混合斑 RNA 并逆转录为 cDNA 后, 针对目标体液的多个特异性 mRNA 标志物设计引物, 通过 PCR 扩增并检测这些转录本上的 cSNP, 即可实现体液来源鉴定与个体识别的一体化分析。张更谦团队利用 SNaPshot 技术, 先后开发了针对血液^[26]、精液^[27], 以及精-阴混合斑^[28]的分型体系, 最高能拆分 1: 10-1: 500 的混合比例。在应用大规模平行测序后, 该团队进一步建立了包含 202 个 cSNP 分型体系, 可同时对外周血、唾液、精液、阴道分泌物、月经血和皮肤 6 种体液/组织进行分型^[29]。与此同时, Hanson 等^[30]也建立了相同的 6 种体液/组织分型体系, 包含 46 个 cSNP, 在欧洲人群中能达到 0.997 的鉴别能力。李成涛团队^[31]首次将编码区插入/缺失多态性 (coding region insertion/deletion polymorphisms, cInDel) 纳入体液特异性 mRNA 分型体系, 进一步拓展了该方法的遗传标记资源。目前该方法仅验证了上述研究提到的 6 种体液/组织, 尚未扩展到其他检材类型。

除此之外, 复合标记也侧重于通过复合 STR、SNP、DIP 等多种遗传标记的实验设计方案, 利用不同标记在突变率、多态性、甲基化修饰等维度的互补优势, 实现不同贡献者的直接拆分 (表 1), 但在法医检案中的应用仍需进一步体系验证和标准化推进。

表1 复合标记特征与挑战总结

Table 1 Summary of features and challenges of compound markers

复合标记	适用特点	核心优势	主要挑战
STR-SNP ^[32]	序列信息指导STR拆分	次要贡献者检出比例可达1: 100	影子峰影响次要贡献者分型解释
DIP-SNP ^[33]	DIP筛选, 靶向扩增SNP	对降解检材友好	单个位点遗传多态性有限
DIP-STR ^[34]	DIP筛选, 靶向扩增STR	次要贡献者检出比例可达1: 1 000	需要更多群体数据支撑
DIP-MH ^[35]	DIP筛选, 靶向扩增MH	对降解且混合比例1: 100-1: 1000的样本较为适用	需要更多群体数据支撑
CpG-MH/ SNP ^[36]	体液来源与个体识别直接关联	可在1: 1等比例混合样本中排除非目标组织等位基因	建模复杂、标记数量不足、甲基化个体差异影响稳健性
DNAm-STR/ SNP ^[37]	体液来源与个体识别联合分析	可同步分析甲基化与STR/SNP	生信分析复杂, 甲基化个体差异影响稳健性

2 后处理策略及其法医学应用

在这种拆分策略中, 先对混合生物检材进行整体检测, 而后通过算法模型对混合分型数据进行拆分, 将混合分型结果还原为各个单一贡献者独立的分型, 即“混合检测+数据拆分”。

2.1 基于概率模型的拆分

为了从混合检测数据中精确拆分出每个贡献者的成分, 研究者们开发出概率基因分型软件 (probabilistic genotyping software, PGS), 引入统计学模型对混合数据进行客观解释, 最终以似然比 (likelihood ratio, LR) 来量化评估不同假设的可信度。目前, 主流的PGS均采用完全连续模型, 主要针对混合STR图谱, 通过模拟PCR扩增和荧光检测的生物学过程, 将峰高、影子峰比例、降解曲线等影响因素纳入统一的概率框架, 从而对观测到的荧光信号进行定量解释。在这个模型框架下, 主要的算法路径包括马尔可夫链蒙特卡罗方法 (Markov chain Monte Carlo methods, MCMC) 和最大似然估计 (maximum likelihood estimation, MLE), 表2归纳了目前主要的PGS。EuroForMix作为一款开源软件在全球范围内得到了广泛应用, 例如葡萄牙科学警察实验室^[38]、巴西国家刑事科学研究所^[39]、德国法医实验室^[40]均有应用报道, 并在哥伦比亚的一起涉及3人混合血样的案件中提供了关键证据^[41]。国内PGS的研发工作虽然起步较晚, 2020年以后陆续开发了针对STR和MH的PGS, 形成国际并跑局面。尤其是自主研发的SMART系统, 可实现2-5人混合STR图谱的高灵敏度、高特异性自动化拆分与似然比评估, 已为基层办理系列盗窃、强奸致孕等案件提供了侦查线索和证据支撑^[42-43]。这些案例表明, 概率分型软件已

从实验室研究走向一线实战, 成为基层法医解决混合拆分难题的有效工具, 但其拆分可靠性受扩增质量、预设参数等多种因素制约, 尚需人工复核与专业判断作为必要补充。

近年来, 可软件拆分的遗传标记已从STR、SNP扩展至mtDNA、DNAm等多种标记。针对mtDNA开发出基于期望最大化算法的Mixemt^[64]、基于贝叶斯统计定相的MMDIT^[65]等拆分工具, 针对DNAm则依据总甲基化水平与各贡献者甲基化水平的加权之和的线性叠加关系, 通过数学建模反推贡献者比例^[66]。然而在真实应用中, mtDNA的异质性、数据库局限, 以及DNAm的组织特异性、年龄效应, 均会对其理想建模前提造成影响, 不利于向实战转化。

2.2 基于聚类模型的拆分

基于聚类模型的数据拆分依托近年来飞速发展的单细胞测序技术, 其核心是通过高通量的微流控液滴或微孔技术将混合样本中各个细胞与带有独特标签 (barcode) 的磁珠或引物精准地包裹在同一反应单元中, 即对单个细胞进行物理分离并添加标记。但这种在实验阶段添加的细胞独特标记并不能直接实现混合拆分, 而是为后续的拆分软件提供数据依据。对所有细胞进行测序并按标签拆分获得每个细胞数据后, 依赖软件通过生物信息学方法, 将相似细胞进行聚类从而还原不同贡献者的数据。

目前, 单细胞基因组测序 (single-cell DNA sequencing, scDNA-seq) 还存在扩增瓶颈^[67], 在法医领域具有应用潜力的单细胞测序技术主要包括单细胞转录组测序 (single-cell RNA sequencing, scRNA-seq) 和单细胞染色质可及性测序 (single-cell ATAC sequencing, scATAC-seq)。这类测序技术通过识别不同细胞来源的差异SNP将具有相似

表2 概率分型软件总结

Table 2 Summary of probabilistic genotyping software

软件名称	发表时间	核心原理	贡献者数量	适用场景	软件优势
STRmix ^[44-45]	2013年	MCMC	2~6人 ^[46]	STR	带有上游数据处理和下游分析套件
EuroForMix ^[47-48]	2016年	MLE+贝叶斯方法	2~6人 ^[33]	STR&SNP	具有高度可扩展性
Kongoh ^[49]	2017年	MLE	2~4人	STR	自动估算贡献者人数
*云算GPM混合图谱分析系统 ^[50]	未明确提及	MCMC	2~5人	STR	支持导入HID或FSA格式的原始数据
*SMART ^[51]	2021年	MCMC	2~5人	STR	可直接利用混合图谱进行数据检索
AH-HA ^[52]	2022年	隐马尔可夫模型	2人	SNP	对低深度测序数据鲁棒
EFMrep ^[53]	2022年	EuroForMix的扩展模型	2~4人 ^[40]	STR	允许联合分析多份样本
STRmix NGS ^[54]	2023年	STRmix的NGS版本	同STRmix ^[55]	STR&SNP ^[56]	相较于STRmix有更高的似然比值 ^[55]
*SinoMixID ^[57]	2025年	MLE	2~4人	STR	可分析亲缘混合DNA及多图谱联合分析
EFMex ^[58]	2025年	EuroForMix的扩展模型	3~4人	STR	一次性评估所有可能的贡献者组合,并排除无关亲属
MixDeR ^[59]	2025年	EuroForMix的扩展模型	2人	SNP	专注两人混合
MPSproto ^[60]	2025年	EuroForMix的扩展模型	2~3人	STR&MH	STR与MH的联合解析
*Truncated Gaussian模型 ^[61]	2025年	截断高斯模型	2~3人	MH	对次要贡献者的推断准确性更高 ^[62]
*SMART-MHmix ^[63]	2026年	MCMC	2~5人	MH	SMART平台下专门针对MH的子模型

*国产软件。

遗传特征的细胞进行聚类,从而实现混合拆分。根据所利用的变异类型不同,现有的单细胞测序混合拆分软件主要利用核基因组 SNP 或线粒体 SNP (表3),输入数据多为BAM格式数据文件和标签列表,对于部分需要已知样本基因型作为先验信息的算法,还需提供相应的VCF格式数据文件作为输入。然而,现有方法大多需要预先提供混合样本中贡献者的数量或其基因型,这在法医学实践中往往难以满足。针对这一法医学应用需求,研究者开

发了de-goulash^[68-69],综合利用两种SNP,先后实现了多至11人等比例混合、2人不等比例混合血液样本的拆分,并可推断贡献者的性别、族群地域、Y染色体和线粒体单倍群信息,为混合检材的法医鉴定提供了技术方向。但单细胞测序技术对检材中的细胞数量及活性要求较高,如何降低该技术的检测门槛、扩大其在法医领域的适用范围,是目前限制该技术实践应用的瓶颈问题。

表3 单细胞测序混合拆分软件总结

Table 3 Summary of software for mixture deconvolution using single-cell sequencing

软件名称	样本限制	预设贡献者数量	适配数据类型	技术短板
Demuxlet ^[70]	核基因组SNP	需要	scRNA-seq	需要贡献者已知基因型
Freemuxlet ^[70]	核基因组SNP	需要	scRNA-seq	易用性、精确度和稳定性略低 ^[71]
Vireo ^[72]	核基因组SNP	需要	scRNA-seq&scATAC-seq	需要每个细胞的已知基因型
scSplit ^[73]	核基因组SNP	需要	scRNA-seq	拆分召回率和精确度略低 ^[71]
Souporcell ^[74]	核基因组SNP	需要	scRNA-seq	计算资源需求较高
scAVENGERS ^[75]	核基因组SNP	需要	scRNA-seq&scATAC-seq	计算资源需求较高
*mitoSplitter ^[76]	线粒体SNP	需要	scRNA-seq	仅适配包含线粒体的数据
*MitoSort ^[77]	线粒体SNP	需要	scRNA-seq&scATAC-seq	仅适配包含线粒体的数据
de-goulash ^[68-69]	二者皆需要	不需要	scRNA-seq&scATAC-seq	仅适配包含线粒体的数据
Souporcell3 ^[78]	核基因组SNP	需要	scRNA-seq	计算资源需求较高

*国产软件。

3 策略对比及应用思路

3.1 混合拆分的学术争议

作为法医遗传学领域持续关注核心难题，混合拆分的学术争议主要集中于如何在实现低模板、高复杂度混合拆分的同时兼顾灵敏度与可靠性。第一是次要贡献者的可检测性。DNAmix2021实验室间研究揭示了当贡献者DNA占比低于8% (≤ 25 pg) 时，PGS的错误排除率 and 不确定率显著升高^[79]。因此对于证据强度不足的检材应谨慎放弃分析，还是继续借助PGS开展拆分，目前尚无统一标准。第二是贡献者数量的估计偏差。低估贡献者数量可能将次要贡献者归为噪声从而被错误排除，高估则可能导致错误包含，虽可引入人工智能辅助推断，但当人工智能参与DNA拆分决策时，如何满足法庭对专家证言“可解释性”的要求仍面临挑战。第三是低遗传多样性导致的假阳性。有研究表明，在遗传多样性低、近交系数高的群体中，更易产生错误包含，且这种效应随贡献者人数增加而放大^[80]。如果分析结果表现出明显的群体差异性，PGS应用的群体普适性可能会受到质疑。第四是完全连续模型的计算复杂性。完全连续模型的计算复杂度极高，即使数学原理公开，具体实现细节仍难以被独立审查。同时，输出结果对先验参数高度敏感，导致不同实验室即使分析相同数据，也可能产生显著差异的似然比^[81]。在尚未建立统一校准标准与透明验证机制的现状下，似然比结果能否作为稳定、客观的法庭科学证据仍存有争议。

3.2 两类策略的对比分析

总体而言，两类拆分策略的根本差异是拆分发生的环节与拆分依赖的信息来源。前处理策略的核心价值是“在检测前通过实验手段辨别混合”，以实验环节的精准投入换取检测结果的确定性。对物理拆分方法来说，一方面难点在于目标细胞的寻找和定位，不具备完整形态的细胞无法进行分选。另一方面，人的单个二倍体细胞DNA含量仅为6.6 pg^[82]，分选细胞过少将导致DNA浓度太低不足以获得分型结果。解决方法除了增加分选的细胞数量，也可以全基因组扩增 (whole genome amplification, WGA) 以获得足量的DNA模板，但目前常用的简并寡核苷酸引物聚合酶链式反应 (degenerate oligonucleotide-primed polymerase chain reaction, DOP-PCR)、多重置换扩增 (multiple displacement amplification, MDA)、多次退火环状

循环扩增 (multiple annealing and looping based amplification cycles, MALBAC) 等WGA技术，易产生扩增偏倚，制约了细胞水平检材的法医学应用。而对基于cSNP的化学拆分方法来说，拆分效能受限于RNA的高降解性，目标转录本检出率的降低直接削弱了该方法对陈旧检材的适用性。此外，前处理策略的弊端还在于对同种类型的细胞或体液无法区分，仍须结合其他技术手段，如Y-STR^[83]、原位杂交^[84]、免疫荧光^[85]或PGS^[86]等，才能实现混合拆分。

后处理策略的核心价值则是“在检测后通过计算手段解析混合”，以算法的拆分能力应对检测结果的复杂性。其突出优势是对检材高度包容，无论细胞是否完整、类型是否一样、是否微量或降解，只要能检测出混合结果，即可进行计算拆分。不足之处是随着混合人数增加或混合比例变化，算法拆分的不确定性也随之增加，影响拆分结果的准确性。法医实践中通常难以获取贡献者的先验信息，因此混合拆分能力不仅体现在拆分的准确性，同时也体现在能否自动准确判断贡献者数量。大多数软件在运行前需预设贡献者数量，且不同的贡献者数量会影响最终结果^[87-88]。针对概率模型的拆分方法，有学者利用支持向量机、随机森林、逻辑回归、决策树、多层感知机等机器学习方法判断贡献者数量，但容易受到复杂背景噪声的影响^[89]。2026年Taylor等^[90]开发出基于深度神经网络的deepNoC，基于STR数据判断贡献者数量，在分别包含1~10名贡献者的样本中判断准确性可达89%。而基于单细胞测序技术的贡献者数量推断具有独特优势，现有R程序包NbClust^[91]、mclust^[92]、HDBSCAN^[93]等可通过计算最佳聚类数直接推断贡献者数量，但对细胞群边界模糊、难以区分的数据集，其推断精度存在一定误差。两类拆分策略在应用层面的详细对比见表4。

3.3 法医实践应用思路

在实际案件中，二者的关系不是非此即彼的排他性选择，而应当根据检材的保存状况、混合检材类型、可用的实验与计算资源，构建前后融合处理策略的分析框架，以最适成本获取最可靠的拆分结果。目前法医鉴定工作开展多依托STR，本文基于上述混合拆分策略凝练形成法医实践应用思路 (图2)。针对精-阴混合斑，首选差异裂解法，如有多个同性贡献者，则进一步对差异裂解法获得的结果进行后处理PGS拆分。而对于其他类型的常量、

表4 两类混合拆分策略的对比分析

Table 4 Comparison of the two deconvolution strategies

拆分策略	前处理策略		后处理策略	
	物理拆分	化学拆分	概率拆分	聚类拆分
适用检材类型	需具有完整细胞形态	血液、唾液、精液、阴道分泌物、月经血、皮肤	无限制	新鲜血液
降解检材适用性	不适用	不适用	适用	不适用
微量检材适用性	适用	适用	适用	不适用
是否依赖先验信息	不依赖	依赖	大部分依赖	大部分依赖
实战普及度	较低	低	中等	仍在研究阶段
实战应用限制因素	依赖特定仪器设备、 配套耗材价格昂贵	尚未形成标准方法	国内软件不开源、 国外软件普及度低	检材类型局限、 检测价格昂贵

保存条件较好的混合检材，可直接采用“DNA 提取-PCR 扩增-STR 检测-后处理 PGS”的常规流程（简称“常规 STR-PGS 流程”）。对于微量、陈旧、保存条件较差的疑难检材，可先在显微镜下观察是否有完整形态的细胞。如果没有，可尝试采用常规 STR-PGS 流程；如果有，则可分离细胞，并利用线性扩增（linear amplification via transposon insertion, LIANTI）^[94] 或引物模板定向扩增（primary template-directed amplification,

PTA）^[95-96] 进行 WGA，这两种方法在覆盖均匀性和等位基因检出率方面较其他 WGA 方法有所改善。获得足量的 DNA 模板后，再进行 PCR 扩增和 STR 检测，提升疑难检材中单一贡献者 STR 的检出率。已有案例报道通过 MDA 从单精子中成功检出 STR 分型^[97]。此外，还可对 WGA 获得的 DNA 进行 SNP 检测，对贡献者进行表型特征刻画，深挖线索服务破案。

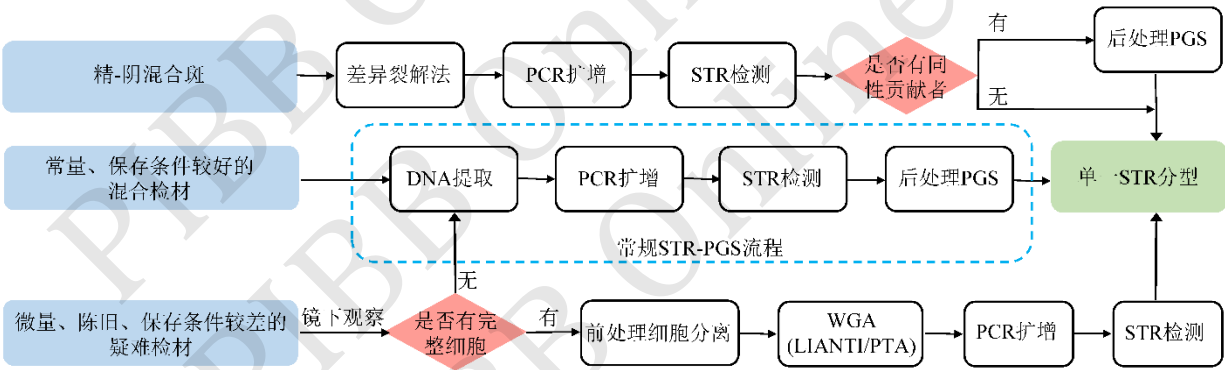


Fig. 2 Schematic diagram of the application strategy in forensic practice

图2 法医实践应用思路示意图

4 总结与展望

本文综述了现有混合拆分策略的流程原理、法医学应用，对两类策略进行了对比分析，并提出了法医实践应用思路。尽管目前已实现混合生物检材的有效拆分，但现有概率模型在贡献者数量增加、混合比例失衡时拆分准确性下降，同时真实案件中现场检材的复杂性使得方法面临稳健性、时效性的考验。如果能进一步提升对复杂混合检材的即时拆分能力，将为基层实战提供更为有效的技术支撑。

4.1 混合检材快检设备的发展路径

当前快检设备的发展主要沿两条技术路径推进。其一为微流控芯片集成化，通过将细胞分选、核酸扩增与产物检测集成于一张芯片，目前已在血卡、腐败肌肉、带毛囊毛发、精斑、烟蒂、胚胎组织等6种检材的分别检测中实现“样本进-结果出”的全流程自动化^[98]，在此基础上，将 PGS 模块内置于快检设备中，即可将现场获取的多重电泳信号直接进行混合拆分，从而为现场混合检材的快速检测分析提供可能；其二为便携式纳米孔测序的现场

化, 例如便携式 MiniION 测序仪^[99], 与可在普通笔记本电脑安装使用的 STRspy2.0 分析软件^[100] 及 PGS 联合使用, 使得基于碱基序列的混合 STR 拆分在资源有限的环境中具备可行性。上述路径目前尚无法完全取代实验室分析, 但其价值在于大大压缩混合检材信息获取的时间, 推动法医混合分析从被动等待走向快速响应。

4.2 人工智能在混合拆分中的应用前景

当前概率基因分型软件在贡献者人数未知、混合比例悬殊等复杂场景下, 拆分准确性仍欠佳。深度学习因具备从数据中自动捕获复杂关联特征的能力, 可以突破传统模型依赖人工预设公式和基因座独立假设的局限, 提升混合拆分准确率与复杂图谱解析能力, 因此将深度学习引入混合拆分是目前主要研究方向之一。例如基于距离加权自注意力机制 (distance-weighted self-attention, DWSA) 的深度学习模型, 是针对混合 STR 图谱分析提出的一种多模态融合算法, 可自动完成贡献者人数推断与混合比例估计^[101]。此外, 多模态大语言模型在法医领域的视觉信息理解与跨模态推理方面的能力已得到初步验证^[102], 虽主要面向图像与文本的多模态融合, 但其在复杂信息整合与端到端学习方面的技术架构, 同样为混合 STR 图谱拆分提供了可借鉴的技术路径。

参考文献

- [1] Zhu Q, Mao Z, Zhang J. DNA mixture deconvolution: a four-strategy framework from physical separation to database searching. *Genes*, 2026, **17**(4): 434
- [2] Butler J M, Iyer H, Press R, *et al.* DNA mixture interpretation: a NIST scientific foundation review: NIST IR 8351. Gaithersburg, MD: National Institute of Standards and Technology, 2024
- [3] Jiang S, Li H, Zhang L, *et al.* Generic Diagramming Platform (GDP): a comprehensive database of high-quality biomedical graphics. *Nucleic Acids Res*, 2025, **53**(D1): D1670-D1676
- [4] Li C, Qi B, Ji A, *et al.* The combination of single cell micromanipulation with LV-PCR system and its application in forensic science. *Forensic Sci Int Genet Suppl Ser*, 2009, **2**(1): 516-517
- [5] Li C X, Wang G Q, Li W S, *et al.* New cell separation technique for the isolation and analysis of cells from biological mixtures in forensic caseworks. *Croat Med J*, 2011, **52**(3): 293-298
- [6] Huffman K, Hanson E, Ballantyne J. Recovery of single source DNA profiles from mixtures by direct single cell subsampling and simplified micromanipulation. *Sci Justice*, 2021, **61**(1): 13-25
- [7] Ostojic L, O'Connor C, Wurmbach E. Micromanipulation of single cells and fingerprints for forensic identification. *Forensic Sci Int Genet*, 2021, **51**: 102430
- [8] Janíček T, Hobza R, Hudzieczek V. Laser capture microdissection: from genomes to chromosomes, from complex tissue to single-cell analysis. *Methods Mol Biol*, 2023, 2672: 163-175
- [9] Guo W, Hu Y, Qian J, *et al.* Laser capture microdissection for biomedical research: towards high-throughput, multi-omics, and single-cell resolution. *J Genet Genom*, 2023, **50**(9): 641-651
- [10] Han J P, Yang F, Xu C, *et al.* A new strategy for sperm isolation and STR typing from multi-donor sperm mixtures. *Forensic Sci Int Genet*, 2014, **13**: 239-246
- [11] 周怀谷, 张晨, 平原. 激光捕获显微切割技术在法医物证检验中的应用. *法医学杂志*, 2004, **20**(4): 231-232
- [12] Zhou H G, Zhang C, Ping Y. *J Forensic Med*, 2004, **20**(4): 231-232
- [12] 俞卫东, 连昌舟, 孙大鹏. 激光捕获显微切割技术在强奸案中的检验研究. *刑事技术*, 2017, **42**(1): 25-27
- [12] Yu W D, Lian C Z, Sun D P. *Forensic Sci Technol*, 2017, **42**(1): 25-27
- [13] 崔丽璇, 杨萌, 陆澄锋, 等. 基于光镊的光学分选理论、技术、应用研究(特邀). *中国激光*, 2025, **52**(15): 157-175
- [13] Cui L X, Yang M, Lu C F, *et al.* *Chin J Lasers*, 2025, **52**(15): 157-175
- [14] Auka N, Valle M, Cox B D, *et al.* Optical tweezers as an effective tool for spermatozoa isolation from mixed forensic samples. *PLoS One*, 2019, **14**(2): e0211810
- [15] Valle M, O'Brien B, Green T D, *et al.* Droplet-based optical trapping for cell separation in mock forensic samples. *J Forensic Sci*, 2024, **69**(1): 273-281
- [16] Schulte J, Caliebe A, Marciano M, *et al.* DEPAArray™ single-cell technology: a validation study for forensic applications. *Forensic Sci Int Genet*, 2024, **70**: 103026
- [17] Schulte J, Marciano M A, Scheurer E, *et al.* A systematic approach to improve downstream single-cell analysis for the DEPAArray™ technology. *J Forensic Sci*, 2023, **68**(6): 1875-1893
- [18] Diepenbroek M, Bayer B, Anslinger K. Pushing the boundaries: forensic DNA phenotyping challenged by single-cell sequencing. *Genes*, 2021, **12**(9): 1362
- [19] Diepenbroek M, Bayer B, Anslinger K. Phenotype predictions of two-person mixture using single cell analysis. *Forensic Sci Int Genet*, 2023, **67**: 102938
- [20] Williamson V R, Laris T M, Romano R, *et al.* Enhanced DNA mixture deconvolution of sexual offense samples using the DEPAArray™ system. *Forensic Sci Int Genet*, 2018, **34**: 265-276
- [21] Xu Y, Xie J, Chen R, *et al.* Fluorescence- and magnetic-activated cell sorting strategies to separate spermatozoa involving plural contributors from biological mixtures for human identification. *Sci Rep*, 2016, **6**: 36515
- [22] Fang W, Liu X, Maiga M, *et al.* Digital PCR for single-cell analysis. *Biosensors*, 2024, **14**(2): 64
- [23] Gill P, Jeffreys A J, Werrett D J. Forensic application of DNA 'fingerprints'. *Nature*, 1985, **318**(6046): 577-579
- [24] Liu Z, Wang J, Li Z, *et al.* mRNA for body fluid and individual identification. *Electrophoresis*, 2025, **46**(1/2): 44-55

- [25] 张更谦, 刘紫东. RNA 多态性个体身份鉴识的方法、应用与展望. 法医学杂志, 2026, **42**(2): 143-150
Zhang G Q, Liu Z D. J Forensic Med, 2026, **42**(2): 143-150
- [26] Liu Z, Gao Z, Wang J, *et al.* A method of identifying the blood contributor in mixture stains through detecting blood-specific mRNA polymorphism. Electrophoresis, 2020, **41**(15): 1364-1373
- [27] Liu J, Cheng X, Liu F, *et al.* Identification of coding region SNPs from specific and sensitive mRNA biomarkers for the deconvolution of the semen donor in a body fluid mixture. Forensic Sci Int Genet, 2021, **52**: 102483
- [28] Zhang X, Li J, Liu J, *et al.* Identification of the vaginal secretion donor in mixture stains using polymorphic cSNPs on mRNA biomarkers. Forensic Sci Int Genet, 2022, **59**: 102703
- [29] Liu Z, Wang J, Li L, *et al.* Identification of the body fluid donor in mixtures through target mRNA cSNP sequencing. Forensic Sci Int Genet, 2024, **71**: 103066
- [30] Hanson E, Dorum G, Zamborlin M, *et al.* Targeted S5 RNA sequencing assay for the identification and direct association of common body fluids with DNA donors in mixtures. Int J Leg Med, 2023, **137**(1): 13-32
- [31] Wang S, Jiang T, Yuan C, *et al.* An mRNA profiling assay incorporating coding region InDels for body fluid identification and the inference of the donor in mixed samples. Forensic Sci Int Genet, 2024, **69**: 102979
- [32] 易少华, 潘超, 黄代新, 等. SNP-STR 标记在不平衡 DNA 混合物次要成分分析中的有限价值. 中国法医学杂志, 2024, **39**(06): 710-716
Yi S H, Pan C, Huang D X, *et al.* Chin J Forensic Med, 2024, **39**(06): 710-716
- [33] Tan M, Xue J, Wu Q, *et al.* Improving DNA mixtures analysis using compound markers composed of InDels and SNPs screened from the whole genome with next-generation sequencing. Electrophoresis, 2024, **45**(5/6): 463-473
- [34] Damour G, Mauffrey F, Hall D. Identification and characterization of novel DIP-STRs from whole-genome sequencing data. Forensic Sci Int Genet, 2023, **64**: 102849
- [35] Liu J, Hao T, Cheng X, *et al.* DIP-microhaplotypes: new markers for detection of unbalanced DNA mixtures. Int J Leg Med, 2021, **135**(1): 13-21
- [36] Park S U, Kim B M, Park S, *et al.* A strategy for body fluid mixture analysis using semen- and epidermis-specific tDMRs and adjacent microhaplotypes/SNPs. Forensic Sci Int Genet, 2026, **82**: 103429
- [37] Schmelzer L, Hoogenboom J, Naue J. Linking STRs/SNPs and DNA methylation using massively parallel sequencing for potential forensic applications. Int J Leg Med, 2026, **140**(2): 621-637
- [38] Costa C, Pereira É, Costa S, *et al.* Stutter modeling in probabilistic genotyping for forensic DNA analysis: a casework-driven assessment. Genes, 2025, **16**(9): 1053
- [39] Minervino A C, de Aragão Gualda C M, Trindade B R, *et al.* Reanalyzing DNA mixture: a evaluation of EuroForMix for deconvolution and weight-of-evidence computing. Forensic Sci Med Pathol, 2025, **21**(4): 1813-1819
- [40] Schiller R, Fleckhaus J, von Máriássy D, *et al.* DNA mixture deconvolution using fully continuous models EuroForMix and EFMrep. Rechtsmedizin, 2025, **35**(1): 22-28
- [41] Gómez N A, Nieto A P R, Chacón G, *et al.* A crime case involving a mixture of three individuals solved using a sample from the daughter of the person of interest. Int J Leg Med, 2025, **139**(4): 1569-1574
- [42] 卢润权, 周家柱. 拆分混合 STR 分型破获入室盗窃案件 1 例. 广东公安科技, 2025, **33**(3): 60-61
Lu R Q, Zhou J Z. Guangdong Gong'an Keji, 2025, **33**(3): 60-61
- [43] 余华光, 黄江平, 刘亚楠, 等. 综合运用混合分型拆分和 Y-STR 破获强奸致孕案 1 例. 法医学杂志, 2025, **41**(6): 612-615
Yu H G, Huang J P, Liu Y N, *et al.* J Forensic Med, 2025, **41**(6): 612-615
- [44] Taylor D, Bright J A, Buckleton J. The interpretation of single source and mixed DNA profiles. Forensic Sci Int Genet, 2013, **7**(5): 516-528
- [45] Bright J A, Taylor D, Curran J M, *et al.* Developing allelic and stutter peak height models for a continuous method of DNA interpretation. Forensic Sci Int Genet, 2013, **7**(2): 296-304
- [46] Bright J A, Richards R, Kruijver M, *et al.* Internal validation of STRmix™ – a multi laboratory response to PCAST. Forensic Sci Int Genet, 2018, **34**: 11-24
- [47] Bleka Ø, González-Bao J, de la Puente M, *et al.* Improved deconvolution of sequenced based mixtures by extending EuroForMix. Forensic Sci Int Genet, 2026, **84**: 103479
- [48] Bleka Ø, Storvik G, Gill P. EuroForMix: an open source software based on a continuous model to evaluate STR DNA profiles from a mixture of contributors with artefacts. Forensic Sci Int Genet, 2016, **21**: 35-44
- [49] Manabe S, Morimoto C, Hamano Y, *et al.* Development and validation of open-source software for DNA mixture interpretation based on a quantitative continuous model. PLoS One, 2017, **12**(11): e0188183
- [50] 李甫, 徐磊, 蔺晨, 等. 两种混合 STR 分型分析软件的应用. 刑事技术, 2022, **47**(1): 18-23
Li F, Xu L, Lin C, *et al.* Forensic Sci Technol, 2022, **47**(1): 18-23
- [51] Ji X, Chi L, Wu L, *et al.* SMART: STR mixture analysis and resolution tools. Forensic Sci Int Genet, 2025, **74**: 103148
- [52] Azhari G, Waldman S, Ofer N, *et al.* Decomposition of individual SNP patterns from mixed DNA samples. Forensic Sci, 2022, **2**(3): 455-472
- [53] Bleka Ø, Prieto L, Gill P. EFMrep: an extension of EuroForMix for improved combination of STR DNA mixture profiles. Forensic Sci Int Genet, 2022, **61**: 102771
- [54] Cheng K, Bright J A, Kelly H, *et al.* Developmental validation of STRmix™ NGS, a probabilistic genotyping tool for the interpretation of autosomal STRs from forensic profiles generated using NGS. Forensic Sci Int Genet, 2023, **62**: 102804
- [55] Duke K, Cuenca D, Cheng K, *et al.* Comparison of key diagnostics for probabilistic interpretation of STR mixture data generated with

- length-based and MPS methodologies. *Forensic Sci Int Genet*, 2027, **86**: 103577
- [56] Cheng K, Bright J A. An evaluation of ForenSeq DNA Signature Prep iiSNP mixture interpretation using a probabilistic genotyping method. *Forensic Sci Int Genet*, 2026, **82**: 103428
- [57] Hu Y H, Yang Q Q, Hou T Y, *et al.* SinoMixID — the establish and validation of a fully continuous probabilistic genotyping system for DNA mixture interpretation. *J Forensic Med*, 2025. DOI: 10.12116/j.issn.11004-15619.12025.550904
- [58] Bleka Ø, Vigeland M D, Gill P. EFMex: using EuroForMix to evaluate DNA mixtures with multiple persons of interest. *Forensic Sci Int Genet*, 2026, **81**: 103365
- [59] Mitchell R, Peck M, Gorden E, *et al.* MixDeR: a SNP mixture deconvolution workflow for forensic genetic genealogy. *Forensic Sci Int Genet*, 2025, **76**: 103224
- [60] Giuffrida M, Rodrigues P, Köksal Z, *et al.* Mixture deconvolution with massively parallel sequencing data: microhaplotypes versus short tandem repeats. *Genes*, 2025, **16**(9): 1105
- [61] Wang Y, Hou T, Zhu Q, *et al.* A continuous model for interpreting microhaplotype profiles of forensic DNA mixtures. *Forensic Sci Int Genet*, 2025, **78**: 103271
- [62] Yang C, Cao Y, Wang Y, *et al.* Transferring a microhaplotype-specific probabilistic genotyping model from targeted amplification to hybridization capture: a proof-of-concept study of a 100-locus panel for DNA mixtures. *Forensic Sci Int Genet*, 2027, **86**: 103549
- [63] Ji X, Feng Y, Chi L, *et al.* SMART-MHmix: a probabilistic model for microhaplotype-based forensic DNA mixture analysis. *Forensic Sci Int Genet*, 2026, **84**: 103509
- [64] Vohr S H, Gordon R, Eizenga J M, *et al.* A phylogenetic approach for haplotype analysis of sequence data from complex mitochondrial mixtures. *Forensic Sci Int Genet*, 2017, **30**: 93-105
- [65] Smart U, Cihlar J C, Mandape S N, *et al.* A continuous statistical phasing framework for the analysis of forensic mitochondrial DNA mixtures. *Genes*, 2021, **12**(2): 128
- [66] Poggiali B, Børsting C, Kampmann M L, *et al.* Age prediction of contributors to two-person DNA mixtures via deconvolution of autosomal DNA methylation profiles. *Forensic Sci Int Genet*, 2026, **83**: 103438
- [67] Estévez-Gómez N, Prieto T, Tomás L, *et al.* Differential performance of strategies for single-cell whole-genome amplification. *Cell Rep Methods*, 2025, **5**(4): 101025
- [68] Kulhankova L, Montiel González D, Bindels E, *et al.* Single-cell transcriptome sequencing allows genetic separation, characterization and identification of individuals in multi-person biological mixtures. *Commun Biol*, 2023, **6**(1): 201
- [69] Kulhankova L, Bindels E, Kayser M, *et al.* Deconvoluting multi-person biological mixtures and accurate characterization and identification of separated contributors using non-targeted single-cell DNA sequencing. *Forensic Sci Int Genet*, 2024, **71**: 103030
- [70] Kang H M, Subramaniam M, Targ S, *et al.* Multiplexed droplet single-cell RNA-sequencing using natural genetic variation. *Nat Biotechnol*, 2018, **36**(1): 89-94
- [71] Fu Y, Youness M, Virzi A, *et al.* Benchmarking of computational demultiplexing methods for single-nucleus RNA sequencing data. *Brief Bioinform*, 2025, **26**(4): bbaf371
- [72] Huang Y, McCarthy D J, Stegle O. Vireo: Bayesian demultiplexing of pooled single-cell RNA-seq data without genotype reference. *Genome Biol*, 2019, **20**(1): 273
- [73] Xu J, Falconer C, Nguyen Q, *et al.* Genotype-free demultiplexing of pooled single-cell RNA-seq. *Genome Biol*, 2019, **20**(1): 290
- [74] Heaton H, Talman A M, Knights A, *et al.* Souporecell: robust clustering of single-cell RNA-seq data by genotype without reference genotypes. *Nat Methods*, 2020, **17**(6): 615-620
- [75] Han S, Kim K, Park S, *et al.* scAVENGERS: a genotype-based deconvolution of individuals in multiplexed single-cell ATAC-seq data without reference genotypes. *NAR Genom Bioinform*, 2022, **4**(4): lqac095
- [76] Lin X, Chen Y, Lin L, *et al.* mitoSplitter: a mitochondrial variants-based method for efficient demultiplexing of pooled single-cell RNA-seq. *Proc Natl Acad Sci USA*, 2023, **120**(39): e2307722120
- [77] Tang Z, Zhang W, Shi P, *et al.* MitoSort: robust demultiplexing of pooled single-cell genomic data using endogenous mitochondrial variants. *Genomics Proteomics Bioinformatics*, 2024, **22**(5): qzac073
- [78] Weerakoon M, Vu H, Behboudi R, *et al.* Souporecell3: robust demultiplexing for high-donor single-cell RNA-seq datasets. *Bioinformatics*, 2026, **42**(3): btg117
- [79] Hicklin R A, Richetelli N, Bever R A, *et al.* Inter-laboratory variation in interpretations, comparisons, and statistical analyses of DNA mixtures. *J Forensic Sci*, 2025, **70**(4): 1292-1314
- [80] Flores M, Ho E, Ly C, *et al.* Decreased accuracy of forensic DNA mixture analysis for groups with lower genetic diversity. *iScience*, 2024, **27**(11): 111067
- [81] Boodoosingh S, Kelly H, Curran J M, *et al.* An inter-laboratory comparison of probabilistic genotyping parameters and evaluation of performance on DNA mixtures from different laboratories. *Forensic Sci Int Genet*, 2024, **71**: 103046
- [82] Huffman K, Ballantyne J. Single cell genomics applications in forensic science: Current state and future directions. *iScience*, 2023, **26**(11): 107961
- [83] Hogg J, Vandepoele A C W, Zaccheo N, *et al.* Targeted recovery of male cells in a male and female same-cell mixture. *J Forensic Sci*, 2024, **69**(4): 1183-1197
- [84] Feng L, Li C X, Han J P, *et al.* Isolating cells from female/male blood mixtures using florescence *in situ* hybridization combined with low volume PCR and its application in forensic science. *Int J Leg Med*, 2015, **129**(6): 1211-1215
- [85] Sun M L, Zheng J L, Wang B J, *et al.* Sperm cell capture based on ABH antigen differences to separate two men in mixed seminal stains. *Biomed Res Int*, 2021, **2021**: 7269237
- [86] Huffman K, Kruijver M, Ballantyne J, *et al.* Carrying out common DNA donor analysis using DBLR™ on two or five-cell mini-mixture subsamples for improved discrimination power in

- complex DNA mixtures. *Forensic Sci Int Genet*, 2023, **66**: 102908
- [87] Costa C, Figueiredo C, Costa S, *et al*. The impact of considering different numbers of contributors in identification problems involving real casework mixture samples. *Int J Leg Med*, 2025, **139** (5): 2099-2107
- [88] Huffman K, Moreno L. Assessing adventitious matches of non-donors related to true contributors. *Biomolecules*, 2025, **15**(3): 398
- [89] 毛哲华, 徐珍, 杨帆, 等. 混合 STR 图谱拆分方法及智能化趋势. *刑事技术*, 2026, **51**(01): 57-66
Mao Z H, Xu Z, Yang F, *et al*. *Forensic Sci Technol*, 2026, **51**(01): 57-66
- [90] Taylor D, Humphries M A. deepNoC: a deep learning system to assign the number of contributors to a short tandem repeat DNA profile. *Forensic Sci Int Genet*, 2026, **83**: 103434
- [91] Charrad M, Ghazzali N, Boiteau V, *et al*. NbClust: an R package for determining the relevant number of clusters in a data set. *J Stat Soft*, 2014, **61**(6): 1-36
- [92] Scrucca L, Fraley C, Murphy T B, *et al*. Model-Based Clustering, Classification, and Density Estimation Using mclust in R. Boca Raton: Chapman and Hall/CRC, 2023
- [93] Hahsler M, Piekenbrock M, Doran D. dbSCAN: fast density-based clustering with R. *J Stat Soft*, 2019, **91**(1): 1-30
- [94] Chen C, Xing D, Tan L, *et al*. Single-cell whole-genome analyses by linear amplification *via* transposon insertion (LIANTI). *Science*, 2017, **356**(6334): 189-194
- [95] Gonzalez-Pena V, Natarajan S, Xia Y, *et al*. Accurate genomic variant detection in single cells with primary template-directed amplification. *Proc Natl Acad Sci US A*, 2021, **118**(24): e2024176118
- [96] Derks L L M, van Leeuwen A J C N, Steemers A S, *et al*. Protocol for genome-wide analysis of somatic variants at single-cell resolution using primary template-directed DNA amplification. *STAR Protoc*, 2025, **6**(1): 103499
- [97] Theunissen G M G, Gibb A, Lin P K T, *et al*. DNA profiling of single sperm cells after whole genome amplification. *Forensic Sci Int Rep*, 2021, **4**: 100240
- [98] Li Y, Zhuang B, Han J, *et al*. A versatile "sample-in-answer-out" microfluidic system for combined detection of multi-functional forensic human genetic markers. *Small*, 2025, **21**(49): e07944
- [99] Luo Y, Zhang J, Ni M, *et al*. Pilot validation of on-field STR typing and human identity testing by MinION nanopore sequencing. *Electrophoresis*, 2024, **45**(9/10): 885-896
- [100] Hall C L, Kesharwani R K, McBroom Henson K E, *et al*. STRspy2.0: unlocking the potential of long reads for forensic DNA profiling. *Int J Mol Sci*, 2026, **27**(4): 1889
- [101] Su Z. A study on hybrid STR profile analysis model based on distance-weighted self-attention mechanism. *TNS*, 2025, **139**(1): 211-220
- [102] Sohail A, Patel O M, Venditti J C S, *et al*. Benchmarking multimodal large language models for forensic sciences: a comprehensive dataset and evaluation framework. *IEEE Access*, 2026, **14**: 68028-68038

Deconvolution Strategies for Mixed Biological Samples and Forensic Applications*

GAO Jing^{1,2,3}, ZANG Li-Li⁴, WANG Chun-Nian^{1,2}, JIA Zhen^{1,2}, LIU Jing²,
JIANG Li²**, LI Cai-Xia^{1,2}**

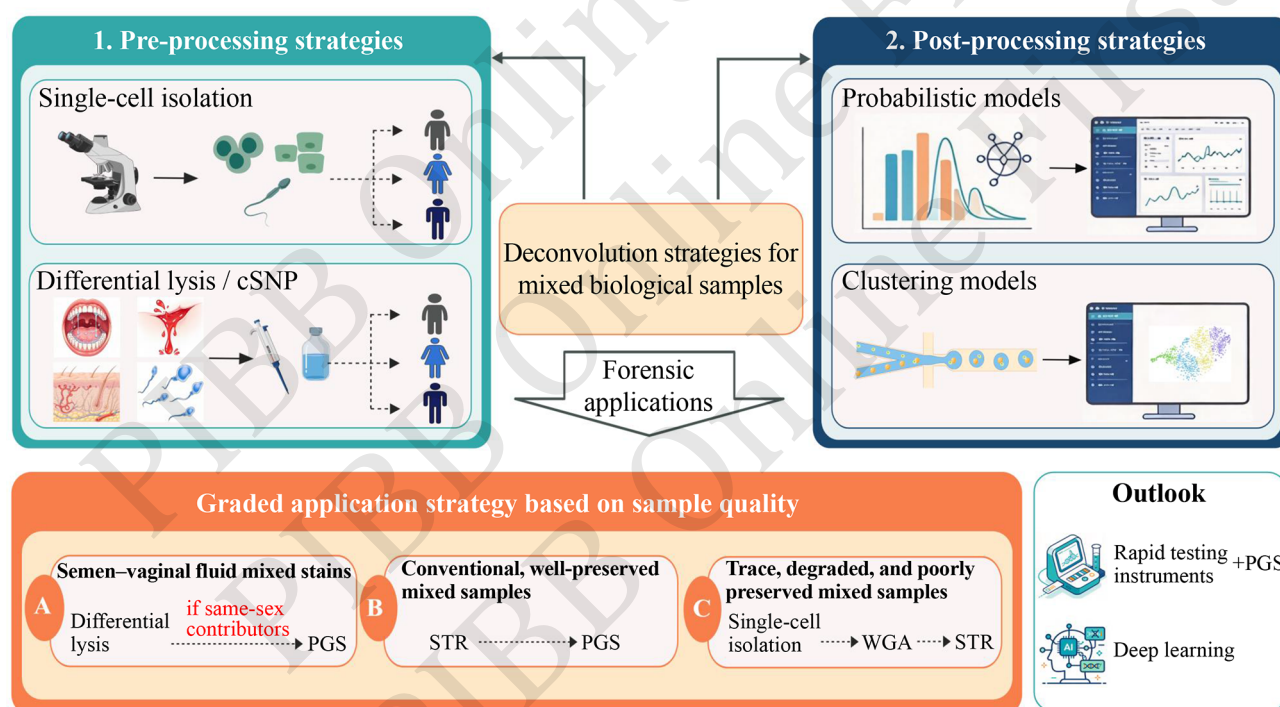
⁽¹⁾School of Investigation, People's Public Security University of China, Beijing 100038, China;

²Institute of Forensic Science, Key Laboratory of Forensic Genetics, Beijing Engineering Research Center of Crime Scene Evidence Examination, National Engineering Laboratory for Forensic Science, Beijing 100038, China;

³Jinchuan Branch of Jinchang Public Security Bureau, Jinchang 737100, China;

⁴Gansu Provincial Public Security Department Criminal Investigation Corps, Lanzhou 730046, China)

Graphical abstract



Abstract The deconvolution of contributors from mixed biological samples represents a long-standing core technical challenge in the field of forensic genetics, as the accuracy of this process directly impacts case investigation efficiency and judicial fairness. Based on whether the deconvolution relies on software as a boundary, this paper systematically reviews the two existing types of strategies for deconvoluting mixed biological samples. (1) Pre-processing strategies based on biological experiments, which achieve physical separation and independent detection of cells from different sources through microscope-guided single-cell isolation, or accomplish differentiation prior to detection *via* specific physicochemical treatments. Among these, physical separation based on differences in cell morphology mainly employs cell capture techniques characterized by low-to-medium throughput but high precision, using visualization tools such as microscopes for single-cell isolation, followed by specialized detection of the isolated cells. Chemical separation based on differences in

biochemical properties mainly targets different body fluids and can directly link a body fluid to its contributing source, including differential lysis, as well as coding-region single nucleotide polymorphisms (cSNPs) and some compound markers. (2) post-processing strategies based on computational biology, which rely on bioinformatics algorithms to analyze mixed DNA testing data, enabling precise discrimination of contributor components. Among these, deconvolution based on probabilistic models mainly depends on probabilistic genotyping software (PGS), currently including software that applies two core algorithms: Markov chain Monte Carlo methods (MCMC) and maximum likelihood estimation (MLE). While clustering-based deconvolution relies on single-cell RNA sequencing (scRNA-seq) and single-cell ATAC sequencing (scATAC-seq), in which each cell is labeled with a distinct barcode, thereby allowing deconvolution of the resulting data and classification of each cell into its cell population of origin. On this basis, this paper summarizes the academic debates, advantages, and disadvantages of the two deconvolution strategies, evaluates their applicable scenarios and limitations in forensic practice, and proposes a graded application strategy based on sample quality assessment, as well as a pathway for the integrated use of the different deconvolution strategies, taking into account the diverse conditions of forensic exhibits and limited resources in routine forensic casework: for sperm-vaginal fluid mixed stains, differential lysis is preferentially adopted to obtain single-source male and female profiles. If same-sex contributors are involved, PGS is further used for deconvolution. For mixed samples that are conventional in amount and well preserved, PGS can be used for post-processing deconvolution after routine STR profiling. For mixed samples that are trace, aged, and poorly preserved, it is necessary to introduce linear amplification via transposon insertion (LIANTI) or primary template-directed amplification (PTA) for whole genome amplification (WGA) on the basis of isolated cells, thereby creating conditions for obtaining STR profiles. Finally, this paper looks forward to the future development of rapid testing instruments for mixed forensic samples and the promising applications of artificial intelligence in mixture deconvolution: the former includes the integration of microfluidic chips, the field deployment of portable nanopore sequencing and its combined use with PGS, while the latter includes the application of artificial intelligence technologies such as deep learning models and multimodal large language models, aiming to provide a methodological reference for the accurate deconvolution of mixed biological samples.

Key words mixture deconvolution, single-cell isolation, single-cell sequencing, probabilistic genotyping, forensic genetics

DOI: 10.3724/j.pibb.2026.0335

CSTR: 32369.14.pibb.20260335

* This work was supported by the Fundamental Research Funds for Basic Scientific Research (2025JB031, 2026JB025).

** Corresponding author.

JIANG Li. Tel: 86-10-83752707, E-mail: jl@mail.bnu.edu.cn

LI Cai-Xia. Tel: 86-10-83752706, E-mail: licaixia@tsinghua.org.cn

Received: June 17, 2026 Accepted: October 8, 2026