



[专家简介] 王艳,解放军总医院医学创新研究部某中心主任,主任技师,硕士研究生导师。中国人体健康科技促进会人类遗传资源保护与利用专业委员会副主任委员,中国合格评定国家认可委员会(CNAS)生物样本库专业技术专家、法庭科学法医物证专业技术评审员、检验检测机构资质认定国家级评审员,全军计划生育优生优育专业委员会委员兼秘书长,北京司法鉴定业协会物证委员会主任委员,中国研究型医院学会临床数据与样本资源库专业委员会常务委员。承担国家重点研发计划课题、国家自然科学基金和军队基金等10余项,主编专著3部,发表学术论文50余篇,获军队医疗成果二等奖等3项奖项,授权发明专利16项。

基于SNP推断表型的法庭科学研究进展与展望

周玲^{1,2}, 康霞¹, 王艳^{1,2}

¹解放军总医院医学创新研究部, 北京 100853; ²山西医科大学法学院, 山西太原 030001

摘要:随着犯罪嫌疑人反侦察能力的增强及检材复杂程度的增加,法医DNA表型推断在学科应用中被发掘,其中单核苷酸多态性(single nucleotide polymorphism, SNP)更是凭借其片段小、检测方法多等优点,在法庭科学中展现出极大发展潜力,有望将法庭科学推向新高度。但由于相关技术及分析方法的不成熟,需要大量的实验研究来提高其精确度,以达到实践运用的标准。本文将从色素特征、身高、面部特征、男性脱发4方面简述SNP推断表型的相关研究进展及其面临的挑战并对未来进行展望,为SNP的法庭科学应用提供参考。

关键词:单核苷酸多态性; 法医遗传学; 表型推断; 法庭科学

中图分类号:R89; DF785.2; D919.2

文献标志码:A

文章编号:2095-5227(2025)01-0037-06

DOI: 10.12435/j.issn.2095-5227.24090801

引用本文:周玲,康霞,王艳.基于SNP推断表型的法庭科学研究进展与展望[J].解放军医学院学报,2025,46(1):37-42.

Research advances and prospect in SNP phenotyping in forensic science

ZHOU Ling^{1,2}, KANG Xia¹, WANG Yan^{1,2}

¹Medical Innovation Research Department of PLA General Hospital, Beijing 100853, China; ²School of Forensic Science, Shanxi Medical University, Taiyuan 030001, Shanxi Province, China

Corresponding author: WANG Yan. Email: 001wangyan@sina.com

Abstract: With the enhancement of the suspect's counter-surveillance ability and the increase of the complexity of the samples, forensic DNA phenotypic inference has been explored in the disciplinary applications, among which single nucleotide polymorphism (SNP) has shown great potential for development in courtroom science by virtue of the advantages of its small fragment size and multiple methods, which is expected to push courtroom science to a new height. However, due to the immaturity of related technology and analysis methods, it needs a lot of experimental research to improve its accuracy to reach the standard of practical application. The progress of SNP inferred phenotypes, the challenges and future prospects of SNPs in terms of pigmentation, height, facial features, male pattern baldness, etc. are discussed in this paper, so as to provide references for the application of SNPs in courtroom science.

Keywords: single nucleotide polymorphism; forensic genetics; phenotypic inference; courtroom science

Cited as: Zhou L, Kang X, Wang Y. Research advances and prospect in SNP phenotyping in forensic science[J]. Acad J Chin PLA Med Sch, 2025, 46(1): 37-42.

收稿日期:2024-09-08

基金项目:法医遗传学公安部重点实验室开放课题(2024FGKFKT07)

第一作者:周玲,硕士。Email:zhou ling990307@163.com

通信作者:王艳,博士,主任法医师,主任技师。Email:001wangyan@sina.com

法庭科学是运用自然科学和社会科学的原理和方法,研究查明事件法律性质、发现犯罪、揭露犯罪、证实犯罪及预防犯罪的科学技术手段与方法的一门综合性应用学科^[1]。自人类基因组计划

开展以来, DNA测序和分析的研究成果层出不穷。研究人员发现单核苷酸多态性(single nucleotide polymorphism, SNP)位点含量丰富, 分布广泛, 在表型特征的推断中起到了重要作用。法医DNA表型推断作为学科的一大热点, 旨在缩小嫌疑人范围, 为案件侦破提供技术依据和线索, 随着测序技术与分析方法的不断完善, 其相关研究也进入了新阶段。本文将从业态特征、身高、面部特征、男性脱发4方面简述SNP推断表型的相关研究进展及其面临的挑战并进行展望。

1 SNP与表型推断的基本概念

单核苷酸多态性主要是指在基因组中特定部位单个碱基序列的变异所引起的DNA序列多态性, 是人类可遗传的变异中最常见的一种, 占有已知多态性的90%以上^[2]。SNP广泛存在于人类基因组中, 约每1 000 bp就存在一个SNP, 在人类基因组中约有 3×10^6 个SNP, 是人类基因组中含量最丰富的DNA多态性。

此外, 研究者们还发现SNP有以下优点: (1)扩增片段可小于100 bp, 能够分析高度降解的生物样本; (2)检测方法众多, 可采用自动化高通量分析技术, 效率高; (3)采用毛细管电泳技术进行等位基因分型数据检测时, 较少存在伪峰, 并可简化等位基因分析; (4)大多为二等位基因, 破译信息相对简单与灵活; (5)可用于推断生物地理祖先及表型特征^[3-4]。因此, SNP也被法医界视为潜在的遗传标记。

人类表型特征包括容貌、身高、发色、肤色、年龄等。表型推断是指通过分析生物样本的遗传信息, 对样本来源的表型特征进行预测^[5]。法医DNA表型推断主要用于分析犯罪现场采集的人类生物样本中提取的DNA, 从而缩小潜在嫌疑人的范围, 为侦破案件、缉拿嫌犯提供依据及线索^[6]。现阶段对犯罪现场遗留DNA推测未知嫌犯表型的技术尚无准确可靠的方法, 因此表型推断在法医工作中具有较好的发掘应用前景。

2 SNP推断表型的技术与方法

基于SNP推断表型主要分为3部分。(1)识别与定位: 运用高通量测序技术快速检测整个基因组中的SNP。(2)与表型的关联研究: 通过全基因组关联研究(Genome-Wide Association Studies, GWAS)等统计方法, 分析SNP与表型之间的关联, 找到与特

定表型显著相关的SNP。(3)构建预测模型: 基于前两者所得, 运用各种算法构建预测模型并对其做风险评估。现行的技术主要有基因芯片技术、高通量测序技术、质谱检测技术、TaqMan探针法、焦磷酸测序技术及基于统计学和机器学习的方法等, 下文将简述时下几种热门的技术与方法。

1975年基于链终止法及光检测法的第一代测序技术问世, 不仅将遗传学相关研究推向了新的高点, 更使得DNA测序技术长期处于技术变革的阵痛中。从高精度再到高通量及低成本, 测序技术不断推陈出新。值得一提的是, 基于三磷酸核苷(nucleoside triphosphate, NTP)的同步试剂洗涤与同步光学检测方法的组合的二代测序技术(next-generation sequencing, NGS)出现, 将DNA测序技术带入了高通量时期。相较一代测序技术, 二代测序技术具有高通量、低成本、快速等优点, 所以也被称为高通量测序技术, 可以同时为数百万甚至数十亿的DNA片段进行测序, 大大提高了SNP的筛查和分析效率, 是目前最主要的测序技术。之后, 为了迎合需求, 第三代测序技术也成功问世, 其中Pacific Biosciences开发的单分子测序技术, 其分辨率具有不可比拟的优势, 因其无需PCR扩增, 对每1条DNA分子进行单独测序, 进一步提高了测序的精度和速度, 该优势使其在特定序列的SNP检测具有优势, 但由于零模波导纳米结构的限制, 其通量有待提高^[7]。同期, Complete Genomics推出了DNA纳米球与图形化陈列结合, 大大提高了测序通量并降低测序成本, 但限制于其极短的读取长度, 可能会出现完整基因的组装受阻的情况。除此以外, Ion Torrent技术使用离子敏感场效应传递器来消除对测序事件的光学检测需求、Oxford Nanopore通过测量纳米孔的电导率变化来消除光学元件以及测序设计中对DNA扩增的需求等, 也不断推动了新型DNA测序技术的发展^[8]。

全基因组关联研究是分析SNP与表型之间关联的关键方法。GWAS通过比较大规模样本群体中个体间的基因组, 寻找与特定表型显著相关的SNP。随着样本量的增加和分析技术的进步, GWAS能够发现更多具有统计学意义的SNP, 这些SNP可以解释更大比例的表型变异^[9]。然而GWAS通常只能解释部分表型遗传变异, 存在“遗传力缺失”的问题及假阳性结果。在此基础上, 多基因风险评估(polygenic risk score, PRS)可以确定表型相关SNP位点的效力大小, 综合评估个体的遗传风险, 在构

建预测模型时发挥一定作用,但其较依赖于GWAS数据质量和样本代表性的影响^[10]。此外,常用的SNP推断表型的方法还有孟德尔随机化(Mendelian randomization, MR)、基因表达数量性状位点分析等。

随着学科发展,越来越多的技术被应用于法庭科学推断表型中,其中机器学习在实际应用中表现良好。其在构建基于SNP位点的表型预测模型中起着至关重要的作用。通过大量的基因组数据和表型数据,机器学习算法能够学习和提取复杂的模式和关系。常用的算法包括回归分析、决策树、随机森林和深度学习等^[11]。此外,机器学习方法还可以对预测模型进行优化和验证,评估模型的性能和准确性,确保其在实际应用中的可靠性。

3 SNP推断表型的研究进展

3.1 色素特征

据评估,人类色素特征性状的遗传力在80%以上,如肤色、发色和眼睛颜色,将已知的色素特征性状进行分类,并结合多项式逻辑回归分析单核苷酸多态性与色素特征性状的关联性,从而构成预测模型,现有的体系有HrisPlex-s、IrisPlex等^[12-13]。

在欧洲人群中,色素特征较为多变,对头发和虹膜颜色的推断有助于缩小侦查范围。为此,研究人员不断发掘与积累色素特征性状的数据,以提高对色素预测的准确性及全面性。2018年Hysi等^[14]对近300 000名欧洲血统参与者的全基因组关联研究进行Meta分析,鉴定了138个SNP位点,其中有7个SNP显示出显著关联,包括1个未经报道的新位点1q32.1 DSTYK rs2369633。2019年Peng等^[15]在VISAGE项目的支持下,首次发表了关于眉毛颜色的GWAS分析,发现眉毛颜色位点不完全与头发颜色相关,如金发碧眼但眉毛却呈现棕色的情况在欧洲并不少见。基于8 500个欧洲样本的测试确定了6个显著相关的SNP位点,其中包括1个新位点。2021年VisiGen联盟发表了一项基于19.5万名欧洲人眼睛颜色的GWAS研究,并确定了124个独立相关基因,其中包括50个未经报道的眼睛颜色相关基因,其研究解释了人群中眼睛颜色总变化的53% (95% CI: 0.45 ~ 0.61)^[16]。同年Kukla-Bartoszek等^[17]对150名波兰受试者进行了全外显子组测序研究,报告了27个新的候选眼睛颜色SNP位点。在新候选的SNP位点与已知色素相关的114个SNP位点

中,选定了137个SNP位点,于849个波兰样本中预测建模,Kukla-Bartoszek等^[17]采用赤池信息量准则(Akaike information criterion, AIC)、贝叶斯信息准则(Bayesian information criterion, BIC)和LASSO方法进行标记选择并建立回归模型,发现位于ARFIP2基因中的rs2253104位点显示出预测潜力。2022年Salvo等^[18]使用ForenSeq™ DNA Signature预处理试剂盒(Verogen)评估了24个HIrisPlex SNPs的分型性能,并在挪威研究人群中使用HIrisPlex网络工具进行了眼睛和头发颜色预测,结果发现红色(0.97)和黑色(0.93)头发颜色以及蓝色(0.85)和棕色(0.94)眼睛颜色的AUC预测精度很高。

尽管我国大部分人群具有典型的黄色人种表型特征,但作为一个多民族国家,验证适用于中国人群的色素特征推断体系,对案件侦破的效力同样不容小觑。2020年,张微等^[19]探讨了HIrisPlex眼睛及头发颜色推断体系在中国亚欧混合人群中的应用,该研究得出HIrisPlex对欧亚混合人群的蓝色和棕色眼睛推断模型评估中AUC值可达0.876和0.874。2021年,杨亚芳等^[20]进行了41-Plex SNP色素推断复合检测体系在中国人群中的适用性研究,该研究在尝试多种肤色分类方法后,得到相对较高的AUC值:白肤色0.831,中间肤色0.661,深肤色0.641,较黑-黑肤色0.768。

近年来,随着研究者对色素特征性状数据的不断深入发掘并积累,色素预测模型有望变得更加精准和全面,并在实践中加以应用。这不仅有助于法庭科学领域的案件侦破,还可以为人类遗传学研究提供更深入的洞察。同时,不同人群间的差异性研究将更加细致,从而提高预测模型在不同地区和民族中的适用性和准确性。

3.2 身高特征

身高是一种高度遗传的多基因性状,据悉,成年人身高遗传力大约80%,其中SNP是最常见的人类遗传变异之一,也是人类身高相关研究的一大热点^[21]。2014年人体性状遗传研究(Genetic Investigation of Anthropometric Traits, GIANT)协会,针对超过25万欧洲个体的全基因组数据进行全基因组关联研究,报告了697个SNP变异位点,并解释了欧洲群体36%的SNP遗传力^[22]。2016年Marouli等^[23]对超过70万人进行了GWAS,报告了83个与身高相关的低频(基因频率0.1%~4.8%)SNP位点,其影响高达2 cm(如IHH、STC2、AR和CRISPLD2的影响)。2021年Isshiki等^[24]在GWAS的基

基础上,对10 840名日本个体进行多基因评分,发现身高存在遗传背景差异且存在南北地理高度梯度。2022年,YENGO团队对540万人进行综合Meta分析,确定了12 111个显著性SNP位点,分别位于7 209个基因座,占基因组的21%。这些基因座(包含超过100万个SNP位点)解释了欧洲群体几乎全部的SNP遗传力和非欧洲血统群体中至少90%的SNP遗传力^[25]。2023年陕西师范大学计算机科学学院与公安部物证中心基于已报道的573个日本人群身高相关SNP位点,在1578名中国北方汉族男性样本中进行关联验证,得到47个显著位点,对身高定量预测中,47个SNP位点的随机森林模型准确性最佳 $[R^2=0.14$ (95% CI: 0.04 ~ 0.24), $MAE=4.87$ (95% CI: 4.68 ~ 5.06)]。在身高定性预测中,基于双向逐步回归筛选出的139个SNP位点,K近邻分类器对于高身高(≥ 178 cm)的预测效果最好(AUC=0.68, 95% CI: 0.66 ~ 0.71);逻辑回归分类器对于矮矮身高(< 165 cm)的预测效果最好(AUC=0.70, 95% CI: 0.68 ~ 0.72)^[26]。

身高预测模型在法庭科学中的应用前景。身高预测不仅为案件侦破提供线索,在失踪人员的身份识别及地理祖先溯源中,也可以作为重要的参考信息。尽管研究者们已发掘大量的身高相关SNP位点,但现有的预测模型预测准确性普遍偏低,仍需要补充大量遗传标记点,以达到较高准确性。

3.3 面部特征

长期以来,畸形面部特征一直是临床和遗传学研究的焦点,自人类基因组和表型组计划的兴起,正常面部特征也逐步得到科研工作者的重视。人类面部结构复杂且多变,受遗传、环境因素及其相互作用的影响。研究表明,大多数面部特征的遗传率超过60%,遗传因素可以解释70%以上的面部表型变异,包括面部形状、鼻子(高度、宽度)、耳朵形状等^[27-28]。面部特征研究主要包括三维脸部图像的数字化以及脸部特征点间距、角度、主成分变量等特征关联遗传位点的挖掘^[29]。自GWAS研究问世以来,涉及面部形态特征性状的研究不断进步,提高并拓宽了科研工作者们对人类外貌遗传基础的认识。

2015年Adhikari等^[30]对来自巴西、智利、哥伦比亚、墨西哥和秘鲁的5 062例个体进行了一项全基因组关联研究(GWAS),以识别影响人类耳朵形态变化的多个基因座。结果显示,7个基因区域与耳朵形态特征(如耳垂大小和附着、反耳轮折叠、耳

郭突出和对耳轮大小等)显著相关。其中,4个特征区域与外胚乳蛋白A受体基因的功能变异有关,该基因在胚胎皮肤附属器官发育中起关键作用。2018年,Claes等^[31]在近2 300名欧洲受试者中进行了一项关于脸型的GWAS研究,发现了38个明显相关的遗传位点,其中15个位点在1 700例独立欧洲样本中得到了重现,包含了4个新SNP位点。在该研究小组2021年的一项后续研究中,将其表型分析方法应用于近8 200名欧洲受试者的扩大样本集,并报告了203个明显相关的遗传位点,其中包含53个未知位点,位于与面部发育或面部表现疾病有关的区域^[32]。2019年Sero等^[33]基于已知的118 420个SNP位点,剔除年龄小于18岁后,对7 150个个体进行分析比较,发掘了2 232个面部特征识别相关的显著SNP位点。同年,Xiong等^[34]基于已知的203个面部特征显著相关遗传位点,对近10 100名欧洲受试者的78种面部形状表型进行了GWAS分析,这些表型是通过使用专用计算机视觉方法在三维数字面部图像上自动放置13个面部地标获得的,并在另外超过7 900名欧洲人和非欧洲人中进行了重现,发现了24个明显相关的遗传位点,其中17个新位点。2021年,Bonfante等^[35]在6 000余名拉丁美洲人中进行了一项面对面的GWAS研究,报告了32个遗传位点的显著相关性,其中9个位点是之前未确定的。2023年,Li等^[36]采用GWAS方法对来自欧洲、亚洲和拉丁美洲14 921例个体进行了分析,识别了16个与耳形态相关的基因位点,其中8个位点是首次被发现的。

2018年李祎等^[37]在包含715例新疆维吾尔族亚欧混合人群的针对EDARV370A及一系列系统的面部特征及耳形态表型的关联分析中发现,EDARV370A与8个面部形态的定量表型、下巴类型以及3个耳形态定序分类表型显著相关(多重检验校正后 $P < 0.05$),进一步明确了EDARV370A的遗传多效性及其在亚欧混合人群中对面部和耳形态的影响。2019年,Wu等^[38]针对我国北方汉族群体个体,使用全外显子组测序和Sanger测序进行研究,采集50例头部3D CT扫描数据,对48种面部和颅骨表型进行生物信息学分析。研究共识别出4个基因区域(RGPD3, IGSF3, SLC28A3, USP40)与颅面形态显著相关,其中RGPD3基因的rs62152530位点与鼻长、耳长和鼻宽密切相关;IGSF3基因的rs647711位点与鼻长、下颌宽度和颞孔间距相关;SLC28A3基因的rs10868138位点与鼻长、耳间距和

鼻翼宽度相关；USP40基因的rs1048603和rs838543位点与梨状孔宽度相关。2020年，刘明等^[39]在1 177名中国汉族成年男性人群的17个脸部特征点中提取出136组欧几里得距离表型，基于已报道的与脸部形态显著相关的125个SNP位点分析相关性，解释了3.89%脸部表型差异。2022年，Zhang等^[40]基于6 968个汉族个体及其2 706例复制队列的3D面部图像进行了一项人脸GWAS研究，发现了166个显著相关的基因位点，其中62个位点之前未发现与人脸变异相关。

近年来相关研究发现了许多与面部形态特征相关的SNP，但这些SNP的效应大小较小，对表型变异的解释能力有限。目前，这些SNP还不足以用于法庭科学的面部形态特征分析。因此，需要更大规模的样本量和更多的GWAS研究，以确定更多具有实际预测价值的小效应SNP，并将其纳入未来的法医面部形态特征分析工具中。

3.4 男性脱发

雄激素性脱发，俗称男性型秃发(male pattern baldness, MPB)，是男性最常见的进行性脱发障碍类型。利用DNA变异预测早发或正常MPB状态的能力可能对治疗策略具有重要意义。此外，MPB作为一项法医DNA表型预测的外部特征，在实际案件中，可结合视频证据、证人证词等进行综合分析，判断是否佩戴假发等信息，为侦查方向提供参考信息。2016年，Liu等^[41]已报道的MPB相关位点，对2725名德国和荷兰男性进行了预测分析。来自12个基因组位点的25个SNP的基因型的Logistic回归模型表明，当模型中包含14个SNP时，早发性MPB风险的准确度为0.74。当使用6~11个SNP位点时，该模型能够以略低的准确度(AUC=0.69~0.71)预测中老年人的正常MPB状态。其中55.8%的早发性MPB遗传易感性可由常染色体SNP解释，23.3%可由X染色体SNP解释。2020年潘思宇等^[42]参考已报道的624个与MPB相关的SNPs，在我国684个亚欧混合人群样本中基于多基因风险评估建立线性回归和逻辑回归模型，效果较好($R^2=28.9%$, AUC=0.82)；年龄对模型效果影响较大($R^2=22.9%$, AUC=0.77)，结合体质量指数及68个SNPs时AUC达到最大(约0.89)。Chen等^[43]参考已报道的相关SNP位点，在超过5.5万名英国男孩中发现了117个SNP预测因子。基于来自85个基因位点的这117个SNP位点，他们用不同的方法在英国生物样本库研究(UKBB)超过10万名男性中建立了MPB预

测模型，并在超过2.6万名UKBB中的男性中进行了验证。不同方法的AUC值范围相似：重度脱发为0.725~0.728，中度脱发为0.631~0.635，轻度脱发为0.598~0.602，无脱发为0.708~0.711；而“有脱发与无脱发”的二元分类预测在纳入年龄时AUC为0.690~0.711，不纳入年龄时略低。在早期发病的富集MPB数据集(n=991)中进行的额外外部验证显示，在不考虑年龄的情况下，预测准确性有所提高，如无脱发与任何脱发的AUC为0.830。Chen等^[43]发现的可预测MPB的SNP位点现已被纳入经法医验证的VISAGE-ET-AA MPS工具，男性脱发预测模型已在VISAGE软件中实现。

现阶段针对MPB的研究大多在治疗与诊断方面，程度预测的研究并不多，虽然已有对早发性脱发预测效果较好的预测模型，但其适用人群多为欧美人群，已在法庭科学中实际应用。适用于我国人群的MPB相关SNP位点以及适用于法庭科学应用的预测模型有待进一步开发和验证。

4 当前的挑战与未来发展方向

尽管国内外研究者在表型推断方面不断深入研究，基于SNP的表型预测仍面临诸多挑战。由于不同种族和地区的遗传背景存在较大差异，这些模型在不同种族中的预测准确性可能不高。因此，需要更多不同种族和地区的人群样本，以提高模型的普适性。此外，表型特征如身高、发色、肤色等通常由多基因控制，这些特征不仅受遗传因素影响，还受到环境因素的调节。因此，基于SNP的预测模型需要整合更多的信息，才能更准确地进行预测。尽管GWAS研究发现了许多与表型相关的SNP，但有些SNP的效应往往较小，对表型变异的解释能力有限。因此，需要更大规模的研究来发现更多具有显著效应的SNP。虽然高通量测序技术和机器学习方法在不断进步，但现有的方法在处理高维基因数据和复杂表型关系时仍存在局限性。如SNP间的相互作用和表型的非线性关系难以被简单的模型捕捉到。

针对现阶段SNP推断表型的难点做出以下展望。(1)进一步扩大样本量及其涵盖的种族和地区，旨在提高SNP与表型关联的发现率，增强预测模型的适用性和准确性。(2)人类表型特征的遗传机制较为复杂，在进行SNP推断表型的研究中，各种类型的非遗传因素也需考虑在内，如生长环境、生活水平、职业等因素，综合分析这些因素有助于构建更

完备的预测模型。(3)表型预测的研究需要遗传学、统计学、计算机科学等多个学科的紧密合作。新一代测序技术(如单细胞全基因组测序等)和先进的机器学习算法,这些技术可以提高测序精度和数据处理能力,从而更好地捕捉SNP间的复杂关系,提高表型预测的准确性和实用性。(4)在实际应用中,尤其是法庭科学中,预测模型需要经过严格的验证和优化,以确保其可靠性和准确性。未来的研究应注重模型的应用测试,并根据实际情况不断改进模型。通过应对这些挑战以及技术的不断进步,基于SNP的表型推断在法庭科学中的应用前景将更加广阔,有望为案件侦破和犯罪嫌疑人追踪提供强有力的科学支持。

作者贡献 周玲: 论文撰写; 康霞: 审读和修订, 监督指导; 王艳: 审读和修订。

利益冲突 所有作者声明无利益冲突。

参考文献

- Hicks T, Buckleton J, Castella V, et al. A logical framework for forensic DNA interpretation [J]. *Genes (Basel)*, 2022, 13 (6): 957.
- Consortium IH. A haplotype map of the human genome [J]. *Nature*, 2005, 437 (7063): 1299-1320.
- 侯一平. 法医学证学 [M]. 4版. 北京: 人民卫生出版社, 2016: 102-103.
- Wen YF, Liu J, Su YL, et al. Forensic biogeographical ancestry inference: recent insights and current trends [J]. *Genes Genomics*, 2023, 45 (10): 1229-1238.
- 朱波峰, 孟昊天, 兰琼. 我国人体表征分子鉴识研究的成果、挑战与展望 [J]. *法医学杂志*, 2019, 35 (5): 507-511.
- Dabas P, Jain S, Khajuria H, et al. Forensic DNA phenotyping: Inferring phenotypic traits from crime scene DNA [J]. *J Forensic Leg Med*, 2022, 88: 102351.
- Hu TS, Chitnis N, Monos D, et al. Next-generation sequencing technologies: an overview [J]. *Hum Immunol*, 2021, 82 (11): 801-811.
- Boemo MA. DNAscent v2: detecting replication Forks in nanopore sequencing data with deep learning [J]. *BMC Genomics*, 2021, 22 (1): 430.
- Tam V, Patel N, Turcotte M, et al. Benefits and limitations of genome-wide association studies [J]. *Nat Rev Genet*, 2019, 20 (8): 467-484.
- Khera AV, Chaffin M, Aragam KG, et al. Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations [J]. *Nat Genet*, 2018, 50 (9): 1219-1224.
- Barash M, McNevin D, Fedorenko V, et al. Machine learning applications in forensic DNA profiling: a critical review [J]. *Forensic Sci Int Genet*, 2024, 69: 102994.
- Paparazzo E, Gozalishvili A, Lagani V, et al. A new approach to broaden the range of eye colour identifiable by IrisPlex in DNA phenotyping [J]. *Sci Rep*, 2022, 12 (1): 12803.
- Brancato D, Coniglio E, Bruno F, et al. Forensic DNA phenotyping: genes and genetic variants for eye color prediction [J]. *Genes (Basel)*, 2023, 14 (8): 1604.
- Hysi PG, Valdes AM, Liu F, et al. Genome-wide association meta-analysis of individuals of European ancestry identifies new loci explaining a substantial fraction of hair color variation and heritability [J]. *Nat Genet*, 2018, 50 (5): 652-656.
- Peng FD, Zhu G, Hysi PG, et al. Genome-wide association studies identify multiple genetic loci influencing eyebrow color variation in Europeans [J]. *J Invest Dermatol*, 2019, 139 (7): 1601-1605.
- Simcoe M, Valdes A, Liu F, et al. Genome-wide association study in almost 195, 000 individuals identifies 50 previously unidentified genetic loci for eye color [J]. *Sci Adv*, 2021, 7 (11): eabd1239.
- Kukla-Bartoszek M, Teisseyre P, Pośpiech E, et al. Searching for improvements in predicting human eye colour from DNA [J]. *Int J Legal Med*, 2021, 135 (6): 2175-2187.
- Salvo NM, Janssen K, Kirsebom MK, et al. Predicting eye and hair colour in a Norwegian population using Verogen's ForenSeq™ DNA signature prep kit [J]. *Forensic Sci Int Genet*, 2022, 56: 102620.
- 张微, 马咪, 许磊, 等. HIRISPLEX眼睛及头发颜色推断体系在中国亚欧混合人群中的应用 [J]. *中国法医学杂志*, 2020, 35 (3): 278-285.
- 杨亚芳, 刘雅洁, 梁鳅双, 等. 41-Plex SNP色素推断复合检测体系在中国人群中的适用性研究 [J]. *中国法医学杂志*, 2021, 36 (2): 148-153.
- Zhong KY, Zhu G, Jing XX, et al. Genome-wide compound heterozygote analysis highlights alleles associated with adult height in Europeans [J]. *Hum Genet*, 2017, 136 (11/12): 1407-1417.
- Wood AR, Esko T, Yang J, et al. Defining the role of common variation in the genomic and biological architecture of adult human height [J]. *Nat Genet*, 2014, 46 (11): 1173-1186.
- Marouli E, Graff M, Medina-Gomez C, et al. Rare and low-frequency coding variants alter human adult height [J]. *Nature*, 2017, 542 (7640): 186-190.
- Isshiki M, Watanabe Y, Ohashi J. Geographic variation in the polygenic score of height in Japan [J]. *Hum Genet*, 2021, 140 (7): 1097-1108.
- Yengo L, Vedantam S, Marouli E, et al. A saturated map of common genetic variants associated with human height [J]. *Nature*, 2022, 610 (7933): 704-712.
- 罗祎倩, 孙亚男, 李彩霞, 等. 基于机器学习的中国北方汉族男性身高遗传预测研究 [J]. *中国法医学杂志*, 2023, 38 (3): 290-297.
- Kienkas K, Jakobsone G, Salms G. The facial characteristics of individuals with posterior crossbite: a cross-sectional study [J]. *Healthcare (Basel)*, 2023, 11 (13): 1881.
- Crouch DJM, Winney B, Koppen WP, et al. Genetics of the human face: identification of large-effect single gene variants [J]. *Proc Natl Acad Sci USA*, 2018, 115 (4): E676-E685.
- 赵雯婷, 江丽, 李彩霞. 精准医疗时代的法医表型特征分子刻画研究进展 [J]. *法医学杂志*, 2019, 35 (5): 525-530.
- Adhikari K, Reales G, Smith AJP, et al. A genome-wide association study identifies multiple loci for variation in human ear morphology [J]. *Nat Commun*, 2015, 6: 7500.
- Claes P, Roosenboom J, White JD, et al. Genome-wide mapping of global-to-local genetic effects on human facial shape [J]. *Nat Genet*, 2018, 50 (3): 414-423.