Review Article

Genomic Portraiture: The Science and Ethics of DNA Phenotyping in Identity Prediction

Anubhav Shukla[1], Prachi Tripathi[2], Piyush Anand[3], Oshin Hathi[4], Keya Pandey[5]

[1]M.Sc. Forensic Science Department of Anthropology, University of Lucknow, UP, [2,3]PG diploma In DNA Forensics Uttar Pradesh State Institute of Forensics Sciences, Lucknow, Uttar Pradesh, [4]Assistant Professor, Department of Anthropology, University of Lucknow, [5]Professor, Department of Anthropology, University of Lucknow.

How to cite this article: Anubhav Shukla, Prachi Tripathi, Piyush Anand et. al. Genomic Portraiture: The Science and Ethics of DNA Phenotyping in Identity Prediction. Indian Journal of Forensic Medicine and Toxicology/ Volume 19 No. 3, July - September 2025.

Abstract

Forensic DNA phenotyping (FDP) has emerged as a groundbreaking tool in criminal investigations, enabling the prediction of externally visible characteristics (EVCs) such as eye color, hair color, and skin color from DNA samples. This capability provides valuable leads in cases with limited traditional evidence. The IrisPlex system, developed to predict eye color with high accuracy, marked a significant advancement in FDP, setting the stage for further developments in the field. This review traces the evolution of FDP technologies, starting with the IrisPlex system, which employs specific single nucleotide polymorphism (SNP) markers to predict eye color. In forensic DNA phenotyping (FDP), the transition from IrisPlex to HIrisPlex and HIrisPlex-S has greatly improved the process. Eye color was correctly predicted by IrisPlex, and forensic application was increased by HIrisPlex's addition of hair color prediction. Skin color prediction is incorporated into the most recent HIrisPlex-S system, which offers a more thorough phenotypic profile. These developments enhance the precision and pertinence of criminal investigations. A new era in forensic science should be ushered in by future studies that improve predictive accuracy, broaden characteristic analysis, and tackle ethical issues.

Keywords: Forensic DNA Phenotyping (FDP), External Visible Characters (EVCs), Single Nucleotide Polymorphism (SNPs), HIrisplex-S, Criminal Investigation

Introduction

In general, forensic DNA phenotyping (FDP) is a collection of methods used to analyze biological materials taken from crime scenes in order to deduce the biogeographical heritage and externally visible physical characteristics of human Individual, such as eye, hair, and skin color. FDP technology have been used in a few high-profile instances across several jurisdictions to gather intelligence for criminal investigations and give information pertinent to particular target[1]. The use of forensic genetic

Corresponding Author: Anubhav Shukla, M.Sc. Forensic Science Department of Anthropology, University of Lucknow, UP.

E-mail: anubhavshukla114@gmail.com

Submission date: February 24, 2025 **Acceptance date:** April 14 , 2025 **Published date: July 10, 2025**

technologies in the criminal justice system has changed historically with the introduction of FDP. First, it causes forensic science to refocus its attention from creating evidence to producing intelligence that is useful for criminal investigations. Second, by grouping "suspect" groups that have genetic heritage and/or outwardly noticeable traits in common, FDP shifts the locus from individualization that is, the identification of single individuals to collectivization. Cole refers to this process as the "convergence of individual and collective identity." Third, single nucleotide polymorphisms (SNPs) are used in FDP. In the realm of forensic genetics, these genetic markers are distinguished by their informational richness, integrate appearance, and race[1]. Outside the realm of forensics, DNA prediction is used in paleogenetics and anthropology to reconstruct the appearance of deceased individuals using (ancient) DNA analysis of (old) human remains. A number of EVCs, including eye, hair, skin color, body height, male head hair loss, head hair type, and face shape, have shown improvements in the genetic understanding of human appearance in recent years. The first genes associated with other EVCs, such as ear morphology and facial hair and graying have recently been discovered[2].

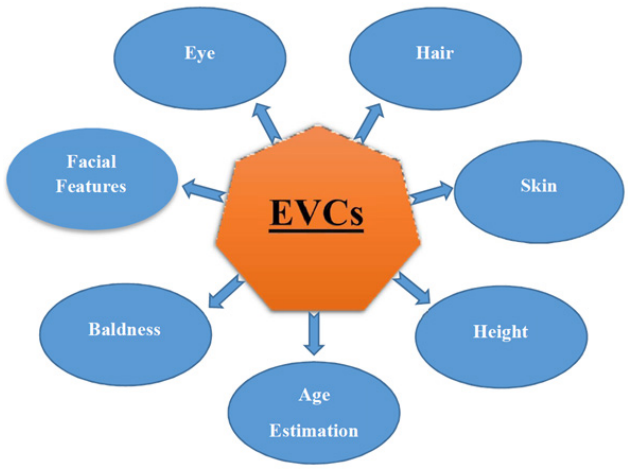

Fig 1. Various Externally Visible Characteristics Predicted by FDP

Actually, FDP analysis may be very important in the forensic domain if normal DNA analysis proves to be unfeasible in the lack of a comparative sample or extensively decomposed, or traces of DNA found on crime scene. By analyzing biological material from unidentified bodies or from identified biological evidence discovered during crime scene investigation, phenotypic studies use DNA sequence analysis to predict exterior somatic features. It helps locate victims of disasters and missing people. The genetic basis of human pigmentation features is well recognized since they are simpler EVCs that are regulated by a small number of genes. Similar to common diseases, complex features involve a large number of genes, each of which contributes only slightly to the variance in phenotype.While EVCs are considered complex traits, multiple genes contribute to the phenotype in addition to environmental factors[3].

Predictions of Phenotype Characteristics Using DNA Markers

FDP is the technique of determining externally visible characteristics for forensic purposes from biological materials. Examples of these include height, facial features, complexion, iris, and hair color. FDP provides a lot of information about the subject (victims and suspects) without requiring a reference sample while comparing EVCs of an individual. Two unknown/reference samples must be compared in order to perform STR profiling analysis. If such a pair is not found, the only option is to search a DNA database that has the profiles of potential suspects. The STR markers' repeats, which range in length from approximately 100 to 300 base pairs, present another drawback. In many cases, the biological evidence from the crime scene is too damaged that it is impractical to extract DNA samples that would provide sufficient information for a precise identification. Due to the challenges associated with applying classic STR markers methods in some circumstances, a number of studies have made it possible for police investigations to benefit from the use of genetic predictions of EVCs for victim identification and suspect tracking.

Eye

Eye color is the most important visible character of humans in forensic sample analysis. It has a very diverse color in the population. The quantity of iris pigment melanin and melanosomes number in the iris's external layer predict this color variation. Blue eyes have low melanin pigment and melanosomes than dark eyes[4]. A complex genetic trait, eye color deviates from the classical paths of Mendelian inheritance and is now considered to be polygenic, i.e., controlled by interactions between multiple genes, any one of which may be incompletely dominant and regulated by other genes[5]. The

IrisPlex System, which comprises of six single nucleotide polymorphisms spread across color genes was one of the earliest phenotyping procedures developed and validated. This approach can identify between brown & blue eyes with a large level of perfection (>90%) with around 30pg DNA input, which has been tested in both homogenous and mixed populations[4]. The European DNA Profiling Group (EDNAP) investigated the IrisPlex assay in a multi-center exercise involving approximately 20 labs, and it was discovered to be simple to apply and extremely authentic. Although, the exactness was lower in the Asiatic population, indicating that future research should be conducted in this community. Upcoming study is also required to uncover new genetic variants and improve the accuracy of existing variants[6].

Hair

Hair colour is among the most obvious distinctive feature with a range of phenotypes. The two types of melanin that cause the majority of variations in hair colour are red/yellow pheomelanin and brown/black eumelanin. People with red hair have more pheomelanin than eumelanin in their hair, whereas people with dark hair have more eumelanin and people with blond hair have less of both types of melanin. In 2013, the HIrisplex software was developed, which added 18 hair color markers to the six pre-existing Irisplex SNPs[4]. Current hair prediction models suffer a hurdle in that they are only accurate for adult populations. As a result, precise hair prediction is difficult for those whose hair color fluctuates over their lives. In the future, quantitative hair color prediction should be prioritized because there is less study on this topic[6].

Skin

Skin pigmentation variation originated as an evolutionary reaction to the strength of ultraviolet light across globe regions. Using 36 indications spread across 16 pigmentation genes, a global prediction model known as the HIrisPlex-S system was developed with this evolutionary barrier taken into consideration. The prediction accuracy ranged from 83% to 97% for the three-category scale and 72% to 97% for the five-category scale. For some of these genes, including HERC2, SLC24A5 and SLC45A2, some of these associations had previously been reported in admixed populations, suggesting that they may have potential uses in the future.

Facial Symmetry

When investigating phenotyping, one of the primary goals is to anticipate the facial form of each EVC. Facial symmetry measurement described as the "Quantification of some facial morphology" such as lip thickness, forehead height or chin protrusion, nose wing width, nose tip shape. Face morphology is investigated using distances between facial markers such as lip thickness, nose width[8]. Some of the genetic markers linked to face traits are first discovered in investigations on syndromes and facial abnormalities. Some of these indicators are then connected to craniofacial development, resulting in the normal variation in facial appearance. For example, the PAX3 gene encodes a transcription factor found in neural crest cells that was linked to Waardenburg syndrome and later identified with the nasion location[9]. Other potential genes, such as PRDM16 and TP63, have been found using PAX3-like patterns. However, like with height determination, each of these genetic markers appears to contribute just slightly to overall facial morphology[4]. The investigation of facial symmetry consist of, LYPLAL1 rs5781117, PRDM16 rs4648379, DKK1 rs1194708, TNFSF12 rs80067372, and SUPT3H rs227833, EDAR rs3827760, PAX3 rs7559271, CACNA2D3 rs56063440[10].

Fig. SNP markers for Eye, Hair and Skin[7]

Age Estimation

It has been discovered that utilising DNA methylation detection technologies is advantageous

for determining an individual's age. Over the course of a person's life, their levels of DNA methylation change; they peak in childhood and decline in adulthood. With as few as seven indicators, these alterations can be assessed and utilised to accurately determine an individual's age from biological samples of various origins (different tissues and bodily fluids) and situations (either from human remains or from a crime scene)[4]. studies have been done to determine the association between T-cell numbers and age as well as to predict the genes that generate age. Age is estimated by quantifying the sjTREC, as they decrease with age, as reported by Zubakov et al. (2010)[11],[12]. Recent advancements in the field of epigenetics have made it possible to estimate an individual's age thanks to the use of DNA methylation detection methods, since methylation directly affects age. Childhood is characterised by a higher degree of methylation, which declines with age. Plus, CpG candidate markers show a lot of potential. This variance can be utilised with biological samples to accurately determine an individual's age[6].

Baldness

FDP predicts male baldness by analyzing specific genetic markers associated with hair loss. The method focuses on genes like AR (androgen receptor) and EDA2R, which influence the likelihood of developing male pattern baldness. It is scientifically established that male pattern baldness, or androgenic alopecia, has a substantial hereditary component. The major loci potentially involved are those on the q12 region of the X chromosome, which contains AR/EDA2R genes directly linked to the production of androgen receptor and ectodysplasin A2 receptor, respectively, on the 20p11 region, and on the genes EBF1, TARDBP, and HDAC9 with predictive potential[4]. The role of 29 SNPs in determining MBP in European individuals of varying ages was confirmed by Marcińska et al: rs929626 in EBF1, rs12565727 (chr1), rs756853 in HDAC9, rs10502861 in SLC12A2, 8 SNPs on chromosome 20 (rs61374441, rs19980761, rs201571, rs6047844, rs913063, rs1160312, rs6113491 and rs2180439) and 17 SNPs on Xq12 (rs4827379, rs1385699, rs1352015, rs1041668, rs2497938, rs2497935, rs962458, rs6152, rs12396249, rs4827545, and rs7885198)[13].

Height

Only a small number of genes have been linked to human height up till 2008. The number of markers increased to 180 in 2010 and nearly 700 in 2014 after more association studies were conducted in 2008, which found 54 loci that directly correlated with variations in height. The majority of these genes are expressed in significant tissues like the growth plate and are connected to growth-signalling pathways like the fibroblast growth factor. Estimation of height is based on ACAN, DNM3, SDR16C5 EFEMP1, FBXW11, GH region, GHSR, GPR126, HHIP, HMGA1, HMGA1, IHH, LCORL, SOCS2, MICA, NOG, NPR3, PML, PPIF etc genes.

CURRENT PHENOTYPIC APPROACH

Sample Collection

Samples for FDP are collected in various ways depending on sample types. The samples and their collection methods are given below:

Table 1: DNA extraction from various biological samples (Richard Li)

Biological Sample	Collection Methods
Bone	Collect in sterilized container
Blood	Collect in EDTA vial or use sterilized swab
Hair	Collect in sterilized container
Tissue	Collect in sterilized container or normal saline
Semen	Collect in sterilized container by swabbing
Saliva	Collect in sterilized container by swabbing

DNA Extraction

- **Tissue -** Extraction of DNA from tissue can be performed using PCI (Phenol Chloroform Isoamyl alcohol) method. In this method, phenol is used for separating DNA from other cell debris. Isoamyl alcohol works as an anti-foaming agent in this method. Extracted DNA is purified using ethanol and eluted in Tris-EDTA for storage.
- **Bone -** Bones should be broken into small pieces or powder using TissueLyser. Bone samples must be treated with diluted bleach, and then irradiated with Ultraviolet light for 30 min, before being powdered as much as possible. Put the bone pieces or powder in EDTA for 2-3 days before performing DNA extraction. After 2-3 days, the PCI method or QIAmp DNA Investigator kit can be used for DNA extraction from bones.

- **Hair -** In hair, mtDNA and nuclear DNA are present which are used for analysis. The nuclear DNA is present in the root of hair which can be extracted using various methods such as PCI method or PrepFiler™ BTA Automated Forensic DNA Extraction Kit etc.

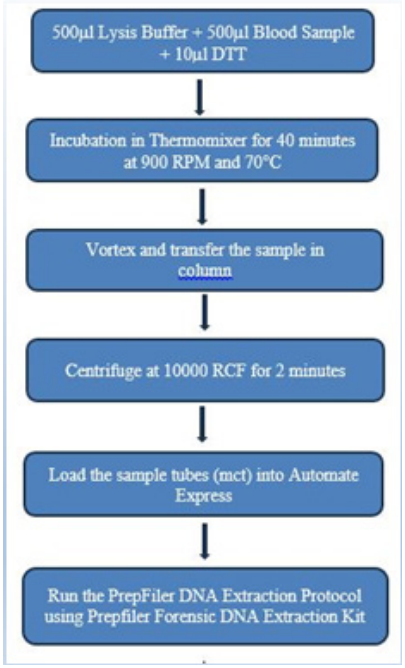

Fig. 3 DNA Extraction from Blood, Tissue, Saliva or Semen (Soft Tissue)

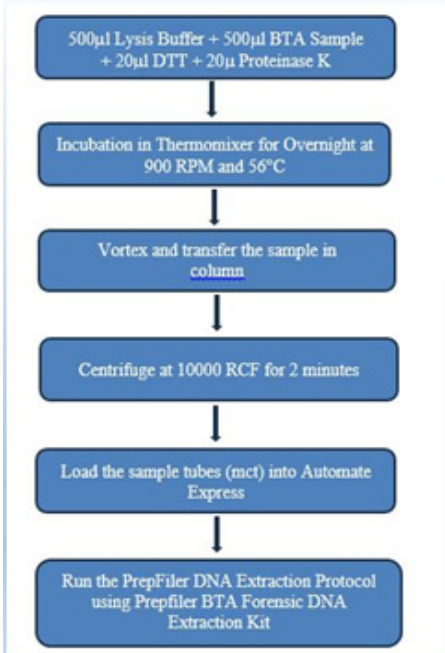

Fig. 4 DNA Extraction from Bone, Teeth or Adhesive (BTA, Hard Tissue)

Phenotyping

In cases where forensic DNA analysis is necessary, STR analysis is generally used for identification person of interest by generating STR profile of person and matching of this profile with reference STR profile or DNA databases such as CODIS. However, if both are unavailable for any reason, we can use DNA phenotyping technique for identification of a person.

The phenotyping for eye color prediction initially used the IrisPlex system having six SNPs; rs12913832, rs1800407, rs12896399, rs16891982, rs1393350 and rs12203592 from the HERC2, OCA2, SLC24A4, SLC45A2 (MATP), TYR and IRF4 genes respectively. The analysis of these SNPs follows a similar procedure as the HIrisPlex system[14].

The HIrisPlex system uses given steps for analyzing forensic samples, which includes hair –

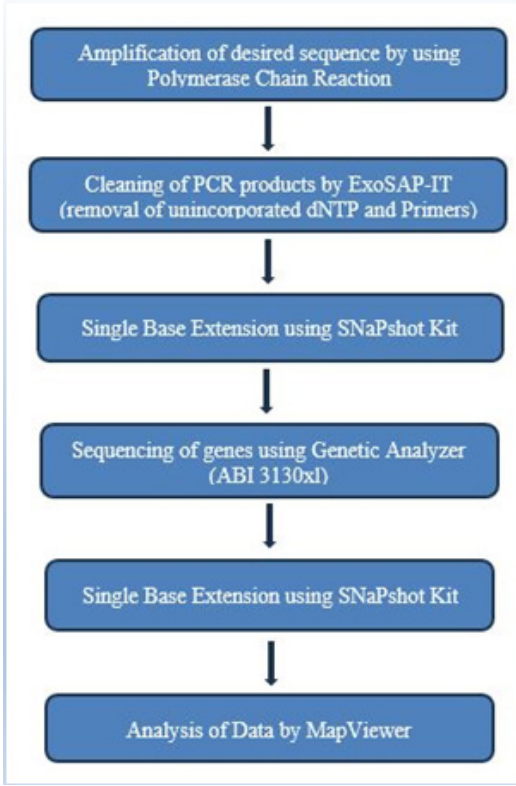

Fig. 5 Sample analysis using HIrisPlex System[16]

The skin colour prediction using HIrisPlex-S system is based on 36 SNPs from 16 genes (**SLC24A5**, **IRF4**, **MC1R**, **OCA2**, **SLC45A2**, **HERC2**, **TYR**, **RALY**, **DEF8**, **PIGU**, **ASIP**, **SLC24A4**, **TYRP1**, **KITLG**, **ANKRD11** and **BNC2**)[16]. The procedure of

the HIrisPlex-S system is performed in several steps, which includes sample collection, DNA extraction, PCR amplification, SBE, Gene sequencing and Phenotype prediction using 36 SNPs.

HOW DOES FDP DIFFER FROM GENOTYPING?

With the development of DNA fingerprinting from RFLP in the 1980s to STR analysis in the 1990s, forensic identification became quicker and more effective. CODIS is compared with STR profiles produced by genetic analyzers to look for matches. However, accuracy decreases in the absence of reference DNA. Forensic DNA Phenotyping (FDP) is employed in these situations. FDP examines SNPs connected to observable characteristics including eye color and facial features, in contrast to STR-based genotyping. Kits like SNapShot are used for sequencing and analysis, which helps in situations when there are no direct DNA matches. In FDP, the reference data is not required because with the help of SNP, The EVC's are predicted with more accuracy. So, if there is only a body part found on a crime scene such as hand, ear, hair etc. and the generated STR profile does not have any match, FDP can construct the EVC and predict the suspect or victim[17].

CHALLENGES AND LIMITATIONS

FDP has limitations due to its probabilistic nature and lack of high accuracy. It predicts the most probable appearance of a suspect or victim based on biological samples but is not conclusive. This is because genes responsible for externally visible characteristics (EVCs) are influenced by various factors, and traits can be modified through hair dye, contact lenses, or surgeries. Moreover, the genomic regions used for phenotyping may only be linked to EVCs rather than causing them. FDP also requires a sufficient DNA sample, and if the amount is too low (approx. 30-40 ng), the results become unreliable. In contrast, DNA genotyping is more accurate as it relies on direct matching, making it more useful in forensic analysis.

LEGAL ASPECTS:

In forensic DNA phenotyping, if the questioned is matched with an innocent person with a slight variation, it will cause a negative impact on the person's social life and not even someone's personal life but the society and its justice system that is directly correlated to us. The oppose of the application of FDP in social view is high, arguing that the information obtained by such analysis could result in racial and ethnic prejudice. In India and other countries, there is no legal framework for forensic DNA phenotyping. In DNA fingerprinting, the non-coding region of DNA is used for discrimination, but in FDP, the coding region is used, which makes it tough to establish and implement regulations.

ETHICAL ASPECTS:

In FDP, ethical considerations are a top priority. In situations involving FDP, the privacy of individuals may become public knowledge. One of the factors contributing to concerns is the fact that some diseases have genetic roots that an individual doesn't wanted to disclose, and FDP does not have any regulations to stop this from happening. Another ethics related question under consideration is the accuracy of FDP data. There is a lack of accuracy and high error rate in FDP, which can be caused by environmental factors or by other means. This affects the FDP data and questions the admissibility of phenotypic data in court as evidence. Finally, one last ethical consideration is the access to the data created by phenotyping and its storage.

RECENT ADVANCEMENTS IN FDP:

Reliability and accuracy in estimating age, ancestry, and appearance from crime scene DNA are critical to the effectiveness of FDP in criminal casework. A prediction model for estimating attributes using epigenetic data and a validated multiplex genotyping tool for evaluating predictive DNA markers are examples of FDP technologies. Probability estimation uses an error-based model to forecast age and a likelihood ratio (LR) framework for ancestry. Improvements in FDP have increased multiplex capacity and sensitivity, especially in MPS technology. Using programs like Geno Geographer and forensic BGA tools, recent MPS-based technologies improve predictions for age, appearance, and ancestry.

VALIDATION

Established in 2017, the VISible Attributes Through GEnomics (VISAGE) Consortium the main

goal of the VISible Attributes through GEnomics (VISAGE) Consortium is to develop and confirm new, reliable molecular and statistical methods for utilizing DNA to predict appearance, ancestry, and age. The VISAGE ET A&A assay was evaluated and validated with assistance from five VISAGE Consortium laboratories. Tests were dispersed among labs to lower MPS expenses while maintaining validity. In order to enable quicker suspect identification from crime scene DNA, advancements are being made to get around forensic restrictions. The VISAGE Consortium uses enormous parallel sequencing to provide tools for predicting the age, appearance, and lineage of suspects. Through international assessments, comparisons, and conferences, the European DNA Profiling (EDNAP) Group guarantees the accuracy of forensic DNA technology.The IrisPlex System's repeatability across 21 laboratories has so far been satisfactorily certified by EDNAP. Additionally, it has evaluated the age prediction using an examination of DNA methylation[18].

Conclusion

In conclusion, FDP stands at the intersection of science, ethics, and law enforcement, offering unprecedented insights into predicting physical traits from genetic data. While it holds huge potential for aiding criminal investigations and providing valuable leads in cases where traditional methods fall short. The implications for privacy, potential misuse, and the inherent limitations of accuracy underscore the importance of careful oversight and implementation. As this continues to advance, it is imperative that we navigate its complexities with a keen awareness of human rights, ensuring that FDP contributes positively to justice while upholding ethical standards and protecting the dignity of all.

Source of Funding: Self

Conflict of interest: There are no conflict of interest.

References

1. Granja R, Machado H. Forensic DNA phenotyping and its politics of legitimation and contestation: Views of forensic geneticists in Europe. Social Studies of Science. 2020;53(6):850-68.
2. Pośpiech E, Chen Y, Kukla-Bartoszek M, Breslin K, Aliferi A, Andersen JD, et al. Towards broadening Forensic DNA Phenotyping beyond pigmentation: Improving the prediction of head hair shape from DNA. Forensic Science International: Genetics. 2018;37:241-51.
3. Fabbri M, Alfieri L, Mazdai L, Frisoni P, Gaudio RM, Neri M. Application of Forensic DNA Phenotyping for Prediction of Eye, Hair and Skin Colour in Highly Decomposed Bodies. Healthcare. 2023;11(5):647.
4. Marano LA, Fridman C. DNA phenotyping: current application in forensic science . Research and Reports in Forensic Medical Science. 2019;Volume 9:1-8.
5. White D, Rabago-Smith M. Genotype–phenotype associations and human eye color. Journal of Human Genetics. 2010;56(1):5-7.
6. Pushpesh K, Umema A, Manisha R. Forensic DNA phenotyping: A promising tool to predict human appearance for forensic purposes. GSC Biological and Pharmaceutical Sciences. 2023;23(2):034-41.
7. External visible characteristics prediction through SNPs analysis in the forensic setting: a review. Frontiers in Bioscience-Landmark. 2021;26(10):828.
8. Luquetti D, Claes P, Liberton DK, Daniels K, Rosana KM, Quillen EE, et al. Modeling 3D Facial Shape from DNA. PLoS Genetics. 2014;10(3):e1004224.
9. Alshehhi A, Almarzooqi A, Alhammadi K, Werghi N, Tay GK, Alsafar H. Advancement in Human Face Prediction Using DNA. Genes. 2023;14(1):136.
10. Li Y, Zhao W, Li D, Tao X, Xiong Z, Liu J, et al. RETRACTED ARTICLE: EDAR, LYPLAL1, PRDM16, PAX3, DKK1, TNFSF12, CACNA2D3, and SUPT3H gene variants influence facial morphology in a Eurasian population. Human Genetics. 2019;138(6):681-9.
11. Yamanoi E, Uchiyama S, Sakurada M, Ueno Y. sjTREC quantification using SYBR quantitative PCR for age estimation of bloodstains in a Japanese population. Legal Medicine. 2018;32:71-4.
12. Cho S, Ge J, Seo SB, Kim K, Lee HY, Lee SD. Age estimation via quantification of signal- joint T cell receptor excision circles in Koreans. Legal Medicine. 2014;16(3):135-8.
13. Prokunina-Olsson L, Marcińska M, Pośpiech E, Abidi S, Andersen JD, van den Berge M, et al. Evaluation of DNA Variants Associated with Androgenetic Alopecia and Their Potential to Predict Male Pattern Baldness. Plos One. 2015;10(5):e0127852.

14. Walsh S, Liu F, Ballantyne KN, van Oven M, Lao O, Kayser M. IrisPlex: A sensitive DNA tool for accurate prediction of blue and brown eye colour in the absence of ancestry information. Forensic Science International: Genetics. 2011;5(3):170-80.
15. Walsh S, Liu F, Wollstein A, Kovatsi L, Ralf A, Kosiniak-Kamysz A, et al. The HIrisPlex system for simultaneous prediction of hair and eye colour from DNA. Forensic Science International: Genetics. 2013;7(1):98-115.Walsh S, Chaitanya L, Breslin K, Muralidharan C, Bronikowska A, Pospiech E, et al. Global skin colour prediction from DNA. Human Genetics. 2017;136(7):847-63.
16. MacLean C, Lamparello A. Forensic DNA Phenotyping in Criminal Investigations and Criminal Courts: Assessing and Mitigating the Dilemmas Inherent in the Science. Recent Advances in DNA & Gene Sequences (Formerly Recent Patents on DNA & Gene Sequences). 2015;8(2):104-12.
17. Canales Serrano A. Forensic DNA phenotyping: A promising tool to aid forensic investigation. Current situation. Spanish Journal of Legal Medicine. 2020;46(4):183-90.
18. Xavier C, de la Puente M, Mosquera-Miguel A, Freire-Aradas A, Kalamara V, Ralf A, et al. Development and inter-laboratory evaluation of the VISAGE Enhanced Tool for Appearance and Ancestry inference from DNA. Forensic Sci Int Genet. 2022;61:102779.