FORENSIC SCIENCE ADVANCES

Open Access

e-ISSN 3136-1021

Review Article

Advances in forensic DNA phenotyping

Caixia Li[1,§], Hao Zhang[1,§], Chao Liu[2] ✉, Lan Hu[1] ✉

[1] *Key Laboratory of Forensic Genetics, Beijing Engineering Research Center of Crime Scene Evidence Examination, National Engineering Laboratory for Forensic Science, Institute of Forensic Science, Beijing 100038, China*

[2] *State Key Laboratory of Bioactive Molecules and Druggability Assessment, Guangdong Basic Research Center of Excellence for Natural Bioactive Molecules and Discovery of Innovative Drugs, Jinan University, Guangzhou 510632, China*

[§] *Caixia Li and Hao Zhang contributed equally to this work.*

Cite this article: Li, C. X., et al. *Forensic Science Advances*, **2026**, 1(1): 33–40. https://doi.org/10.26599/FSA.2026.9750002.

ABSTRACT: Forensic DNA phenotyping (FDP) refers to a suite of technologies that infer phenotypic information from the genetic material of biological samples. These technologies play a critical role in criminal investigations. This review systematically summarizes recent advances in four key areas: biogeographical ancestry inference, forensic genetic genealogy, appearance prediction, and age prediction. We discuss the underlying genetic principles, current methodological approaches, major achievements, persistent challenges, and outline directions for future research to bridge the gap between forensic DNA phenotyping technologies and operational casework.

KEYWORDS: forensic genetics; forensic DNA phenotyping; biogeographical ancestry; forensic genetic genealogy

1 Introduction

Forensic DNA phenotyping (FDP) refers to a collection of technologies that infer phenotypic information—including biogeographical ancestry, genealogy, appearance, age, and lifestyle—from the genetic information contained in biological samples, thereby providing investigative leads [1,2]. This technological framework (Fig. 1) relies on the construction of large-scale population cohorts, through which both genetic and phenotypic data are collected. Genome-wide association studies (GWAS) and epigenome-wide association studies (EWAS) identify molecular markers associated with phenotypes, which in turn serve as the basis for constructing predictive tools for forensic DNA phenotyping [1,3].

Traditionally, forensic genetics has relied on short tandem repeat (STR) analysis of biological samples collected from crime scenes to identify perpetrators, but this approach requires a match with a suspect or a DNA database entry. Advances in omics technologies have driven significant progress in forensic genetics, with multi-omics tools expanding the scope of information obtainable from evidentiary samples [4]. Even in the absence of a match, FDP can predict various phenotypes from biological samples [5], thereby narrowing the pool of potential suspects, guiding investigative directions, and providing new leads particularly in cold cases and long-standing unsolved cases.

This review covers four core domains of FDP: biogeographical ancestry (BGA) inference, forensic genetic genealogy (FGG), appearance inference, and age prediction. These domains share fundamental methodological principles [1,3]. However, they differ substantially in the genetic architecture and definitional measurement of the predicted traits, leading to marked variations in predictive accuracy. BGA inference yields qualitative ancestry labels [5,6]. In contrast, forensic genetic genealogy provides quantitative estimates of kinship degrees and is a relatively mature domain [7,8]. Within appearance inference, pigmentation traits governed by oligogenic architectures can be predicted with reasonable accuracy [9–11], whereas facial morphology—a complex polygenic trait—remains at an experimental stage [12]. Unlike other domains, which depend on genomic variants such as single nucleotide polymorphisms (SNPs), age prediction leverages epigenetic DNA methylation markers [13]. The subsequent sections will elaborate on the basic concepts and recent advances in each of these four areas, and discuss current challenges and future perspectives for bridging the gap between FDP technologies and operational casework.

2 Biogeographical ancestry inference

The core objective of biogeographical ancestry inference is to utilize ancestry-informative signals carried in an individual's genome to estimate their continental or sub-continental biogeographical ancestry [5,6]. The genetic basis of BGA inference lies in allele frequency differences among different populations, which result from long-term processes of human migration history, genetic drift, and natural selection. Ancestry informative markers (AIMs) are genetic loci exhibiting significant allele frequency differences among ancestral populations and constitute the core toolkit for BGA inference [14]. AIMs are selected using various statistical metrics, including the fixation index (F_{st}). Selected AIM panels can then

Received: 29 July 2026; **Revised:** 21 August 2026

Accepted: 24 August 2026

✉Address correspondence to Lan Hu, hulan328@126.com;

Chao Liu, liuchaogzf@163.com

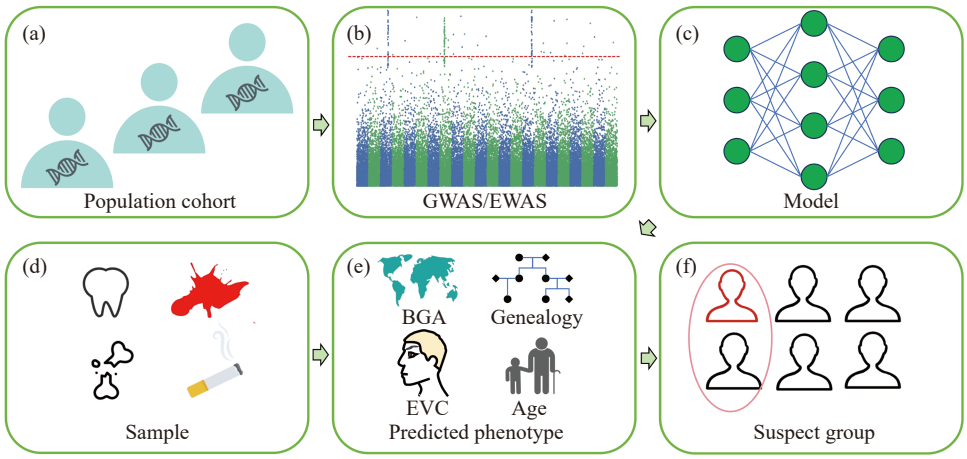


Figure 1 Schematic overview of the forensic DNA phenotyping (FDP) workflow. (a) Large-scale population cohorts are established to collect both genetic and phenotypic data. (b) Genome-wide association studies (GWAS) and epigenome-wide association studies (EWAS) are performed to identify molecular markers significantly associated with phenotypic traits. (c) Predictive models are constructed using statistics and machine learning approaches based on the identified markers. (d) Biological samples recovered from crime scenes (e.g., bloodstains, bones, teeth, and touch DNA samples) are genotyped. (e) The models are applied to infer phenotypic information of the sample donor, including biogeographical ancestry (BGA), genealogy, externally visible characteristics (EVC), and age. (f) The predicted phenotypes provide investigative leads that help narrow down the suspect group.

be applied in practice. For example, a single-tube 27-plex SNP assay for estimating individual ancestry and admixture across three continents has been previously described [15]. Furthermore, the authors developed a panel of 74 AISNPs that achieved differentiation of populations from ten global regions and could further distinguish northern and southern genetic components within East Asian populations [16].

A study evaluating the performance of three popular forensic BGA panels—MAPlex, Precision ID Ancestry Panel (PIDAP), and VISAGE Basic Tool (VISAGE BT)—found that at the continental-level classification (K = 6), all panels achieved an area under the precision-recall curve (AUPR) greater than 0.97; however, at sub-continental classification levels (K = 7 and K = 8), the panels exhibited bias [17]. Such bias may be attributable to reference bias: first, current reference populations over-represent individuals of European ancestry while under-representing non-European populations; second, the geographical labeling of genetic populations is an assumption, as the genetic ancestry of a population in a given region may have changed significantly over time due to migration and drift [18].

To improve the understanding of genetic diversity of Asian populations, several projects have been launched. The Central Asian Genomic Diversity Project (CAGDP) systematically investigated and reconstructed the evolutionary history of Central Asian populations [19]. The 10K Chinese People Genomic Diversity Project (10K_CPGDP) revealed the genetic structure and adaptation history of ethnic groups along the Silk Road region of China [20]. Such studies provide upstream theoretical support for BGA inference.

Current methods for estimating biogeographical ancestry from SNP data typically rely on principal component analysis (PCA) and model-based clustering implemented in STRUCTURE [21] or ADMIXTURE [22]. PCA reduces genome-wide SNP genotypes to orthogonal axes of maximal variance, onto which query samples are projected for cluster-based assignment, whereas model-based clustering implemented in STRUCTURE and ADMIXTURE applies Bayesian inference to estimate each individual's ancestry proportions from K hypothetical ancestral populations, thereby explicitly accommodating admixed individuals. In contrast to the aforementioned unsupervised methods, several studies have proposed supervised learning approaches based on labeled datasets, such as partial least squares discriminant analysis (PLS-DA) or extreme gradient boosting (XGBoost) for ancestry inference based on SNP data, with PLS-DA considered more robust for this purpose [23]. STRs [24], insertion-deletion polymorphisms [25,26], and microhaplotypes [27] have also shown advances in BGA inference. The microfluidic system, based on a 38-DIPs (insertion/deletion polymorphism) panel, achieved fully automated sample-in-answer-out rapid detection within 124 minutes [26].

3 Forensic genetic genealogy

Forensic genetic genealogy, also referred to as investigative genetic genealogy (IGG), generally involves using genome-wide SNP data, in conjunction with large-scale population genomic databases and population genetics algorithms, to infer distant genetic relationships of an individual of interest [28–30]. Degrees of relatedness are defined by the expected amount of shared DNA [31]. The identical-by-descent (IBD) algorithm is a commonly used approach in genealogical analysis, which calculates the length and number of DNA segments inherited from a common ancestor without recombination, measured in centimorgans (cM), to infer kinship relationships. By excluding chance matches and long-range linkage disequilibrium from segments with identical-by-state (IBS) nucleotide consistency, IBD segments are identified [29]. IBD-based algorithms can infer relationships spanning one to seven degrees of relatedness, or even more distant kinship [8].

FGG has demonstrated substantial utility in criminal investigations. In 2018, American investigators uploaded SNP genotype data from crime scene DNA to the GEDmatch public genealogy database, identified distant relatives of the suspect, and ultimately apprehended the perpetrator, successfully

resolving the decades-long "Golden State Killer" case—an achievement recognized by Science as one of the top ten scientific breakthroughs of the year [32]. In China, a 14-year-long cold case was resolved using FGG techniques, whereby investigators narrowed down from a large pedigree to a smaller lineage and ultimately apprehended the suspect. This application of the technology avoided large-scale population screening and expedited case resolution [7].

In practice, FGG depends on SNP detection. SNP microarrays offer comparatively low cost and ease of variant calling, making them the predominant method in FGG [33]. Previously, FGG techniques and applications based on SNP microarrays have been established. However, common crime scene samples such as bone or teeth often yield DNA of poor quality, whereas SNP microarrays require relatively high DNA quality and are not well suited to low-template or degraded DNA samples [34]. Whole-genome sequencing (WGS) technology offers advantages in analyzing such samples [35–37]. In the application of WGS, an evaluation of sequencing depth demonstrated that 8× coverage did not differ significantly from higher-depth data in terms of prediction accuracy [38]. An assessment of samples with varying DNA quality found that 0.5 ng of DNA input performed comparably to 200 ng, and 200 bp DNA fragments showed no difference from non-degraded DNA[35]; these findings provide practical guidance for WGS-based applications.

Nevertheless, WGS is susceptible to interference from microorganisms, inhibitors, and other contaminants, and its detection capability remains limited for trace degraded biological samples such as touch DNA. Moreover, the costs associated with testing, analysis, and storage remain prohibitively high, precluding large-scale adoption. Targeted sequencing represents another technique applicable to kinship inference and has now emerged as a leading research frontier in this field. The Kintelligence Kit [39] containing approximately 10,000 SNPs is used for kinship analysis within the 5th degree. To establish a panel for analyzing distant relatives within East Asian populations, researchers screened 9,000 SNPs and evaluated their efficacy in kinship analysis [40]. Building on this foundation, they further evaluated the impact of different SNP densities on the effectiveness of kinship inference [41] and developed a kit containing 20,838 SNPs.

4 Appearance inference

Reconstructing human appearance traits from genetic information is highly valuable for forensic identification [42], yet research progress in this area remains limited [43]. Externally visible traits are influenced by both environmental and genetic factors, and most traits are polygenic, which adds to the complexity of appearance prediction from genetic data.

4.1 Pigmentation traits

Among externally visible traits, pigmentation traits are relatively well studied due to the presence of prominent major-effect genes. HIrisPlex-S [11] is currently the most widely used combined pigmentation prediction tool. Researchers first developed IrisPlex [9], which predicts eye color based on six pigmentation-associated SNPs. This was followed by HIrisPlex [10], which extended prediction to hair color. The model was subsequently expanded to include skin color prediction, resulting in HIrisPlex-S [11], which enables reasonably accurate simultaneous prediction of eye, hair, and skin color. The HIrisPlex-S system employs a multinomial logistic regression (MLR) model to calculate the probability of each pigmentation category from SNPs, and reports the category with the highest probability.

4.2 Facial morphology

Compared with pigmentation traits, reconstructing facial topology from genetic information poses a far greater challenge. Facial morphology is influenced by both skeletal and soft-tissue morphology, and various two-dimensional and three-dimensional methods have been developed to characterize facial features [44,45]. During embryonic development, cranial neural crest cells contribute to the formation of most craniofacial structures [46]. A meta-analysis of GWAS based on European populations found that regions surrounding significant signals associated with facial morphology showed enrichment in enhancer activity in cranial neural crest cells and craniofacial tissues [12]. GWAS and EWAS conducted in East Asian populations have also deepened the understanding of the genetic basis of facial morphology across different populations [47–49]. Nevertheless, overall understanding of facial genetics remains incomplete, and reconstructing facial images from genetic data continues to present difficulties.

Given that facial morphology is a complex trait composed of multiple sub-traits, C-GWAS represents a promising approach, as it combines GWAS summary statistics of multiple potentially correlated traits to improve statistical power [50]. More recently, researchers proposed Diffface, a method inspired by text-image multimodal models that attempts to reconstruct facial images from genetic information. It first maps SNPs and 3D facial images to the same low-dimensional feature space using a transformer and a spiral convolutional network, respectively, and establishes associations between the two modalities in latent features through contrastive learning. Subsequently, a diffusion model is introduced to reconstruct facial structures from the characteristics of SNPs [51].

4.3 Height

Height is also of considerable importance in forensic DNA phenotyping. A landmark study [52] published in Nature in 2022 utilized cross-ancestry GWAS data from 5.4 million individuals (encompassing European, East Asian, Hispanic, African, and South Asian populations) and identified 12,111 independent height-associated SNPs, explaining nearly all common SNP-based heritability. However, this map explained 40% of phenotypic variance in European populations but only 10%–20% in other ancestral groups [52]. In studies of height estimation in Han Chinese populations, a model employing partially connected neural networks (PCNN) achieved the highest accuracy [53]. An EWAS conducted in a Chinese population modeled DNA methylation sites in relation to height phenotype [54], suggesting the potential of extending beyond SNPs for height estimation.

5 Age prediction

Age prediction is an important component of FDP. DNA methylation, an epigenetic modification, is a commonly used biomarker for age estimation [13]. In the human genome, cytosine followed by guanine at a "CG" dinucleotide position is termed a CpG site. The addition of a methyl group to the fifth carbon of cytosine, forming 5-methylcytosine (5-mC), constitutes the process of DNA methylation. CpG islands, which are regions enriched in CpG sites, are typically located in promoter regions and are generally unmethylated. Promoter methylation is associated with gene silencing [55]. Unmethylated cytosines can be converted to uracil by bisulfite treatment and subsequently converted to thymine after PCR amplification, whereas methylated cytosines remain unchanged. Bisulfite conversion is thus employed for the detection of DNA methylation [56].

DNA methylation patterns change with age throughout the human lifespan and can therefore be used to estimate an individual's age [57]. Several age prediction models and methods based on DNA methylation have been developed, and researchers have tested a variety of detection approaches [58]. Recently, researchers developed a novel 17-plex assay on an advanced five-color fluorescence droplet digital PCR (ddPCR) platform. Coupled with a support vector regression (SVR) model, it explained over 98% of age variation, achieving mean absolute deviations (MADs) of 2.396 years (saliva), 2.006 years (blood), and 2.044 years (buccal swabs), while maintaining robust performance across a wide range of age groups [59]. Future efforts should expand multi-tissue and multi-omics analytical methods to realize practical age prediction in forensic applications [60].

6 Discussion

Forensic DNA phenotyping, as a frontier area of forensic genetics, has made substantial progress in recent years in terms of theoretical foundations, technological methods, and practical applications. Leveraging the inherent correlations among phenotypes enables joint inference across multiple traits. The VISAGE Basic Tool (BT) [61] and the VISAGE Enhanced Tool (ET) [62] are tools capable of simultaneously predicting eye, hair, and skin color, along with BGA. Efforts have also been directed toward integrating age estimation with the prediction of white hair, a visible age-related trait [63]. Furthermore, in FGG applications, BGA inference can help narrow the pedigree search space, thereby improving investigative efficiency [7]. These developments collectively reflect the growing cross-fertilization and integration across different FDP domains. Beyond the phenotypes discussed above, lifestyle factors such as caffeinated beverage consumption [64] and smoking [65] can be predicted based on DNA methylation levels. Voice pitch and vowel acoustics have also been found to have a heritable component [66]. The association between fingerprint patterns and the genome has also been investigated [67]. These advances broaden the informational dimensions of DNA evidence and provide richer leads for criminal investigations.

Nevertheless, this field still faces multiple challenges. Currently, population coverage in available databases is uneven. Existing studies and databases are disproportionately composed of European individuals, while genetic and phenotypic data for non-European populations remain relatively scarce, leading to reduced applicability and accuracy of prediction models in other populations. Furthermore, the heritability explained for complex traits remains limited, and the impact of environmental factors and gene-environment interactions is difficult to quantify accurately.

More importantly, FDP carries potential societal risks, including exacerbating racial inequality, infringing privacy, and operating within lagging regulatory frameworks. First, the biogeographical ancestry labels produced by BGA inference may be misinterpreted as socially constructed racial categories, thereby stigmatizing specific populations [68–70]. In some jurisdictions, certain communities are already disproportionately represented in forensic DNA databases [68]. Applying FDP without safeguards could create a feedback loop that further entrenches judicial discrimination against these groups. Second, privacy concerns have drawn significant attention [69,70]. In the Golden State Killer case, investigators uploaded a suspect's DNA profile to a public genealogy database under a fake identity, bypassing the informed consent of database contributors [71]. This practice sparked widespread ethical debate. Moreover, FDP should be strictly limited to externally visible characteristics (EVC) [69] and should never be used to predict disease susceptibility or other private medical information. Third, there is currently no unified international legal framework governing FDP, resulting in fragmented regulation across jurisdictions [69,70]. For example, the Netherlands became the first country to legislate FDP in 2003, but it requires that all other investigative methods be exhausted before deployment. Germany explicitly prohibits BGA inference and permits only pigmentation and age estimation. Because FDP yields probabilistic results that cannot serve as sole evidence for conviction but only as investigative leads, it often circumvents the Daubert standard and the Frye standard, leaving it in a legal vacuum in many regions. Taken together, the ethical, legal, and social implications (ELSI) of FDP demand rigorous and ongoing scrutiny [1,3,33,63,68–72]. In 2023, the National Technology Validation and Implementation Collaborative (NTVIC) provided guiding principles and reference recommendations for FGG [73].

Looking forward, forensic DNA phenotyping technologies will advance toward higher accuracy, broader applicability, and greater standardization. With the continued progress of global population genome projects and in-depth phenomics research, the precision and population applicability of prediction models will continue to improve. In addition, the deep integration of artificial intelligence and big data analytics will drive a transition from single-trait prediction to multi-trait joint inference, potentially improving prediction accuracy when biogeographical ancestry, appearance features, and age are considered jointly. Moreover, it is essential to narrow the gap between forensic DNA phenotyping technologies and operational practice [72], thereby facilitating the translation of these technologies to enhance real-world case resolution capabilities.

To address population biases, it is imperative to expand genetically diverse cohorts of non-European ancestry and

construct locally representative reference databases [19,20]. On this foundation, the development and internal validation of FDP assays should follow established guidelines—such as the SWGDAM Validation Guidelines for DNA Analysis Method [74] and be conducted in laboratories accredited under ISO/IEC 17 025 [75]. The scope of cases in which FDP is deployed should be carefully restricted. Given public distrust regarding the use of FGG in non-violent offenses [68], its application should be prioritized in violent crimes such as homicides, particularly in cold cases [7,37] where conventional investigative methods have failed. FDP reports must adopt standardized, probabilistic language and clearly articulate the limitations of each prediction. Additionally, FDP should draw lessons from traditional forensic practices; for example, appearance prediction could incorporate the approach used in hand-drawn facial composites, which emphasize salient features to facilitate recognition [76]. Finally, fostering a more accurate understanding of FDP technologies among law enforcement personnel and the general public requires targeted training and science communication efforts [72].

Author contribution statement

Caixia Li:Conceptualization, writing–review & editing, project administration, and funding acquisition; Hao Zhang: Writing original draft; Chao Liu: Project administration, review & editing; Lan Hu: project administration, review & editing

Acknowledgements

This work is funded by the Fundamental Research Fund for Institute of Forensic Science (Nos. 2026JB025 and 2025JB032).

Declaration of competing interest

All the contributing authors report no conflict of interest in this work.

Ethics statement

Not applicable.

Use of AI statement

During the preparation of this work, the authors used DeepSeek-V4-Pro in order to polish the English language. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

References

[1] Butler, J. M. INTERPOL review of forensic biology and DNA, 2023-2025. *Forensic Science International: Synergy*, **2026**, 13: 100705. https://doi.org/10.1016/j.fsisyn.2026.100705

[2] Zhao, W. T., Jiang, L., Li, C. X. Research progress on forensic molecular phenotyping in the age of precision medicine. *Fa Yi Xue Za Zhi*, **2019**, 35(5): 525–530.

[3] Branicki, W. DNA testing for investigative purposes: description of the perpetrator. *Problems of Forensic Sciences*, **2024**, 137: 17–29. https://doi.org/10.4467/12307483PFS.24.002.19858

[4] Kayser, M. Forensic genetics in the omics era. *Nature Reviews Genetics*, **2026**, 27(2): 170–186. https://doi.org/10.1038/s41576-025-00896-1

[5] Schneider, P. M., Prainsack, B., Kayser, M. The use of forensic DNA phenotyping in predicting appearance and biogeographic ancestry. *Deutsches Ärzteblatt International*, **2019**, 116: 873–880. https://doi.org/10.3238/arztebl.2019.0873

[6] Kidd, K. K., Speed, W. C., Pakstis, A. J., Furtado, M. R., Fang, R. X., Madbouly, A., Maiers, M., Middha, M., Friedlaender, F. R., Kidd, J. R. Progress toward an efficient panel of SNPs for ancestry inference. *Forensic Science International: Genetics*, **2014**, 10: 23–32. https://doi.org/10.1016/j.fsigen.2014.01.002

[7] Liu, J., Ma, M., Wei, Y. L., Li, Y. X., Jiang, L., Wang, G. Q., Ye, J., Ji, A. Q., Li, C. X. Forensic SNP genealogical inference helping ferret out a 14-year long unsolved cold case. *Forensic Science and Technology*, **2021**, 46(6): 652–656.

[8] Liu, J., Li, J., Yang, L., Guan, S.-S., Wei, Y.-L., Zhao, W.-T., Jiang, L., Zhao, D., Li, C.-X. Accuracy research on the distant kinship relationship prediction by IBD algorithm. *Progress in Biochemistry and Biophysics*, **2023**, 50(12): 2980–2990.

[9] Walsh, S., Lindenbergh, A., Zuniga, S. B., Sijen, T., de Knijff, P., Kayser, M., Ballantyne, K. N. Developmental validation of the IrisPlex system: Determination of blue and brown iris colour for forensic intelligence. *Forensic Science International: Genetics*, **2011**, 5(5): 464–471. https://doi.org/10.1016/j.fsigen.2010.09.008

[10] Walsh, S., Liu, F., Wollstein, A., Kovatsi, L., Ralf, A., Kosiniak-Kamysz, A., Branicki, W., Kayser, M. The HIrisPlex system for simultaneous prediction of hair and eye colour from DNA. *Forensic Science International: Genetics*, **2013**, 7(1): 98–115. https://doi.org/10.1016/j.fsigen.2012.07.005

[11] Breslin, K., Wills, B., Ralf, A., Ventayol Garcia, M., Kukla-Bartoszek, M., Pospiech, E., Freire-Aradas, A., Xavier, C., Ingold, S., de la Puente, M. et al. HIrisPlex-S system for eye, hair, and skin color prediction from DNA: Massively parallel sequencing solutions for two common forensically used platforms. *Forensic Science International: Genetics*, **2019**, 43: 102152. https://doi.org/10.1016/j.fsigen.2019.102152

[12] White, J. D., Indencleef, K., Naqvi, S., Eller, R. J., Hoskens, H., Roosenboom, J., Lee, M. K., Li, J. R., Mohammed, J., Richmond, S. et al. Insights into the genetic architecture of the human face. *Nature Genetics*, **2021**, 53(1): 45–53. https://doi.org/10.1038/s41588-020-00741-7

[13] Boullón-Cassau, M., Ambroa-Conde, A., Casares de Cal, M. A., Gómez-Tato, A., Mosquera-Miguel, A., Ruiz-Ramírez, J., Cabrejas-Olalla, A., González-Bao, J., Casanova-Adán, L., de la Puente, M. et al. Exploring legal age estimation using DNA methylation. *Forensic Science International: Genetics*, **2025**, 74: 103142. https://doi.org/10.1016/j.fsigen.2024.103142

[14] Cheung, E. Y. Y., Gahan, M. E., McNevin, D. Prediction of biogeographical ancestry from genotype: A comparison of classifiers. *International Journal of Legal Medicine*, **2017**, 131(4): 901–912. https://doi.org/10.1007/s00414-016-1504-3

[15] Wei, Y. L., Wei, L., Zhao, L., Sun, Q. F., Jiang, L., Zhang, T., Liu, H. B., Chen, J. G., Ye, J., Hu, L. et al. A single-tube 27-plex SNP assay for estimating individual ancestry and admixture from three continents. *International Journal of Legal Medicine*, **2016**, 130(1): 27–37. https://doi.org/10.1007/s00414-015-1183-5

[16] Li, C. X., Pakstis, A. J., Jiang, L., Wei, Y. L., Sun, Q. F., Wu, H., Bulbul, O., Wang, P., Kang, L. L., Kidd, J. R. et al. A panel of 74 AISNPs: Improved ancestry inference within Eastern Asia. *Forensic Science International: Genetics*, **2016**, 23: 101–110. https://doi.org/10.1016/j.fsigen.2016.04.002

[17] Resutik, P., Aeschbacher, S., Krützen, M., Kratzer, A., Haas, C., Phillips, C., Arora, N. Comparative evaluation of the MAPlex, Precision ID Ancestry Panel, and VISAGE Basic Tool for biogeographical ancestry inference. *Forensic Science International: Genetics*, **2023**, 64: 102850. https://doi.org/10.1016/

j.fsigen.2023.102850

[18] Bird, N., King, T., Hellenthal, G. Power and limitations of inferring genetic ancestry. *Annals of Human Genetics*, **2025**, 89(5): 264–273. https://doi.org/10.1111/ahg.70007

[19] Wang, M. G., Duan, S. H., Sun, Q. X., Tang, R. K., Luo, L. T., Zhong, J., Sabitov, Z., Consortium, C., Wei, L. H., Liu, C. et al. Whole-genome sequencing pilot of the central Asian genomic diversity project reveals distinct histories, adaptation, and introgression. *Advanced Science*, **2026**: e76320. https://doi.org/10.1002/advs.76320

[20] He, G. L., Yao, H. B., Duan, S. H., Luo, L. T., Sun, Q. X., Tang, R. K., Chen, J., Wang, Z. Y., Sun, Y. T., Li, X. P. et al. Pilot work of the 10K Chinese People Genomic Diversity Project along the Silk Road suggests a complex east-west admixture landscape and biological adaptations. *Science China Life Sciences*, **2025**, 68(4): 914–933. https://doi.org/10.1007/s11427-024-2748-4

[21] Pritchard, J. K., Stephens, M., Donnelly, P. Inference of population structure using multilocus genotype data. *Genetics*, **2000**, 155(2): 945–959. https://doi.org/10.1093/genetics/155.2.945

[22] Tang, H., Peng, J., Wang, P., Risch, N. J. Estimation of individual admixture: Analytical and study design considerations. *Genetic Epidemiology*, **2005**, 28(4): 289–301. https://doi.org/10.1002/gepi.20064

[23] Alladio, E., Poggiali, B., Cosenza, G., Cisana, S., Omedei, M., Garofano, P., Pilli, E. Machine Learning overview for biogeographical ancestry prediction - a PLS-DA approach. *Forensic Science International: Genetics Supplement Series*, **2022**, 8: 306–307. https://doi.org/10.1016/j.fsigss.2022.10.071

[24] Casanova-Adán, L., González-Bao, J., Mosquera-Miguel, A., Ambroa-Conde, A., Ruiz-Ramírez, J., Cabrejas-Olalla, A., Boullón-Cassau, M., Freire-Aradas, A., Rodríguez-López, A., Phillips, C., et al. Exploring STRs as an approach to continental biogeographical ancestry inference. In: *Proceedings of the 30th Congress of the International Society for Forensic Genetics* Universidade de Santiago de Compostela, **2025**: 161–168. https://doi.org/10.15304/cc.2025.1869

[25] Neyra-Rivera, C., Sam R. L., Ramos, E. D., De La Cruz-Vargas, J. A., Gusmão, L. Analysis of the ancestry and admixture proportions in Peruvian Mestizo populations using an insertion deletion multiplex. In: Proceedings of the 30th Congress of the International Society for Forensic Genetics Universidade de Santiago de Compostela, **2025**: 535–540. https://doi.org/10.15304/cc.2025.1869

[26] Li, Y. J., Zhuang, B., Han, J. P., Feng, J. H., Zhao, L. J., Liu, Z. Y., Jiang, L., Yao, Y. R., Xing, W. L., Li, C. X. A versatile "sample-in-answer-out" microfluidic system for combined detection of multi-functional forensic human genetic markers. *Small*, **2025**, 21(49): e07944. https://doi.org/10.1002/smll.202507944

[27] de Barros Rodrigues, M. L., Rodrigues, M. P., Norton, H. L., Mendes-Junior, C. T., Simões, A. L., Lawson, D. J. Large-scale selection of highly informative microhaplotypes for ancestry inference and population specific informativeness. *Forensic Science International: Genetics*, **2025**, 74: 103153. https://doi.org/10.1016/j.fsigen.2024.103153

[28] Zeng, K., Liu, J., Zhao, D., Li, C. X. Research progress in forensic SNP genealogy inference. *Chinese Journal of Forensic Medicine*, **2023**, 38(5): 486–491, 500. (in Chinese)

[29] Zhang, J. Y., Sun, S. Y., Zhang, F. H., Niu, Q. S. Current status and prospects of forensic genetic genealogy. *Guangdong Gongan Keji*, **2025**, 33(4): 21–23. (in Chinese)

[30] Chen, J. T., Chen, H. Forensic genealogy: Identifying the unknown through long-range kinship networks. *Science Magazine*, **2026**, 78(2): 21–24, 69. (in Chinese)

[31] Greytak, E. M., Moore, C., Armentrout, S. L. Genetic genealogy for cold case and active investigations. *Forensic Science International*, **2019**, 299: 103–113. https://doi.org/10.1016/j.forsciint.2019.03.039

[32] Erlich, Y., Shor, T., Pe'er, I., Carmi, S. Identity inference of genomic data using long-range familial searches. *Science*, **2018**, 362(6415): 690–694. https://doi.org/10.1126/science.aau4832

[33] Kling, D., Phillips, C., Kennett, D., Tillmar, A. Investigative genetic genealogy: Current methods, knowledge and practice. *Forensic Science International: Genetics*, **2021**, 52: 102474. https://doi.org/10.1016/j.fsigen.2021.102474

[34] de Vries, J. H., Kling, D., Vidaki, A., Arp, P., Kalamara, V., Verbiest, M. M. P. J., Piniewska-Róg, D., Parsons, T. J., Uitterlinden, A. G., Kayser, M. Impact of SNP microarray analysis of compromised DNA on kinship classification success in the context of investigative genetic genealogy. *Forensic Science International: Genetics*, **2022**, 56: 102625. https://doi.org/10.1016/j.fsigen.2021.102625

[35] Lu, J. W., Liu, J., Li, J., Jiang, L., Ni, M., Zhao, W. T., Dai, L., Zhao, C. T., Li, C. X. Forensic investigative genetic genealogy based on low-quality DNA whole genome sequencing data. *Forensic Science International: Genetics*, **2026**, 82: 103417. https://doi.org/10.1016/j.fsigen.2025.103417

[36] Loreille, O., Tillmar, A., Brandhagen, M. D., Otterstatter, L., Irwin, J. A. Improved DNA extraction and illumina sequencing of DNA recovered from aged rootless hair shafts found in relics associated with the romanov family. *Genes*, **2022**, 13(2): 202. https://doi.org/10.3390/genes13020202

[37] Tillmar, A., Fagerholm, S. A., Staaf, J., Sjölund, P., Ansell, R. Getting the conclusive lead with investigative genetic genealogy–A successful case study of a 16 year old double murder in Sweden. *Forensic Science International: Genetics*, **2021**, 53: 102525. https://doi.org/10.1016/j.fsigen.2021.102525

[38] Xie, Q., Zhao, W. T., Liu, W. H., Zhao, Y. R., Chen, X. Z., Li, J., Fang, Z. X., Hu, L., Li, C. X. Forensic SNP genealogy inference using whole genome sequencing data of varying depths. *Forensic Science International: Genetics*, **2025**, 79: 103296. https://doi.org/10.1016/j.fsigen.2025.103296

[39] Watson, J., McNevin, D., Grisedale, K., Spiden, M., Seddon, S., Ward, J. Operationalisation of the ForenSeq® Kintelligence Kit for Australian unidentified and missing persons casework. *Forensic Science International: Genetics*, **2024**, 68: 102972. https://doi.org/10.1016/j.fsigen.2023.102972

[40] Zeng, K., Zhao, W. T., Fang, Z. X., Li, J., Liu, J., Zhao, D., Zhu, B. F., Li, C. X. Development and validation of a capture sequencing panel containing 9000 SNPs for inferring distant relatives in East Asian populations. *Forensic Science International: Genetics*, **2026**, 81: 103341. https://doi.org/10.1016/j.fsigen.2025.103341

[41] Dai, L., Tang, Z.-C., Jia, Z., Jiang, L., Zhao, C.-T., Zhao, Z.-Y., Zhao, W.-T., Li, C.-X. SNP density impact on kinship inference and IBS-machine learning optimization. *Yi Chuan*, **2026**, 48(6): 570–588. (in Chinese)

[42] Alshehhi, A., Almarzooqi, A., Alhammadi, K., Werghi, N., Tay, G. K., Alsafar, H. Advancement in human face prediction using DNA. *Genes*, **2023**, 14(1): 136. https://doi.org/10.3390/genes14010136

[43] Chemeris, A. V., Khalikov, A. A., Garafutdinov, R. R., Chemeris, D. A., Sakhabutdinova, A. R., Khaliullina, A. F., Galyautdinov, R. R., Sagidullin, R. H., Aminev, F. G. Genetic problems of DNA portrait as part of DNA phenotyping: A review. *Russian Journal of Forensic Medicine*, **2024**, 10(3): 398–410. https://doi.org/10.17816/fm16167

[44] Naqvi, S., Hoskens, H., Wilke, F., Weinberg, S. M., Shaffer, J. R., Walsh, S., Shriver, M. D., Wysocka, J., Claes, P. Decoding the human face: Progress and challenges in understanding the genetics of craniofacial morphology. *Annual Review of Genomics and*

Human Genetics, **2022**, 23: 383–412. https://doi.org/10.1146/annurev-genom-120121-102607

[45] Sforza, C., de Menezes, M., Ferrario, V. Soft- and hard-tissue facial anthropometry in three dimensions: What's new. *Journal of Anthropological Sciences*, **2013**, 91: 159–184. https://doi.org/10.4436/jass.91007

[46] Selleri, L., Rijli, F. M. Shaping faces: Genetic and epigenetic control of craniofacial morphogenesis. *Nature Reviews Genetics*, **2023**, 24(9): 610–626. https://doi.org/10.1038/s41576-023-00594-w

[47] Huang, Y., Li, D., Qiao, L., Peng, Q. Q., Wu, S. J., Zhang, Y. J., Tan, J. Z., Xu, S. H., Jin, L., Wang, S. J., et al. A genome-wide association study of facial morphology identifies novel genetic loci in Han Chinese. *Journal of Genetic Genomics*, **2021**, 48(3): 198–207. https://doi.org/10.1016/j.jgg.2020.10.004

[48] Zhang, M. F., Wu, S. J., Du, S. Y., Qian, W., Chen, J. Y., Qiao, L., Yang, Y. J., Tan, J. Z., Yuan, Z. Y., Peng, Q. Q. et al. Genetic variants underlying differences in facial morphology in East Asian and European populations. *Nature Genetics*, **2022**, 54(4): 403–411. https://doi.org/10.1038/s41588-022-01038-7

[49] Kong, Y. Q., Zou, L. X., Zhao, Y. R., Ji, P. C., Chen, X. Z., Wu, S. Y., Zhao, W. T., Wei, Y. L., Li, C. X. Integrative genomic and epigenomic dissection of facial morphology in Chinese populations. *Genomics, Proteomics & Bioinformatics*, **2026**: qzag041. https://doi.org/10.1093/gpbjnl/qzag041

[50] Xiong, Z. Y., Gao, X. J., Chen, Y., Feng, Z. Y., Pan, S. Y., Lu, H. J., Uitterlinden, A. G., Nijsten, T., Ikram, A., Rivadeneira, F. et al. Combining genome-wide association studies highlight novel loci involved in human facial variation. *Nature Communications*, **2022**, 13: 7832. https://doi.org/10.1038/s41467-022-35328-9

[51] Jiao, M. Q., Li, J. R., Zhong, B. X., Du, S. Y., Li, S. N., Zhang, M. F., Zhang, Q. B., Liang, Z. M., Liu, F., Zuo, C. M. et al. *De novo* reconstruction of 3D human facial images from DNA sequence. *Advanced Science*, **2025**, 12(29): 2414507. https://doi.org/10.1002/advs.202414507

[52] Yengo, L., Vedantam, S., Marouli, E., Sidorenko, J., Bartell, E., Sakaue, S., Graff, M., Eliasen, A. U., Jiang, Y., Raghavan, S. et al. A saturated map of common genetic variants associated with human height. *Nature*, **2022**, 610(7933): 704–712. https://doi.org/10.1038/s41586-022-05275-y

[53] Weng, H. Y., Jiang, L., Zheng, Z. F., Tang, K. C., Zhang, D., Zhao, W. T., Song, J., Bi, M. X., Tang, S. W., Li, T. et al. Partially connected neural networks for complex trait prediction: Application to human height. *Human Genetics*, **2026**, 145: 4. https://doi.org/10.1007/s00439-025-02794-2

[54] Wang, Z. H., Fu, G. P., Ma, G. J., Wang, C. Y., Wang, Q., Lu, C. L., Fu, L. H., Zhang, X. J., Cong, B., Li, S. J. The association between DNA methylation and human height and a prospective model of DNA methylation-based height prediction. *Human Genetics*, **2024**, 143(3): 401–421. https://doi.org/10.1007/s00439-024-02659-0

[55] Doerfler, W. DNA methylation—a regulatory signal in eukaryotic gene expression. *Journal of General Virology*, **1981**, 57(1): 1–20. https://doi.org/10.1099/0022-1317-57-1-1

[56] Cao, Y. L., Bai, Y. L., Yuan, T. J., Song, L. Y., Fan, Y., Ren, L. H., Song, W. L., Peng, J. H., An, R., Gu, Q. Q. et al. Single-cell bisulfite-free 5mC and 5hmC sequencing with high sensitivity and scalability. *Proceedings of the National Academy of Sciences of the United States of America*, **2023**, 120(49): e2310367120. https://doi.org/10.1073/pnas.2310367120

[57] Carlsen, L., Holländer, O., Danzer, M. F., Vennemann, M., Augustin, C. DNA methylation-based age estimation for adults and minors: Considering sex-specific differences and non-linear correlations. *International Journal of Legal Medicine*, **2023**, 137(3): 635–643. https://doi.org/10.1007/s00414-023-02967-6

[58] Lee, J. E., Lee, J. M., Naue, J., Fleckhaus, J., Freire-Aradas, A., Neubauer, J., Pośpiech, E., McCord, B., Kalamara, V., Gauthier, Q. et al. A collaborative exercise on DNA methylation-based age prediction and body fluid typing. *Forensic Science International: Genetics*, **2022**, 57: 102656. https://doi.org/10.1016/j.fsigen.2021.102656

[59] Zhang, J., Zhao, C. T., Jiang, L., Fan, P., Li, C. X., Liu, R., Zhao, W. T. A five-color multiplex droplet digital PCR assay with support vector regression for forensic age prediction. *Electrophoresis*, **2026**, 47(6): 489–501. https://doi.org/10.1002/elps.70102

[60] Castagnola, M. J., Medina-Paz, F., Zapico, S. C. Uncovering forensic evidence: A path to age estimation through DNA methylation. *International Journal of Molecular Sciences*, **2024**, 25(9): 4917. https://doi.org/10.3390/ijms25094917

[61] Xavier, C., de la Puente, M., Mosquera-Miguel, A., Freire-Aradas, A., Kalamara, V., Vidaki, A., E Gross, T., Revoir, A., Pośpiech, E., Kartasińska, E. et al. Development and validation of the VISAGE AmpliSeq basic tool to predict appearance and ancestry from DNA. *Forensic Science International: Genetics*, **2020**, 48: 102336. https://doi.org/10.1016/j.fsigen.2020.102336

[62] Ruiz-Ramírez, J., de la Puente, M., Xavier, C., Ambroa-Conde, A., Álvarez-Dios, J., Freire-Aradas, A., Mosquera-Miguel, A., Ralf, A., Amory, C., Katsara, M. A. et al. Development and evaluations of the ancestry informative markers of the VISAGE Enhanced Tool for Appearance and Ancestry. *Forensic Science International: Genetics*, **2023**, 64: 102853. https://doi.org/10.1016/j.fsigen.2023.102853

[63] Terrado-Ortuño, N., May, P. Forensic DNA phenotyping: A review on SNP panels, genotyping techniques, and prediction models. *Forensic Sciences Research*, **2025**, 10(1): owae013. https://doi.org/10.1093/fsr/owae013

[64] Karabegović, I., Portilla-Fernandez, E., Li, Y., Ma, J. T., Maas, S. C. E., Sun, D. K., Hu, E. A., Kühnel, B., Zhang, Y., Ambatipudi, S. et al. Epigenome-wide association meta-analysis of DNA methylation with coffee and tea consumption. *Nature Communications*, **2021**, 12: 2830. https://doi.org/10.1038/s41467-021-22752-6

[65] You, C. L., Wu, S. J., Zheng, S. C., Zhu, T. Y., Jing, H., Flagg, K., Wang, G. Y., Jin, L., Wang, S. J., Teschendorff, A. E. A cell-type deconvolution meta-analysis of whole blood EWAS reveals lineage-specific smoking-associated DNA methylation changes. *Nature Communications*, **2020**, 11: 4779. https://doi.org/10.1038/s41467-020-18618-y

[66] Gisladottir, R. S., Helgason, A., Halldorsson, B. V., Helgason, H., Borsky, M., Chien, Y. R., Gudnason, J., Gudjonsson, S. A., Moisik, S., Dan, D. D. et al. Sequence variants affecting voice pitch in humans. *Science Advances*, **2023**, 9(23): eabq2969. https://doi.org/10.1126/sciadv.abq2969

[67] Li, J. X., Glover, J. D., Zhang, H. G., Peng, M. F., Tan, J. Z., Mallick, C. B., Hou, D., Yang, Y. J., Wu, S. J., Liu, Y., et al. Limb development genes underlie variation in human fingerprint patterns. *Cell*, **2022**, 185(1): 95–112.e18. https://doi.org/10.1016/j.cell.2021.12.008

[68] Committee on Law and Justice, Computer Science and Telecommunications Board, Division of Behavioral and Social Sciences and Education, et al. Law Enforcement Use of Probabilistic Genotyping, Forensic DNA Phenotyping, and Forensic Investigative Genetic Genealogy Technologies: Proceedings of a Workshop. Washington, D.C.: National Academies Press; **2024**.

[69] Granja, R., Machado, H. Forensic DNA phenotyping and its politics of legitimation and contestation: Views of forensic geneticists in Europe. *Social Studies of Science*, **2023**, 53(6): 850–868. https://doi.org/10.1177/0306312720945033

[70] Coquet, M., Terrado-Ortuño, N. Forensic DNA phenotyping: Privacy breach, bias reification and the pitfalls of abstract assessments of rights. *International Journal of Police Science & Management*, **2023**, 25(3): 262–279. https://doi.org/10.1177/14613557231184707

[71] Kaiser, J. New federal rules limit police searches of family tree DNA databases. *Science*, **2019**: eaaz6336. https://doi.org/10.1126/science.aaz6336

[72] Gareau-Léonard, A., Mousseau, V., Crispino, F., Milot, E. Forensic DNA Phenotyping: Examining knowledge and operational view from police officers. *Forensic Science International: Synergy*, **2025**, 10: 100586. https://doi.org/10.1016/j.fsisyn.2025.100586

[73] Wickenheiser, R. A., Naugle, J., Hoey, B., Nowlin, R., Kumar, S. A., Minton, A., Allen, L., Glynn, C. National technology validation and implementation collaborative (NTVIC) policies and procedures for forensic investigative genetic genealogy (FIGG). *Forensic Science International: Synergy*, **2023**, 6: 100316. https://doi.org/10.1016/j.fsisyn.2023.100316

[74] Scientific Working Group on DNA Analysis Methods (SWGDAM). SWGDAM Validation Guidelines for DNA Analysis Methods. *SWGDAM*. **2016**.

[75] ISO/IEC 17025:2017. General requirements for the competence of testing and calibration laboratories. **2017**.

[76] Hopman, R., Bleumink, R. Between pencils and genetic markers: Rethinking innovation in policing through forensic face-making technologies. *International Journal of Police Science & Management*, **2023**, 25(3): 280–296. https://doi.org/10.1177/14613557231173213

© The Editorial Office of *Forensic Science Advances* 2026. Published by Tsinghua University Press. The articles published in this open access journal are distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International Public License (CC BY-NC-ND 4.0). See https://creativecommons.org/licenses/by-nc-nd/4.0/.