Recent Developments in the Forensic Phenotyping of Age, Ancestry, and Appearance

Shajie Xing*

Department of Gerontology, Guangxi University, Nanning, People's Republic of China

Corresponding Author*

Shajie Xing
Department of Gerontology, Guangxi University,
Nanning, People's Republic of China,
E-mail: shajiexing@gmail.com

Received: May 03, 2023, Manuscript No. JFP-23-23873; **Editor assigned:** March 05, 2023, PreQC No. JFP-23-23873 (PQ); **Reviewed:** May 19, 2023, QC No. JFP-23-23873; **Revised:** July 03, 2023, Manuscript No. JFP-23-23873 (R); **Published:** August 10, 2023, 10.35248/2684-1312.23.8.2.355

Abstract

In order to help with the search for unidentified perpetrators who cannot be found using forensic STR profiling, Forensic DNA Phenotyping (FDP) predicts a person's externally visible characteristics, such as appearance, biogeographic ancestry, and age, from DNA of crime scene samples. All three of FDP's components have made significant recent advancements, which we summarise in this review article. The ability to predict appearance from DNA now includes qualities other than only eye, hair, and skin colour, such as eyebrow colour, freckles, hair structure, male pattern baldness, and tall stature. From determining continental ancestry to detecting sub-continental ancestry and resolving co-ancestry patterns in genetically admixed individuals, biogeographic ancestry inference from DNA has advanced. Beyond blood, various somatic tissues like saliva and bones, as well as new markers and techniques for semen, can now be used to estimate age from DNA. With targeted Massively Parallel Sequencing (MPS), technological advancements have made it possible to use forensically appropriate DNA technology with greatly increased multiplex capacity for the simultaneous analysis of hundreds of DNA predictors. Several appearance traits, multi-regional ancestry, several appearance traits combined with multi-regional ancestry, and age from various tissue types can all be predicted from crime scene DNA using forensically validated MPS based FDP tools. Despite recent developments that will probably increase the impact of FDP in criminal casework in the near future, it will take more funding, technical advancements, and intensified scientific research to move reliable appearance, ancestry, and age prediction from crime scene DNA to the level of detail and accuracy police investigators may desire.

Keywords: Forensic DNA phenotyping • Predictive DNA analysis • Appearance • Biogeographic ancestry • Age • Massively parallel sequencing

Introduction

Forensic DNA Phenotyping (FDP) is the process of determining a person's age, biogeographic ancestry, and Externally Visible Features (EVCs) from DNA extracted from human biological samples gathered at crime scenes. Because the sample donor is unknown to the investigating authorities and its STR profile is not available for comparative matching, FDP is used in criminal cases where forensic STR profiling fails to find a match [1]. By narrowing down the list of probable suspects to a small group of people who match the EVC information predicted from the crime scene DNA, FDP hopes to provide investigative leads to assist in finding unidentified criminals. FDP afterwards permits targeted police inquiry based on data collected straight from the evidence. FDP applications are always followed by forensic STR profiling for final person identification because present FDP techniques cannot give appearance on the individual specific level, which is also unlikely to become available in the near future [2]. The combination of all three FDP components is the most informative way to identify unidentified perpetrators using DNA because appearance, ancestry, and age individually describe a person's EVCs, and because some appearance traits depend on certain biogeographic ancestries and/or a certain age range. Therefore, if the legal situation permits, it is advised to combine DNA based appearance, ancestry, and age prediction in forensic practice. As previously said, FDP has moral, cultural, and legal ramifications. In recent years, a number of nations have updated their forensic DNA laws to permit and control FDP, while in some other nations it is permitted with no legal restrictions [3].

Literature Review

The degree of specificity, precision, and dependability with which appearance, ancestry, and age can be anticipated from crime scene DNA determines the FDP's ultimate practical success in criminal casework. If the culprit is local, another aspect is the frequency of the projected EVC trait in the area where the crime was committed. Predicted EVC traits that are less prevalent in the area where the crime occurred will aid the police investigation more than common ones in identifying the unknown offender (if the culprit is from that region). This does not imply, however, that forecasting common qualities is useless because it allows for the exclusion of people who do not share such traits, such as members of minority groups [4]. The ultimate application of the FDP results during the police inquiry will determine how well the FDP is implemented. Combining FDP with patrilineal familial search, where only those men are addressed for voluntarily participating in the Y-STR profiling that meet the FDP outcome is an efficient method in cases with unidentified male criminals. Compared to DNA mass screening or DNA dragnets used without combining with FDP, this combined strategy enables the emphasis on a smaller subset of volunteers who match the FDP outcome. The murder and rape case of Milica van Doorn in the Netherlands serves as an example of how well this combination strategy worked. FDP tools typically have two parts: A forensically validated multiplex genotyping tool for analysing all predictive DNA markers in the crime scene sample based on a forensically appropriate DNA technology that allows low quality and low quantity DNA analysis, and a prediction tool based on a validated prediction model for obtaining probabilities in appearance and ancestry prediction, as well as for estimating the age from the epigenetic data, established with the mu-seq method [5]. If a prediction model or tool is accurate enough for practical application, it can be determined from the average prediction accuracy estimates that are available from the validation of the prediction models. Even models with lower average accuracy, meanwhile, can nonetheless produce high probability in some people, albeit in fewer people than with accurate models. Current appearance DNA prediction techniques provide probability of trait categories for all appearance traits for which DNA predictors are present in the genotyping methods in practical casework applications. Since the science of continuous appearance prediction from genetic data is still in its infancy or is hindered by the extremely high number of DNA predictors required, appearance DNA prediction in forensic applications thus far only reflects categorical prediction. When determining biogeographic ancestry from crime scene DNA, a Likelihood Ratio (LR) framework is typically used, which mirrors

the estimation of a probability? When estimating an individual's age using DNA, the prediction model's average error is used to calculate the age estimate. The most likely category of all appearance qualities, the most likely geographic area of bio-geographic heritage, and the most likely age of the unknown sample donor are therefore provided by FDP, along with information on the flaws of these DNA based predictions. This is a benefit compared to eyewitness accounts, which are notoriously very subjective and subject to alter over time, but the mistake in an eyewitness account of any particular case is entirely unknown. Police investigators can choose what weight to give the generated FDP information in the inquiry based on the size of the probability or LRs obtained in any given case [6]. Reference data is always used, either directly or indirectly, by FDP. Directly performing inference analysis on the case sample data along with the reference population sample data, as with the majority of biogeographic ancestry prediction tools. Indirectly, as in the case of age and appearance prediction, when the probabilities are derived from the reference data by the prediction models rather than directly by the prediction tools. Therefore, when FDP results are reported to the investigating authorities, reference data should always be described along with the prediction outcome. Age prediction from crime scene DNA, appearance, and ancestry have all shown considerable improvements recently, which we summarise in this paper. We cite the prior review articles on the forensically motivated prediction of appearance and ancestry published in 2015 and age published in 2016–17 for earlier accomplishments in this field. The more DNA predictors that have become accessible in recent years have been a crucial component of the enhanced FDP systems. This quantity, however, exceeded the multiplex capacities of the forensic DNA technologies that were previously utilised in FDP instruments. Targeted MPS technologies have improved multiplex capacity as compared to all other forensic DNA methods now in use, and this, along with its sensitivity and dependability, makes MPS the primary technology for FDP applications. When compared to DNA methylation markers, which are used to predict age, SNPs, which are used to predict appearance and ancestry, have the highest multiplex capacity of targeted MPS. Significant advancements have been made in recent years regarding the integration of MPS technologies into the forensic workflow for all forensic objectives, including FDP. As we discuss in this review, a variety of MPS-based multiplex genotyping tools have recently been developed for predicting from crime scene DNA

- Several appearance traits combined.
- Multi-region biogeographic ancestry.
- Several appearance traits combined with multi-regional ancestry.
- Age from different tissues.

These tools provide improved FDP solutions However, there are important points that apply to numerous recent and earlier studies on DNA based appearance, ancestry, and age prediction used for forensic purposes. First, there is uncertainty regarding the predictive power of the markers utilised because the sample size of the dataset used to discover the predictive DNA marker was frequently tiny. Second, the datasets used to create and test the prediction model were small in scope and not totally independent from the dataset used to uncover the markers. When the same dataset is applied to the different parts of the prediction process, especially when it is applied to all three steps, the prediction accuracy may be overestimated. As long as the same (big) dataset was not also used for marker discovery, utilizing the same (large) dataset for model development and model validation using cross validation is a viable option. However, using separate datasets for all three processes is always preferable. Third, many of the reported prediction models were not made available as tools for making predictions, which prevents people besides the authors of the reports from using them in real world situations. Fourth, a dedicated multiplex genotyping tool is frequently absent from studies that discuss prediction markers and models, which makes them less useful for real-world applications. Fifth, multiplex genotyping methods frequently lacked forensic validation studies when they were announced, which prevents their use in forensic practice. The best method for creating and validating FDP tools includes the following elements:

- DNA predictors identified from a sizable dataset not used in prediction modelling
- Prediction model created in a sizable independent dataset and validated in a sizable independent dataset
- Delivering high enough prediction accuracies for appearance and ancestry and low enough error for age
- A single-multiplex genotyping method is built for all DNA predictors utilised in the prediction model (s) based on technology suited for analyzing low amount and low quality DNA, and The generated prediction model is made available for practical applications as a prediction tool.
- The prediction tool is made accessible along with the forensically verified multiplex genotyping tool, enabling practical FDP applications in forensic casework.

Discussion

The multiplex genotyping tool has undergone forensic validation and successfully passed the major validation steps. However, in practice, a lot of studies did not satisfy some or all of these criteria, which restricts the use of FDP. All of the aforementioned points were raised when designing, developing, and forensically validating the VISAGE toolbox by the Visible Attributes through GEnomic (VISAGE) consortium and project, which focused on enhancing, integrating, implementing, disseminating, and assessing the societal and ethical implications of Forensic DNA phenotyping on appearance, bio-geographic ancestry, and age. A summary of recent advancements in forensically motivated DNA prediction of appearance features is available elsewhere. At that time, several predictive DNA marker sets, multiplex genotyping technologies, some with forensic validation, and statistical prediction models, some with prediction tools, have established categorical eye and hair colour prediction from crime scene DNA. While it was becoming possible to predict skin colour from DNA, other physical characteristics were not, due to severe limitations in the genetic information that was then available for these characteristics. Since 2015, DNA has become a more reliable predictor of skin tone, and (more) predictive DNA markers for additional physical characteristics have been found, including markers for freckles, hair structure, and hair loss in men, tall stature, and grey hair. After giving a quick update on developments in eye, hair, and skin colour DNA prediction, we summarise the most recent developments in appearance DNA prediction for these freshly included features in the following sections. The number of reported associated DNA variants for other appearance traits, such as ear morphology, facial hair traits, and facial shape, has increased over the past few years, but the phenotypic variance they explain is still insufficient for the development of FDP tools at this time, which we summarise in one section. In the final half of this chapter, we cover the recently developed topic of epigenetic prediction of outwardly apparent traits or behaviours that are influenced by exposure to outside influences like tobacco use. Prior to 2015, the key steps in the construction of crime scene DNA analysis for predicting eye and hair colour were completed. Numerous validation studies of these previously validated eye and hair colour DNA prediction tools have been conducted in recent years in various populations from the same and different continents, including admixed groups, and with various statistical approaches, including machine learning methods. The realistic length of this evaluation would be exceeded by discussing these topics here. The IrisPlex model for predicting eye colour and the HIrisPlex model for predicting hair colour were both updated in 2018 by boosting the model's reference data, respectively. With nearly 9500 samples now included in the upgraded IrisPlex model, it now has prediction accuracy values for brown, blue, and intermediate eye colours of 0.95, 0.94, and 0.74, respectively. Notably, the relatively low AUC for intermediate eye colour understates the ability of IrisPlex to predict non-blue and non-brown eye colour. This is possible by deviating from the standard practise of determining the most likely eye colour from the category with the highest probability, as is done for blue and brown eye colour. IrisPlex frequently enables drawing.

Conclusion

Conclusions about intermediate eye colour from probability that is comparably high for the blue and brown eye colour groups. Based on nearly 1900 samples, the new HIrisPlex model

AUCs of 0.92, 0.83, 0.80, and 0.72 for red, black, blond, and brown hair colours, respectively. The Erasmus MC Hirisplex makes the upgraded IrisPlex and HIrisPlex models accessible to the general public for use as practical prediction tools. Based on dynamic IrisPlex and HIrisPlex prediction models, both of the prediction tools accessible *via* the HirisPlex website enable missing data depending on which SNP is absent in the partial profiles derived from low quality and/or quantity crime scene DNA. The updated IrisPlex and HIrisPlex models are implemented in non-dynamic forms in the VISAGE Software for Appearance, Ancestry and Age prediction from DNA (VISAGE software), which uses MPS data produced by the two VISAGE enhanced tools for age and the VISAGE enhanced tool for appearance as input. Expert forensic genetic practitioners can get the VISAGE Software through the European Network of Forensic Science Institutes (ENFSI).

References

1. Kayser, M., et al. "Recent advances in Forensic DNA Phenotyping of appearance, ancestry and age." *Forensic Sci Int Genet*. 65 (2023):102870.
2. Schneider, P.M., Prainsack, B., & Kayser, M. "The use of forensic DNA phenotyping in predicting appearance and biogeographic ancestry." *Dtsch Arztebl Int*. 116.51-52 (2019): 873.
3. Xavier, C., et al. "Development and validation of the visage AmpliSeq basic tool to predict appearance and ancestry from DNA." *Forensic Sci Int Genet*. 48 (2020):102336.
4. Samuel, G., & Prainsack, B. "Forensic DNA phenotyping in Europe: Views "on the ground" from those who have a professional stake in the technology." *New Genet Soc*. 38.2 (2019): 119-141.
5. Kayser, M. "Forensic DNA phenotyping: Predicting human appearance from crime scene material for investigative purposes." *Forensic Sci Int Genet*. 18 (2015):33-48.
6. Butler, J.M. "Recent advances in forensic biology and forensic DNA typing: INTERPOL review 2019-2022." *Forensic Sci Int Synergy*. 6 (2023):100311.

Cite this article: Xing S. "Recent Developments in the Forensic Phenotyping of Age, Ancestry, and Appearance". J Forensic Pathol, 2023, 8(2), 1-3.