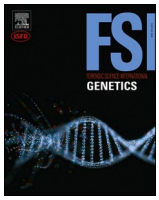
Forensic Science International: Genetics 65 (2023) 102870

Contents lists available at ScienceDirect

Forensic Science International: Genetics

journal homepage: www.elsevier.com/locate/fsigen



Recent advances in Forensic DNA Phenotyping of appearance, ancestry and age☆

Manfred Kayser [a,*], Wojciech Branicki [b,c], Walther Parson [d,e], Christopher Phillips [f]

[a] Department of Genetic Identification, Erasmus MC, University Medical Center Rotterdam, Rotterdam, the Netherlands
[b] Institute of Zoology and Biomedical Research, Jagiellonian University, Kraków, Poland,
[c] Institute of Forensic Research, Kraków, Poland
[d] Institute of Legal Medicine, Medical University of Innsbruck, Innsbruck, Austria
[e] Forensic Science Program, The Pennsylvania State University, PA, USA
[f] Forensic Genetics Unit, Institute of Forensic Sciences, University of Santiago de Compostela, Spain

ARTICLE INFO

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

ABSTRACT

Forensic DNA Phenotyping (FDP) comprises the prediction of a person's externally visible characteristics regarding appearance, biogeographic ancestry and age from DNA of crime scene samples, to provide investigative leads to help find unknown perpetrators that cannot be identified with forensic STR-profiling. In recent years, FDP has advanced considerably in all of its three components, which we summarize in this review article. Appearance prediction from DNA has broadened beyond eye, hair and skin color to additionally comprise other traits such as eyebrow color, freckles, hair structure, hair loss in men, and tall stature. Biogeographic ancestry inference from DNA has progressed from continental ancestry to sub-continental ancestry detection and the resolving of co-ancestry patterns in genetically admixed individuals. Age estimation from DNA has widened beyond blood to more somatic tissues such as saliva and bones as well as new markers and tools for semen. Technological progress has allowed forensically suitable DNA technology with largely increased multiplex capacity for the simultaneous analysis of hundreds of DNA predictors with targeted massively parallel sequencing (MPS). Forensically validated MPS-based FDP tools for predicting from crime scene DNA i) several appearance traits, ii) multi-regional ancestry, iii) several appearance traits together with multi-regional ancestry, and iv) age from different tissue types, are already available. Despite recent advances that will likely increase the impact of FDP in criminal casework in the near future, moving reliable appearance, ancestry and age prediction from crime scene DNA to the level of detail and accuracy police investigators may desire, requires further intensified scientific research together with technical developments and forensic validations as well as the necessary funding.

1. Introduction

Forensic DNA Phenotyping (FDP) refers to the prediction of a person's externally visible characteristics (EVCs) regarding appearance, biogeographic ancestry and age from DNA extracted from human biological samples collected at crime scenes [1]. FDP is applied in criminal cases where no match in forensic STR-profiling is found, because the sample donor is unknown to the investigative authorities and thus its STR-profile is unavailable for comparative matching [2]. FDP aims to provide investigative leads to help find unknown perpetrators of crime by reducing the number of potential suspects to a group of people that match the EVC information predicted from the crime scene DNA. FDP therefore allows focused police investigation based on information obtained directly from the evidential material. Because current FDP tools cannot deliver appearance on the individual-specific level, which also is unlikely to become possible in the foreseeable future, FDP applications are always followed by forensic STR-profiling for final individual identification. Because appearance, ancestry and age by themselves describe a person's EVCs, and since some appearance traits depend on certain biogeographic ancestries and/or a certain age range, the combination of all three FDP components is the most informative way to find unknown perpetrators with the help of DNA. Therefore, it is recommended to

☆ This paper is dedicated to our esteemed colleague and dear friend Peter M. Schneider, who sadly passed away in September 2022; Forensic DNA Phenotyping is one of the many forensic genetic areas he had impacted with his 35 years' work in the field.

* Corresponding author.
E-mail address: m.kayser@erasmusmc.nl (M. Kayser).

https://doi.org/10.1016/j.fsigen.2023.102870
Received 17 February 2023; Accepted 4 April 2023
Available online 6 April 2023
1872-4973/© 2023 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

combine DNA-based appearance, ancestry and age prediction in forensic practice, provided the legal situation allows [2]. FDP does not come without ethical, societal and legal implications as described elsewhere [3]. In recent years, several countries have revised their forensic DNA legislation to allow and regulate FDP, while in some other countries FDP is allowed without law specifications [2].

The final practical success of FDP in criminal casework depends on the level of detail, accuracy and reliability with which appearance, ancestry and age can be predicted from crime scene DNA. Another factor is the frequency of the predicted EVC feature in the region where the crime was committed, assuming the criminal is local. Predicted EVC features that are less frequent in the region where the crime happened will help more in the police investigation to find the unknown perpetrator (provided the perpetrator comes from that region), than common ones. However, this does not mean that predicting common features has no value because it allows the exclusion of individuals not matching these common features, for instance members of minority groups [2]. Practical success of FDP also depends on how the FDP outcome is finally utilized during the police investigation. An effective way in cases with unknown male perpetrators is to combine FDP with patrilineal familial search [4], where only those men are approached for voluntary participation in the Y-STR profiling that meet the FDP outcome. This combined approach allows to focus on a smaller number of volunteers matching the FDP outcome than would DNA mass screening or DNA dragnets without combining with FDP. The success of this combined approach is exemplified in the murder and rape case of Milica van Doorn in the Netherlands [5].

FDP tools generally consist of two components: i) a forensically validated multiplex genotyping tool for analyzing all predictive DNA markers in the crime scene sample based on a forensically suitable DNA technology allowing low quality and low quantity DNA analysis, and ii) a prediction tool based on a validated prediction model for obtaining probabilities in appearance and ancestry prediction, and for estimating the age from the epigenetic data, established with the multiplex genotyping tool from the crime scene DNA. Average prediction accuracy estimates available from the validation of the prediction models indicate if a prediction model / tool is accurate enough for practical application. However, even models with lower average accuracy can yield high probabilities in certain individuals, albeit in less individuals than with accurate models. In practical casework application, current appearance DNA prediction tools deliver probabilities of trait categories for all appearance traits for which DNA predictors are included in the genotyping tools. Thus far, appearance DNA prediction in forensic applications purely reflects categorical prediction as the science of continuous appearance prediction from genetic data is not yet advanced enough or troubled by the very large number of DNA predictors needed. The estimation of a probability is mirrored in the inference of biogeographic ancestry from crime scene DNA, where typically a likelihood ratio (LR) framework is applied. In DNA-based age prediction, the prediction model delivers an age estimate for which the error comes as average error of the prediction model. Thus, FDP not only provides information on the unknown sample donor's most likely category of all appearance traits, the most likely geographic region of bio-geographic ancestry, and the most like age, but also on the errors of these DNA-based predictions. This marks an advantage over eyewitness descriptions (when these are available) that are known to be highly subjective and prone to change over time [6], but the error in an eye witness report of any specific case is completely unknown. Based on the magnitude of the probabilities or LRs obtained in any one case, police investigators can decide what weight to give the generated FDP information in the investigation. FDP always relies on reference data used either directly or indirectly. Directly, as with most biogeographic ancestry prediction tools performing inference analysis of the case sample data alongside reference population sample data. Indirectly, as with appearance and age prediction, where the reference data are used in the prediction models to obtain the probabilities, but not by the prediction tools directly. Therefore, reference data should always be described together with the prediction outcome, when FDP results are reported to the investigative authorities.

In recent years there have been significant advances in FDP in all three components: appearance, ancestry and age prediction from crime scene DNA, which we summarize in this review. For earlier achievements in this field, we refer to the previous review articles on the forensically-motivated prediction of appearance and ancestry published in 2015 [7,8] and age published 2016–17 [9,10]. A key element for the improved FDP solutions was the increased number of DNA predictors that have become available over recent years. However, this number went beyond the limits of the multiplex capacities of the forensic DNA technologies previously used in FDP tools. Advances in targeted MPS technologies demonstrated increased multiplex capacity compared to all previously used forensic DNA technologies, which together with its sensitivity and reliability makes MPS the key technology for FDP purposes. The multiplex capacity of targeted MPS is highest for SNPs, used for appearance and ancestry prediction, compared to DNA methylation markers, used for age prediction. In the last years, significant progress has been made in the implementation of MPS technologies in the forensic workflow for all types of forensic purposes including for FDP [11]. Various MPS-based multiplex genotyping tools have been developed recently for predicting from crime scene DNA i) several appearance traits combined, ii) multi-region biogeographic ancestry, iii) several appearance traits combined with multi-regional ancestry, and iv) age from different tissues, which provide improved FDP solutions, as we discuss in this review.

However, there are critical points that apply to many of the recent and earlier studies on forensically motivated DNA-based appearance, ancestry and age prediction. First, the sample size of the dataset used for discovering the predictive DNA marker was often small, which leads to uncertainty in the predictive value of the markers used. Second, datasets used for building and validating the prediction model were also small in scale and not independent from each other and/or the marker discovery dataset. Applying the same datasets for the different steps in the prediction modelling can result in overestimation of the prediction accuracy, especially when the same dataset is used for all three steps. Using the same (large) dataset for model building and model validation by performing cross-validation is a valid option as long as it such dataset was not additional used for marker discovery, but using independent datasets for all three steps is always better. Third, many of the reported prediction models were not made available as prediction tools, which prevents others, besides the reporting authors, to apply them in practice. Fourth, many articles that describe prediction markers and models do not provide a dedicated multiplex genotyping tool, which hinders practical applications. Fifth, when multiplex genotyping tools were reported, they often lacked forensic validation studies, which prevents applications in forensic practice.An optimum approach for developing and validating FDP tools includes: i) DNA predictors ascertained from a large dataset not used in prediction modeling, ii) a prediction model built in a large independent dataset and validated in a large independent dataset, while the model validation delivers high-enough prediction accuracies for appearance and ancestry and low-enough error for age, iii) the developed prediction model is made available for practical applications as a prediction tool, iv) a single-multiplex genotyping tool is developed for all DNA predictors used in the prediction model(s) based on a technology suitable for analyzing low quantity and low quality DNA, v) the multiplex genotyping tool has undergone forensic validation and successfully passed the major validation steps, and vi) the forensically validated multiplex genotyping tool is made available together with the prediction tool, thereby allowing practical FDP applications in forensic casework. However, in reality, many studies did not meet several or all of these points, which limits FDP applications. The VISible Attributes through GEnomic (VISAGE) Consortium and Project, which worked on improving, integrating, implementing, disseminating and assessing the societal and ethical implications of

Forensic DNA Phenotyping on appearance, bio-geographic ancestry and age (http://www.visage-h2020.eu), considered all the above raised points when designing, developing, and forensically validating the VISAGE Toolbox.

2. Recent progress in predicting appearance traits from crime scene DNA

Developments in forensically motivated DNA prediction of appearance traits published up to 2014–15 are summarized elsewhere [7]. At that time, categorical eye and hair color prediction from crime scene DNA had become established with several predictive DNA marker sets, multiplex genotyping tools - some with forensic validation, and statistical prediction models – some with prediction tools [7]. Skin color DNA prediction was emerging, while other appearance traits were not predictable from DNA due to strong limitations in the genetic knowledge available for these traits at that time [7]. Since 2015, DNA prediction of skin color became more established [12,13], and (more) predictive DNA markers for more appearance traits were discovered, such as for eyebrow color, freckles, hair structure, hair loss in men, tall stature, and grey hair. In the following sections, we summarize the recent advances in appearance DNA prediction for these recently added traits after we provide a brief update on advances in eye, hair, and skin color DNA prediction. For other appearance traits, such as ear morphology, facial hair traits and facial shape, the number of reported associated DNA variants has increased over the last years, but the phenotypic variance they explain is not large enough for developing FDP tools as of yet, which we summarize in one section. In the last section of this chapter, we discuss the newly emerging field of epigenetic prediction of externally visible characteristics or habits that are determined by the exposure to external factors such as tobacco smoking.

2.1. Recent advances in eye color and hair color DNA prediction

The main steps in the establishment of eye and hair color DNA prediction from crime scene DNA were accomplished prior to 2015 as described before [7]. The recent years have seen many validation studies of these previously established eye and hair color DNA prediction tools in different populations from the same and different continents including admixed groups, and with different statistical approaches including machine-learning methods. Discussing these here would go beyond the practical length of this review.

In 2018, the IrisPlex model for eye color prediction and the HIrisPlex model for hair color prediction were both revised by increasing the model-underlying reference data, respectively [13]. The updated IrisPlex model is now based on close to 9500 samples and yields prediction accuracies expressed as cross-validated AUC (area under the receiver operating curve) of 0.95, 0.94 and 0.74 for brown, blue, and intermediate eye color, respectively (Table 1). Notably, the relatively low AUC for intermediate eye color understates the ability to predict non-blue and non-brown eye color with IrisPlex, which is possible by deviating from concluding the most likely eye color from the category with the highest probability, as typically done for blue and brown eye color. In many cases, IrisPlex allows concluding intermediate eye color from similarly high probabilities obtained for the blue and the brown eye color categories. The updated HIrisPlex model is based on close to 1900 samples and gives cross-validated AUCs of 0.92, 0.83, 0.80 and 0.72 for red, black, blond and brown hair color, respectively (Table 1). The updated IrisPlex and HIrisPlex models are publicly available for practical use as prediction tools via the Erasmus MC Hirisplex website at https://hirisplex.erasmusmc.nl. Both prediction tools available via the HirisPlex website are based on a dynamic IrisPlex and HIrisPlex prediction models, thereby allowing missing data depending on which SNP is missing in the incomplete profiles obtained from low quality and/or quantity crime scene DNA. Non-dynamic versions of the updated IrisPlex and HIrisPlex models are implemented in the VISAGE Software for Appearance, Ancestry and Age prediction from DNA (VISAGE Software, Table 1), which uses as input the MPS data generated with the VISAGE Enhanced Tool for Appearance and Ancestry [14] and the two VISAGE Enhanced Tools for Age [15,16]. The VISAGE Software is available for expert forensic genetic practitioners via the European Network of Forensic Science Institutes (ENFSI).

The IrisPlex and HIrisPlex SNPs can be analyzed with the forensically validated IrisPlex and HIrisPlex multiplex genotyping tools based on a single SNaPshot assay, respectively [17–19]. They are included together with autosomal ancestry informative SNPs and Y-chromosome SNPs in a hybridization capture based MPS assay [20] (Table 1). They are part of the forensically validated commercial MPS-based ForenSeq DNA Signature Prep Kit (Verogen) together with genetic markers for biogeographic ancestry and other forensic purposes [21] (Table 1). The commercial Universal Analysis Software (Verogen) for data analysis allows, amongst other things, to generate eye and hair color probabilities from the data established with the kit. Notably, the prediction parameters used by this commercial prediction tool are the ones from the first IrisPlex and HIrisPlex models, which were based on much smaller datasets and deliver lower accuracies then the updated current models available via the HirisPlex webtool (Table 1). The Iris and HIrisPlex SNPs can also be analyzed together with additional skin color predicting SNPs of the HIrisPlex-S system with several MPS tools (see below under skin color).

The last years have also seen significant progress in genetically understanding eye and hair color variation more completely via outcomes of genome-wide association studies (GWASs) with enlarged sample size and consequent increased statistical power. Published in 2018 [22], the International Visible Trait Genetics (VisiGen) Consortium performed a GWAS on hair color in almost 300,000 Europeans and identified 124 significantly associated genetic loci, of which 111 were not previously known for hair color. Next to GWAS, the authors tested the predictive value of the newly identified SNPs in a total of $< 15{,}000$ Europeans from two cohorts. Based on 252 of the 258 independently associated SNPs discovered in their GWAS together with 18 available HIrisPlex SNPs, and based on data from two cohorts combined (split 80 %:20 % for model building and testing), AUCs of 0.86 for red, 0.86 for black, 0.74 for blond and 0.68 for brown were achieved, while the incomplete 18-SNP HIrisPlex model gave 0.85, 0.78, 0.67 and 0.62, respectively, in the same data [22]. Thus, a prediction accuracy increase was achieved for all four hair color categories, which, however, was relatively small for the price of adding 234 (14-fold) more SNPs to the model, with AUC increases of 0.08 for black, 0.07 for blond, 0.06 for brown, and 0.01 for red hair [22]. A dedicated multiplex genotyping tool for the SNP predictors and a prediction tool were not made available by the authors, who used GWAS data from SNP microarrays. Provided enough multiplexing capacity of the genotyping method, the hair color SNP predictor identified by Hysi et al. should be considered in the development of future FDP tools.

In 2021, the VisiGen Consortium published a GWAS on eye color in 195,000 Europeans and identified 61 significantly associated genetic loci of which 50 were previously unknown for eye color [23]. Although the eye color predictive value of the identified SNPs was not tested in this study, the authors quantified the amount of eye color variation explained by these SNPs, which was 53 % (95 % confidence interval 45–61 %) of the total eye color variation in their study population. Also published in 2021, Kukla-Bartoszek et al. [24] performed a whole exome sequencing study in a (for association studies very small) number of 150 Polish subjects and reported 27 new candidate SNPs for eye color. Testing 137 newly and previously discovered eye color SNPs in 849 Polish samples enabled the authors to develop new predictive models for eye color following a two-step modelling approach. In the first step, AIC, BIC and LASSO methods were applied for marker selection, and in the second step, regression models were built based on the selected markers. The regression model based on LASSO-selected markers consisted of 10 SNPs and that based on BIC-selected markers only had 4 SNPs; both

Table 1
MPS-based tools for predicting appearance traits from crime scene DNA.

Tool name	Appearance Traits	Appearance Markers	Composition, MPS technology	Forensic validation*	Prediction models with reference data and prediction accuracy**	Prediction tool	References
ForenSeq DNA Signature Prep Kit (Verogen) #	Eye color, hair color	24 HIrisPlex SNPs	Appearance + BGA + other forensic purposes, ForenSeq	Yes	Eye color: first IrisPlex model: model building N = 3804, model validation N = 2364, AUCs blue 0.91, brown 0.93, interm. 0.72; Hair color: first HIrisPlex model: model building and cross-validation N = 385, AUCs red 0.9, black 0.78, blond 0.75, brown 0.72	Universal Analysis software, UAS (Verogen)#	[21,214,215]
Bulbul & Filoglu	Eye color, hair color, skin color	41 HIrisPlex-S SNPs	Appearance + BGA, AmpliSeq	Yes	Eye color: updated IrisPlex model; Hair color: updated HIrisPlex model; Skin color: HIrisPlex-S models (all see above)	HirisPlex webtool https://hirisplex.erasmusmc.nl/	[13,27]
HIrisPlex-S	Eye color, hair color, skin color	41 HIrisPlex-S SNPs	Appearance; AmpliSeq, Illumina	Yes	Eye color: updated IrisPlex model: N = 9466 cross-validation, AUCs: blue 0.94, brown 0.95, intermediate 0.74; Hair color: updated HIrisPlex model: model building and cross-validation N = 1878, AUCs red 0.92, black 0.83, blond 0.80, brown 0.72; Skin color: IrisPlex-S model: N = 1423, cross-validation: AUCs very light 0.74, light 0.72, interm. 0.73, dark 0.88, dark to black 0.96	Hirisplex webtool https://hirisplex.erasmusmc.nl/	[13,26]
VISAGE Basic Tool for Appearance and Ancestry[$]	Eye color, hair color, skin color	41 HIrisPlex-S SNPs	Appearance + BGA, AmpliSeq, ForenSeq, PowerSeq	Yes	Eye color: updated IrisPlex model; Hair color: updated HIrisPlex model; Skin color: HIrisPlex-S models (all see above)	Hirisplex webtool https://hirisplex.erasmusmc.nl/	[13,28–30]
Forensic Capture Enrichment (FORCE)	Eye color, hair color, skin color	41 HIrisPlex-S SNPs	Appearance + BGA + other forensic purposes, capture enrichment MPS, Illumina Next Seq,[$, $$]	No	Eye color: updated IrisPlex model; Hair color: updated HIrisPlex model; Skin color: HIrisPlex-S models (all see above)	HirisPlex webtool https://hirisplex.erasmusmc.nl/	[13,31]
Rauf et al.	Eye color, hair color	23 HIrisPlex SNPs	Appearance + BGA, capture enrichment MPS, Illumina	No	Eye color: updated IrisPlex model; Hair color: updated HIrisPlex model; Skin color: HIrisPlex-S models (all see above)	HirisPlex webtool https://hirisplex.erasmusmc.nl/	[13,20]
VISAGE Enhanced Tool for Appearance and Ancestry	Eye color, hair color, skin color, eyebrow color, freckles, hair shape, male hair loss	199 appearance SNPs	Appearance + BGA, AmpliSeq	Yes	Eye color: updated IrisPlex (see above); Hair color: updated HIrisPlex model, partial model with 20-SNPs, N = 1878, cross-validation, AUCs: red 0.91, black 0.83, blond 0.80, brown 0.72; Skin color: HIrisPlex-S model, partial model with 34 SNP, N = 1423, cross-validation, AUCs very light 0.74, light 0.72, interm. 0.73, dark 0.89, dark to black 0.96; Eyebrow color: Peng model (Peng et al.), partial model with 24 SNPs + sex, N = 3114 model building, N = 779 model validation, AUCs blond 0.7, black 0.68, brown 0.62; Freckles: Kukla-Bartoszek model (Kukla-Bartoszek et al.), 10 SNPs + MC1R r/R + SNP-SNP interactions + sex, N = 960, cross-validation, AUC yes-no 0.75; Hair shape: Pospiech model (Pospiech et al.), partial model with 28 SNPs + sex, N = 9674 model building, N = 2138 model validation, AUC straight / non-straight 0.68 (all), 0.66 (Europeans), 0.81 (non-	VISAGE Software	[14,35,37,42,46]

(*continued on next page*)

Table 1 (*continued*)

Tool name	Appearance Traits	Appearance Markers	Composition, MPS technology	Forensic validation*	Prediction models with reference data and prediction accuracy**	Prediction tool	References
					Europeans); Male hair loss: Chen model (Chen et al.), partial model with 102 SNPs, N = 104,694 model building, N = 26,177 model validation, AUC no / any 0.69.		

MPS-tools are listed in time-wise order of publication, * in the sense of published forensic validation studies of the MPS tool including testing of sensitivity, specificity, degradation, concordance, population samples, mock casework samples; ** prediction accuracy expressed as AUC where 1.0 means perfect prediction and 0.5 random prediction, for additional prediction accuracy parameters, see original references, also see Caliebe et al. [216] for discussion on prediction accuracy parameters, # commercial tool, $ AmpliSeq primer pool commercially available from Thermo Fisher Scientific, $$ QIAseq primer pool commercially available from QIAGEN.

models included 3 of the 6 IrisPlex SNPs. These two best models achieved similarly high AUCs for blue and brown eye color as available with the IrisPlex markers and model, while for intermediate eye color the AUC was significantly higher with AUC of 0.85 [24]. No multiplex genotyping tool and no prediction tool were made available in this study. Provided enough multiplexing capacity of the genotyping method, the newly identified eye color SNPs should be considered in the development of future FDP tools. Eventually, DNA-based eye and hair color prediction should be moved from the categorical to the continuous level in the future.

2.2. Skin color DNA prediction

The first comprehensive system for skin color prediction DNA became available in 2014 from Maronas et al. [25] as previously discussed together with earlier developments [7]. In 2018, Chaitanya et al. [13] published an extension of the HIrisPlex system by additionally including skin color predicting SNPs previously identified by Walsh et al. [12]. This HIrisPlex-S (S for skin) system includes 41 SNPs of which 36 are predictive for skin color, some of which were already included in HIrisPlex. The predictive value for skin color of the 36 SNPs included in the HIrisPlex-S system was identified by Walsh et al. [12] in a comprehensive study testing 77 pigmentation SNPs from 37 genetic loci in 2025 individuals from 31 global populations. By applying both marker sets in the same 194 DNA samples, Wlash et al. demonstrated that these 36 SNPs outperformed the 10 SNPs from the skin color prediction system of Maronas et al. [25].

The current HIrisPlex-S model for skin color prediction consists of genotype and phenotype data of 1423 globally distributed individuals [13] and achieves prediction accuracies expressed as cross-validated AUC of 0.96 for dark to black, 0.88 for dark, 0.73 for intermediate, 0.72 for light, and 0.74 for very light skin color (Table 1). The dynamic HIrisPlex-S model is publicly available for practical use as prediction tool via the Erasmus MC Hirisplex website at https://hirisplex.erasmusmc.nl together with the IrisPlex model for eye color and HIrisPlex model for hair color prediction. A non-dynamic version of the HIrisPlex-S model is implemented in the VISAGE Software (Table 1).

The HIrisPlex-S SNPs can be analyzed via two forensically validated SNaPshot multiplex assays [13] or via targeted MPS. Targeted MPS assays for HIrisPlex-S are available as i) one multiplex based on AmpliSeq and Illumina sequencing with the HIrisPlex-S MPS tool [26] (Table 1), ii) one multiplex in combination with ancestry-informative SNPs based on AmpliSeq [27] (Table 1), iii) one multiplex in combination with ancestry-informative SNPs with the VISAGE-BT-AA based on AmpliSeq [28], ForenSeq [29], and PowerSeq [30], for which the AmpliSeq primer pool is commercially available as a community panel from Thermo Fisher Scientific, iv) one multiplex in combination with SNPs for biogeographic ancestry and other forensic purposes based on hybridization capture enrichment with the Forensic Capture Enrichment (FORCE) tool [31], and v) one multiplex in combination with SNPs for four additional appearance traits and biogeographic ancestry with the VISAGE Enhanced Tool for Appearance and Ancestry (VISAGE-ET-AA) based on AmpliSeq technology [14] (Table 1). All these MPS tools, except FORCE, have been involved in forensic validation studies, albeit not to the same degree (Table 1).

Genetic knowledge on skin color further improved in the recent years. In 2018, Visconti et al. [32] published a GWAS on sun sensitivity as a proxy phenotype for skin color in over 175,000 Europeans and identified 20 significantly associated genetic loci of which 14 were previously unknown for skin color. Skin color GWASs were also performed recently in non-Europeans such as in Africans [33] and South Asians [34] that re-discovered genetic skin color loci previously known from Europeans as well as genetic loci previously unknown from Europeans, which, however, still require independent replication. These newly identified skin color associated SNPs have not been tested yet for their skin color predictive value, which should be done in the future. However, given that prediction accuracies available with current systems, such as HIrisPlex-S, are considerably lower for light skin categories than those of the dark categories, finding more predictive SNPs for light skin is more essential to improve skin color DNA prediction, than adding more dark skin SNP predictors. Eventually, as with eye and hair color, also skin color DNA prediction should also be moved from the categorical to the continuous level in the future.

2.3. Eyebrow color DNA prediction

In 2019, Peng et al. [35], with support by the VISAGE Project, published the first GWAS on eyebrow color, a trait which is largely but not completely correlated with head hair color (e.g., some blond individuals have dark eyebrows). They used >8500 European samples and identified six significantly associated genetic loci, one of which had not been previously linked with eyebrow color or other human pigmentation traits. In addition to GWAS, the authors used all identified significantly associated SNPs for prediction modelling of eyebrow color by building the prediction model in >3000 Europeans and validating it in >750 subjects. Their best model based on 25 SNPs achieved AUCs of 0.7, 0.67 and 0.62 for blond, black and brown eyebrow color, respectively (red eyebrow color was excluded due to the extremely small sample size in the study). Dedicated multiplex genotyping and prediction tools for eyebrow color were not made available by the authors, who used GWAS data from SNP microarrays. However, the eyebrow color predictive SNPs reported by Peng et al. are included in the forensically validated VISAGE-ET-AA MPS tool (Table 1), and the eyebrow color prediction model is implemented in the VISAGE Software (Table 1). Given the medium level accuracy achieved by the current model, future efforts should focus on identifying additional independently predictive SNPs for eyebrow color and consider them in developing future FDP tools.

2.4. Freckle DNA prediction

In 2018, Hernando et al. [36] reported the first genetic prediction model for freckles, for which the same dataset of 458 Spanish subjects

were used for prediction model building and model validation. Their final prediction model was based on three SNPs from three genes plus the compound R/r markers from *MC1R*. The model achieved a cross-validated AUC of 0.77 for presence versus (vs.) absence of freckles without considering sex, and 0.78 with sex. Independent model testing in an additional > 190 samples gave an AUC of 0.81. Dedicated multiplex genotyping and prediction tools for freckles were not made available by the authors. In 2019, Kukla-Bartoszek et al. [37], with support by the VISAGE Project, published the second prediction model for freckles. In their study, the authors first screened 113 DNA variants from 46 genes previously associated with pigmentation traits in 960 Polish samples to identify genetic freckle predictors. They used the same dataset for prediction model building and model validation. Their 2-category model for presence vs. absence of freckles based on 12 variables achieved a cross-validated AUC of 0.75. Their 3-category model based on 14 variables revealed a cross-validated AUC of 0.79 for heavy freckling, 0.66 for medium freckling, and 0.75 for absence of freckles. As variables in their prediction models, the authors considered SNPs, compound *MC1R* R/r markers, sex, SNP-SNP interactions, and sex-SNP interactions. This model achieved a prediction accuracy increase of 0.085 AUC compared to the previous model from Hernando et al. Dedicated multiplex genotyping and prediction tools for freckles were not made available by the authors. However, the freckle predictive SNPs reported by Kukla-Bartoszek et al. for their 2-category model are included in the forensically validated VISAGE-ET-AA MPS tool (Table 1), and the freckles prediction model is implemented in the VISAGE Software (Table 1). Based on the medium-level AUC achieved by the current models, future efforts should focus on identifying additional independently predictive SNPs for freckles and consider them in future FDP tools.

2.5. Grey hair DNA prediction

In 2016, a GWAS on hair phenotypes including head hair greying was published [38] based on over 6000 Latin Americans, which discovered one significantly associated genetic locus for grey hair harboring *IRF4*, a known pigmentation gene. In 2020, Pospiech et al. [39] published an association study based on whole-exome sequencing data in a (for association testing very small) number of 150 Polish samples, followed by targeted MPS of 378 newly identified exonic and literature-based SNPs in over 849 Polish subjects. The authors used the same dataset for prediction model building and model validation. Their 2-category prediction model for presence vs. absence of hair greying based on ten SNPs, age and sex achieved a cross-validated AUC of 0.87, while their 3-category model based on twelve SNPs, age, and sex cross-validated AUCs of 0.86 for no, 0.79 for mild, and 0.88 for severe hair greying. However, the authors reported that their SNP predictors explained only 10 % of the hair greying variation in their study population, while age explained 48 % and sex > 5 %. Thus, age alone was responsible for most of the prediction accuracy these models achieved. Dedicated multiplex genotyping and prediction tools for grey hair were not made available by the authors. Future efforts will need to concentrate on finding more independently predicting SNPs for hair greying to develop more accurate prediction models and tools. Moreover, because of its strong age dependency, hair greying should additionally be investigated via epigenome-wide association studies (EWAS) to find DNA methylation sites associated with grey hair that may serve as epigenetic predictors for grey hair and consider them in future epigenetic FDP tools.

2.6. Hair shape DNA prediction

In 2015, Pospiech et al. [40], as part of the EUROFORGEN-NoE Consortium, reported the first prediction model for hair shape based on three SNPs in 528 Polish samples using the same dataset for prediction model building and model validation. Based on different methods used, their models achieved cross-validated AUCs between 0.589 and 0.688 for straight vs. non-straight hair. Dedicated multiplex genotyping and prediction tools for hair shape were not made available by the authors. In 2018, Liu et al. [41] published a GWAS on hair shape in almost 29,000 individuals from different continental populations, which identified 12 significantly associated genetic loci, of which 8 were not previously involved in hair shape. The authors reported a prediction model for hair shape based on 14 SNPs and sex, which was built from over 6000 Europeans that were part of the discover GWAS and achieved a cross-validated AUC of 0.66, while the external validation in almost 1000 independent Europeans gave 0.64. Dedicated multiplex genotyping and prediction tools for hair shape were not made available by the authors, who used GWAS data from SNP microarrays. As part of the EUROFORGEN-NoE Consortium, Pospiech et al. [42] published in 2018 what is currently the most comprehensive prediction model for hair shape. For model building, the authors used data from over 9600 European and non-European subjects previously used by Liu et al. [41] and tested 90 candidate SNPs. Model validation employed an independent dataset of nearly 2500 European and non-European samples. The best 2-category prediction model for straight vs. non-straight based on 32 SNPs, sex and age achieved an AUC of 0.7 for Europeans and non-Europeans combined. A considerably higher AUC of 0.80 was obtained from non-Europeans (N = 277) compared to 0.68 Europeans (N = 2138). For combined Europeans and non-Europeans, the best 3-category model based on 33 SNPs, mostly overlapping with those in the 2-category model, yielded AUCs of 0.68, 0.6 and 0.62 for straight, wavy and curly hair shape, respectively, without sex and age. For non-Europeans and Europeans separately, the AUCs were 0.8, 0.61, 0.74 and 0.67, 0.6, 0.6, respectively. The increased prediction accuracy in non-Europeans is explained by a strong SNP predictor in the *EDAR* gene, for which the predictive allele does not exist in Europeans [41]. Dedicated multiplex genotyping and prediction tools for hair shape were not made available by the authors. However, the SNP predictors of the 2-category model of hair shape prediction by Pospiech et al. are included in the forensically validated VISAGE-ET-AA MPS tool (Table 1), and the hair shape prediction model is implemented in the VISAGE Software (Table 1). The medium-level AUC currently achieved for hair shape prediction should prompt future efforts to identify additional independently predictive SNPs and consider them in the development of future FDP tools.

2.7. Male hair loss DNA prediction

In 2015/16, the first two genetic prediction models for hair loss in men, or male pattern baldness (MPB), were independently published. Marcinska et al. [43], as part of the EUROFORGEN-NoE Consortium, tested 50 MPB-associated SNPs for their predictive value in > 600 Europeans from different European populations. The authors reported a 5-SNP model and an extended 20-SNP model, both built in 305 samples and validated in independent 300 samples. Their 20-SNP 2-category model achieved an AUC of 0.66 for no baldness vs. significant baldness without considering age, while when selecting males over 50 years of age, the AUC increased to 0.76. Dedicated multiplex genotyping and prediction tools for MPB were not made available by the authors. In a study carried out in parallel, Liu et al. [44] tested 25 SNPs previously associated with MPB for their predictive value in 2455 samples from three European populations. Prediction modelling was done separately in three cohorts and yielded overlapping sets of 6–14 SNP predictors and age as predictor. The same data were used for model building and validation. In an early-onset enriched MPB dataset (N = 727), the best 2-category model based on 14 SNPs achieved a cross-validated AUC of 0.741 for baldness vs. no baldness. In a population-based dataset (N = 1161), the best 2-category model based on 11 SNPs plus age achieved a cross-validated AUC of 0.711 while in an independent smaller population-based dataset (N = 567) the AUC was lower with 0.685 based on a smaller number of 6 SNPs plus age. Dedicated multiplex genotyping and prediction tools for MPB were not made available by the

authors, who used GWAS data from SNP microarrays. In 2017, Hagenaars et al. [45] reported a MPB prediction model based on 331 SNPs they identified via GWAS in 40,000 samples from the UK Biobank (UKBB) study they also used for model building, which was validated in > 12,000 independent UKBB males. The authors reported AUCs of 0.78 for no MPB vs. severe MPB, 0.68 for no MPB vs. moderate MPB, and 0.61 for no MPB vs. slight MPB without considering age, and 0.79, 0.70, and 0.61, respectively, when considering age. Dedicated multiplex genotyping and prediction tools for MPB were not made available by the authors, who used GWAS data from SNP microarrays.

In 2022, Chen et al. [46] with support by the VISAGE Project, published a genetic prediction model for MPB that currently represents the most data-supported model available for the genetic prediction of MPB or any other appearance trait, because they used large and independent datasets for all different analytical steps. Based on the associated SNPs reported by Hagenaars et al. [45], the authors identified 117 SNP predictors in over 55,500 UKBB males that largely overlapped with those previously used by Hagenaars et al. for discovering the MPB-association of these SNPs. Based on these 117 SNPs from 85 genetic loci, they build MPB prediction models with different methods in over 100,000 independent UKBB males and validated them in over 26,000 independent UKBB males. The reported AUC were similar across methods in the range of 0.725–0.728 for severe, 0.631–0.635 for moderate, 0.598–0.602 for slight, and 0.708–0.711 for no hair loss with age, and slightly lower without. Two-category prediction of any versus no hair loss gave AUCs of 0.690–0.711 with age and slightly lower without. Additional external validation in an early-onset enriched MPB dataset (N = 991) showed improved prediction accuracy without considering age such as AUC of 0.830 for no vs. any hair loss. Dedicated multiplex genotyping and prediction tools for MPB were not made available by the authors, who used GWAS data from SNP microarrays. However, the MPB-predictive SNPs identified by Chen et al. are included in the forensically validated VISAGE-ET-AA MPS tool (Table 1), and the male hair loss prediction model is implemented in the VISAGE Software (Table 1). A recent GWAS on MPB [47] in over 200,000 UKBB males identified 624 significantly associated genetic loci, which in the future should be tested for their predictive value, provided the necessary independent dataset for prediction model building and validation become available.

2.8. Body height DNA prediction

In 2014, Liu et al. [48] used 180 previously height-associated SNPs in a total of > 10,300 Europeans enriched with 770 very tall individuals for building and validating a prediction model for tall stature, which achieved a cross-validated AUC of 0.75 for tall vs. non-tall stature. A dedicated multiplex genotyping tool for the SNP predictors and a prediction tool were not made available by the authors, who used GWAS data from SNP microarrays. In 2019, the same group [49], with support by the VISAGE Project, published an update on the genetic predictability of tall stature by using the same cohort for testing the predictive value of 697 height-associated SNPs that were identified in a previous GWAS on over 250,000 subjects published in 2014 [50]. Based on 689 available SNPs, the model achieved a cross-validated AUC of 0.79 for tall vs. non-tall prediction, representing an AUC increase of 0.04 for the price of 509 (3.8-fold) more SNPs. The authors also demonstrated that the most informative subset of 412 SNPs achieved an AUC of 0.76 i.e., 0.01 AUC increase for 232 (2.3-fold) more SNPs. Of note, these two models for genetic prediction of tall stature have almost no value for predicting continuous height as indicated by the obtained correlations between genetically predicted height and observed height of $R^2 = 0.12$ for the 180-SNP model and 0.21 for the 689-SNP model [49]. Provided enough multiplexing capacity of the applied genotyping method, tall stature prediction should be included in future FDP tools.

Such studies indicate the enormous difficulty in predicting normal height from DNA, which is caused by the large genetic complexity of height together with the minimal effect size in the millimeter and submillimeter range that associated SNPs have on body height. This problem can only be overcome by including a very large number of SNPs in the prediction model i.e., via genomic prediction. A rare example for genomic appearance prediction is the study by Lello et al. published in 2018 [51], which presented prediction of continuous body height using > 630,000 SNPs in > 460,000 UKBB samples. Based on 5000 samples not previously used for prediction model building and with their best 100,000 SNP predictors, the authors reported a correlation between predicted height and observed height of $R^2 < 0.7$, where a subset of 20,000 SNPs were described as optimal height predictors. Dedicated multiplex genotyping and prediction tools for height were not made available by the authors, who used GWAS data from SNP microarrays. In 2022, Yengo et al. [52] published what currently is the largest GWAS on body height in 5.4 million individuals, which revealed over 12,000 significantly associated SNPs from over 7200 genetic loci accounting for 40 % of height variations in their European study population.

Reliable genotyping of many thousands of SNPs from low quality and quantity DNA typically obtained from crime scene samples with currently available targeted MPS technologies and chemistries is expected to be challenging. Targeted MPS involving hybridization capture enrichment allows to drastically increase the number of simultaneously analyzable SNPs and is already used since many years in the field of ancient DNA for hundreds of thousands of SNPs. Since recently, capture-based targeted MPS is started to be used for thousands of SNPs for forensic purposes [53]. In 2021, Tillmar et al. [31] published the Forensic Capture Enrichment (FORCE) panel involving several thousands of SNPs for different forensic purposes including appearance (pigmentation) and ancestry prediction.

2.9. Recent genetic progress on appearance traits not yet applicable for DNA prediction

Since 2015, several GWASs were published that have improved and broadened our knowledge on the genetic basis of human appearance also regarding additional traits not discussed above, such as on ear morphology, facial hair traits, and facial shape. However, the number and effect sizes of the SNPs associated with these traits are too small so that they explain too little of the phenotypic variance; hence, these traits are not applicable for FDP purposes yet. Therefore, a clear need exists for more GWASs based on larger sample sizes to identify more small-effect SNPs that serve as DNA predictors, which may eventually provide practically useful prediction accuracies and if so, include them in future FDP tools in case the multiplex capacity of the genotyping method allows.

In 2015, Adhikari et al. [54] published a GWAS of ear morphology in >5000 Latin Americans and reported seven significantly associated genetic loci for different ear phenotypes including earlobe attachment and others. In 2017, a large GWAS on earlobe attachment in nearly 75,000 individuals identified 49 significantly associated genetic loci for whether the earlobe is free hanging, partially, or fully attached [55]. A large GWAS on multiple ear morphology traits is currently underway by a group including one of the authors. In 2016, Adhikari et al. published a GWAS on scalp and facial hair features in the same Latin American cohort and found significant associations with scalp hair shape and balding, hair greying (see above), mono-eyebrow, beard thickness, hair color and eyebrow thickness at 18 genetic loci, of which 10 were novel findings [38]. In 2018, Wu et al. [56] reported two additional significantly associated genetic loci involved in eyebrow thickness identified via GWAS in < 3000 Chinese. A GWAS on eyebrow thickness in Europeans is currently underway by a group including one of the authors. In 2022, Pospiech et al. [57] published a candidate SNP association study and a prediction study on several hair-related phenotypes considering 240 SNPs in 999 Polish samples, demonstrating clear evidence of pleiotropy and epistasis in the genetics of hair traits. The reported prediction models achieved low to medium cross-validated AUCs for hairiness in females (0.69–76) and males (0.51–0.59), mono-eyebrow (0.62–0.70),

eyebrow thickness in males (0.5–0.63), head hair thickness (0.6–0.63), and head hair density (0.56–0.65).

Recent progress has also been made in increasing the genetic knowledge underlying phenotypic variation of facial shape, building on the first two facial GWASs published before 2015 [58,59] summarized elsewhere [7]. In 2016, three GWASs on facial shape were published. Adhikari et al. [60] identified four significantly associated genetic loci mostly for the nose in >6200 Latin American subjects. Cole et al. [61] identified two significantly associated genetic loci with measures of facial size in >3500 African children that were replicated in <2400 African children. Shaffer et al. [62] reported seven significantly associated genetic loci with different facial shape measurements in >3000 European subjects. In 2018, Claes et al. [63] published a GWAS on facial shape in >2300 European subjects that identified 38 significantly associated genetic loci of which 15 were reported to replicate in an independent European sample of >1700; four of the reported loci were novel. In a follow-up study published in 2021 [64], this group applied their phenotyping approach to an enlarged sample set of >8200 European subjects and reported a large number of 203 significantly associated genetic loci of which 53 were located in regions not previously known to be involved in facial development or diseases with facial manifestation. In a parallel study published in 2019, Xiong et al. [65], on behalf of the VisiGen Consortium, performed a GWAS on 78 facial shape phenotypes, obtained from 13 facial landmarks placed mostly automatically on the 3-dimensional digital facial images using dedicated computer vision methods, in > 10,100 European subjects with replication in an additional >7900 Europeans and non-Europeans. The authors identified 24 significantly associated genetic loci, of which 17 were novel. In 2021, Bonfante et al. [66] published a face GWAS in more than 6000 Latin Americans and reported significant associations at 32 genetic loci of which 9 were previously unidentified. In 2022, Zhang et al. [67] published a face GWAS in Chinese based on nearly 7000 samples for discovery and over 2700 for replication, which revealed 166 significantly associated genetic loci, of which 62 were not previously involved in facial variation. End of 2022, Xiong et al. [68] published a new method for combining GWASs of multiple traits (C-GWAS) and presented its first application to facial shape, which identified 56 significantly associated genetic loci, of which 17 were not involved in facial variation before.

In their 2019 paper, Xiong et al. [65] reported on the quantification of the facial phenotypic variance genetically explained by the SNPs they identified in their GWAS. A multiple-regression analysis conditioning on the effects of the lead SNPs from 24 genetic loci identified 31 SNPs with significant independent effects on sex- and age-adjusted facial distance phenotypes, which per each SNP explained less than 1 % of the phenotypic facial variation and all together 4.62 %. In their 2022 paper, Xiong et al. [68] performed a polygenic risk score analysis based on their C-GWAS findings using 57 significantly face-associated SNPs, which explained on average 2.28 % and up to 4.51 % sex- and age-adjusted facial variance. These very small proportions of facial variance explained by face-associated SNPs illustrate the major problem of moving from face GWASs identifying facial SNPs to predicting faces with SNPs. Therefore, scientific publications that claim to predict human faces from genetic data [69–71], and companies that provide commercial service testing on predicting faces from crime scene DNA, must be viewed very critically [72,73].

2.10. DNA prediction of externally visible characteristics induced by exposure to external factors

A recent newcomer in the field of (extended) FDP is the epigenetic prediction of externally visible characteristics (or habits) that are determined by individual interaction with external factors, such as tobacco smoke, which is summarized elsewhere [74]. Its combination with genetic prediction of appearance and ancestry (see chapter 2) and the epigenetic prediction of age (see chapter 3) is expected to characterize the externals of an unknown person from DNA in a more complete way. Several recent epigenome-wide association studies (EWASs) revealed DNA methylation sites showing significant association with (non)consumption of substances such as tobacco, alcohol, coffee, tea, cocaine, heroin etc. The first epigenetic prediction models for smoking and drinking were already reported of which we mention here the two most recent ones. In 2019, Maas et al. [75] published an epigenetic prediction model for smoking habits based on 13 DNA methylation sites that was built and internally validated on 3764 samples and delivered cross-validated AUCs of 0.925 for current smokers, 0.83 for never smokers, and 0.766 for former smokers, while external validation in 1608 independent samples gave 0.914, 0.781, and 0.699, respectively. In 2021, Maas et al. [76] reported epigenetic prediction models for alcohol consumption based on different sets of DNA methylation markers that were built in 2883 samples and for heavy and at-risk drinkers vs. light and non-drinkers delivered mean (across marker sets) cross-validated AUCs of 0.67–0.68, and 0.6–0.7 in the external validation in 1794 independent samples. Continues progress in understanding the impact of environmental exposure on the human epigenome will likely lead to increasingly accurate epigenetic prediction models for the named and other environmentally determined externally visible habits.

3. Recent progress in inferring biogeographic ancestry from crime scene DNA

In this chapter, we review recent advances in biogeographic ancestry (BGA) inference from forensic DNA published since the 2015 review of this topic [8]. We will describe the key elements of recently reported forensic BGA tools, which comprise marker selection, genotyping multiplex design, and statistical analysis of the resulting data. Ancestry-informative DNA marker (AIM) selection must compile a suitable panel for a defined set of population differentiations. The consequent statistical regime applied to the genotype data must be able to predict BGA using reference population datasets, but also ideally having the capacity to detect co-ancestry in individuals with admixed backgrounds. Because the resolution of BGA inferences from DNA largely depends on the number of AIMs used, and targeted MPS has the largest multiplex capacity of all DNA technologies currently available for forensic DNA analysis, we concentrate here on forensic BGA tools based exclusively on targeted MPS. Recent efforts in developing such tools have concentrated on autosomal SNPs as the ancestry-informative markers (AIMs) of choice, but there is increasing interest in the ancestry informativeness of autosomal microhaplotypes (MHs) – combinations of closely sited SNPs in short sequences which are readily detected with the single-strand sequencing of MPS. Insertion/deletion polymorphisms (Indels) for forensic BGA tools were reviewed previously [8] and have not developed further during the last ten years. Recent progress has also been made in targeted MPS tools for hundreds of Y-chromosome SNPs, many of which are ancestry-informative for a male's paternal lineage [77], and for complete mitochondrial genome analysis that provides the maximal level of maternal BGA information [78]. Since we focus here on bi-parental BGA inferred with autosomal AIMs, we will only mention such markers if they are part of MPS tools that focus on autosomal AIMs. Autosomal STRs used in forensic DNA profiling for individual genetic identification can give viable population differentiations especially when the entire sequence information is available, as with targeted MPS (Fig. 5 of [79]. However, this form of variation has less power than autosomal AIM SNPs, so STR tests have not been adapted specifically for BGA and therefore are not discussed further (but see section 6.2 of [8]).

3.1. AIM SNP panels in forensic MPS tools for BGA

Several MPS tools have been developed since 2015 that are either dedicated to forensic BGA analysis or include DNA markers for BGA combined with those for other forensic applications. Each of the MPS

tools discussed below is summarized in Table 2.

3.1.1. AIMs in commercial MPS-based forensic BGA tools

The compact set of 56 AIMs developed by the Kidd lab at Yale University were described in [1] and have been the subject of several studies since [80,81], but this panel is only a portion of two larger scale SNP sets for MPS analysis: i) MPSplex (developed jointly by QIAGEN and ICMP, but not yet commercially available) combining more than 1400 SNPs and MHs for individual identification, but adding the 56 AIMs to provide BGA inference for missing persons identification [82]; and ii) Verogen's ForenSeq DNA Signature Prep Kit (FDSPK) comprising SNPs and STRs for individual identification in 'primer pool A′ with a multiplex extension possible using 'primer pool B′ which adds eye and hair color predictive SNPs and AIM SNPs [83,84]. As well as being smaller in scale than other MPS-based AIM panels, the 56 Kiddlab AIM-SNPs in FDSPK do not benefit from a comprehensive statistical system, as the universal analysis software of FDSPK (UAS) just runs a simplified principal component analysis (PCA) based on limited 1000 Genomes Phase-I reference populations (now superseded by more comprehensive 1000 Genomes Phase-III datasets, but not used by UAS), and distance-to-closest-centroid calculations (see Section 5 and Fig. 6 of [8]). Therefore, users make BGA assessments based on scrutinizing the position of the forensic sample in relation to reference clusters in the PCA plot, but without accompanying statistical output. To compensate for a lack of likelihood-ratio (LR) analysis in UAS, more complete reference population data from 1000 Genomes Phase-III and CEPH human genome diversity panel (HGDP-CEPH) samples are available in Snipper for the same AIM SNPs (http://mathgene.usc.es/snipper/forensic_mps_aims.html). Snipper analysis provides similar 2D PCA but with more extensive reference data than used by UAS, plus coupled LR calculations, both of which can analyze customizable combinations of reference populations. Adapting reference data in this way can fine-tune the populations compared to evaluate the forensic sample with more limited ranges of ancestries (e.g., comparing African, Europe, Middle East and

Table 2
MPS-based tools for inferring biogeographic ancestry from crime scene DNA.

Name	BGA resolution	Ancestry Informative Markers	Composition, MPS Technology	Forensic validation	Prediction Approach and Prediction Tool	Original Publication	Validation Publication
EUROFORGEN gAIMs	5-groups*	127 autosomal SNPs	BGA, AmpliSeq	Yes	Snipper-based LR tests and PCA (http://mathgene.usc.es/snipper/index.php), STRUCTURE analysis (https://web.stanford.edu/group/pritchardlab/structure.html)	[87]	[217]
ForenSeq DNA Signature Prep Kit (Verogen)#	5-groups*	56 autosomal SNPs	BGA + Appearance + other forensic purposes, ForenSeq (MiSeq)	Yes	PCA with centroids marked generated in Verogen Universal Analysis Software (UAS)	-	[21]
Precision ID Ancestry Panel (Thermo Fisher Scientific)#	5-groups*	165 autosomal SNPs	BGA, AmpliSeq	Yes	Distribution of simulated likelihoods using inferred admixture ratios generated in TFS Converge software	-	[125]
Bulbul & Filoglu	6 groups: 5-groups*+ Middle East / SW-Asia	117 autosomal SNPs	BGA + Appearance, AmpliSeq	Yes	Snipper and/or FROG-kb (https://frog.med.yale.edu/FrogKB/), STRUCTURE analysis	[27]	[27]
MAPlex	7 groups: 5-groups*+ S-Asia + Middle East / SW-Asia	144 autosomal SNPs + 20 MHs	BGA, AmpliSeq	Yes	Snipper-based LR tests and PCA, STRUCTURE analysis	[88]	[218]
NAME	1 group: North African-Middle East	111 autosomal SNPs	Complementary BGA, MiSeq, AmpliSeq	Yes	Snipper-based LR tests and PCA, STRUCTURE analysis	[89]	[89]
MPSplex	6 groups: 5 groups*+ S-Asia	1270 tri-allelic autosomal SNPs +46 MHs**	BGA + individual ID, QIAseq, MiSeq	Yes	STRUCTURE analysis	[82]	[82]
PhenoTrivium	7 groups: 5 groups*+ S-Asia + Middle East / SW-Asia	163 autosomal SNPs +120 Y-SNPs	BGA + Appearance, AmpliSeq	Yes	Distribution of simulated likelihoods using inferred admixture ratios generated in TFS Converge software	[90]	[90]
VISAGE Basic Tool for Appearance and Ancestry$	6-groups: 5 groups*+ S-Asia	115 autosomal SNPs	BGA + Appearance, AmpliSeq, ForenSeq, PowerSeq	Yes	Snipper-based LR tests and PCA, STRUCTURE analysis	[91]	[28–30]
Forensic Capture Enrichment (FORCE)	5 groups* with South Asians instead of Oceanians	241 autosomal SNP, 829 Y-SNPs	BGA + Appearance + other forensic purposes, capture enrichment MPS, Illumina NextSeq,$, $$	No	PCA and Naïve Bayes LR test, Yleaf	[31]	[31]
Rauf et al.	5 groups*	67 autosomal SNPs, 35 Y-SNPs	BGA + Appearance, capture enrichment MPS, Illumina	No	Snipper-based LR tests and PCA	[20]	-
VISAGE Enhanced Tool for Appearance and Ancestry	7 groups: 5 groups*+ S-Asia + Middle East	104 autosomal SNPs +87 Y-SNPs +16 X-SNPs +123 MH SNPs	BGA + Appearance, AmpliSeq	Yes	VISAGE Software	[92]	[14]

MPS-tools are listed in time-wise order of publication, # Commercial tool, which lack original scientific article describing it, $ AmpliSeq primer pool commercially available from Thermo Fisher Scientific, $$ QIAseq primer pool commercially available from QIAGEN, MHs: autosomal microhaplotypes consisting of several SNPs in close physical proximity. * 5-groups: distinguishing Sub-Saharan African, European, East Asian, Oceanian and Native American population groups. * * Markers not selected for ancestry but for individual identification purpose; because of their number they also provide BGA information.

South Asian data alone). Snipper also handles missing genotypes effectively by highlighting the relative informativeness of those AIMs dropping out, by listing the SNPs in order of informativeness and marking absent genotypes in red. FROG-kb, the Kidd lab's own population database for SNPs useful in the forensic setting including for BGA, also offers comprehensive sets of population data and the means to perform likelihood-based ancestry calculations on multiple SNP profiles [85]. Both MPSplex and FDSPK MPS assays constitute forensic ancestry tools but use a relatively small portion of the total genotypes that the test generates and for detailed statistical analysis the user must bring in complementary analysis systems in Snipper or FROG-kb to infer BGA.

Of the 56 AIM SNPs described above, 55 were combined with 128 from Seldin's non-forensic panel (the Kiddlab SNP rs1919550 is redundant as it is completely linked to rs12498138 selected independently by Seldin), creating the Thermo Fisher Scientific (TFS) Precision ID Ancestry Panel (PIAP), which represents the first commercial MPS test dedicated wholly to BGA inference [86]. With 37 SNPs common to both panels, 165 autosomal AIM SNPs are genotyped in a single MPS assay. TFS provide an ancestry analysis plug-in for use with PIAP and forms part of the Converge software suite which performs LR tests using population data to generate a ranked list of likelihoods for individual populations and/or continental regions. Likelihood comparisons also enable an estimate of co-ancestry proportions which are used to evaluate individual ratios of admixture components using simulation-based modelling (described below). Limitations of this commercial BGA tool include: i) a requirement to link to the allele frequencies held in FROG-kb via the ALFRED database to make the likelihood calculations, although these are quite extensive in geographic scope, this limits a user's flexibility to apply specific population comparisons; and, ii) an emphasis on sub-population comparisons, which do not always accurately reflect a forensic DNA donor's ancestry (e.g., likelihood ratios less than 100 that suggest 'more likely Japanese than Chinese Han') which give the impression that the test can reliably make such distinctions. The reporting forensic scientist can moderate this statistical output to reduce the risk of overstated accuracy but may also be tempted to provide the investigators with ancestry information that has too high a classification error rate when the compared sub-populations are closely related and/or lack geographic separation.

3.1.2. AIMs in non-commercial MPS-based forensic BGA tools

Dedicated MPS-based forensic AIM panels which are not currently commercially developed are, in order of publication: the 126-SNP EUROFORGEN Global AIMs (gAIMs) [87], a 154-SNP combined BGA-pigmentation phenotype predictive panel [27], the 164-marker MAPlex panel [88], the 111-SNP EUROFORGEN NAME panel [89], the PhenoTrivium combined ancestry-appearance panel [90], the 153-SNP VISAGE-BT-AA tool [28–30,91], and the 524-SNP VISAGE-ET-AA tool (counting all SNPs for appearance and ancestry) [14,92]. Available statistical data analysis tools such as Snipper and GenoGeographer (see below) can be applied to data generated by all these MPS tools. Snipper provides LR and PCA analysis frameworks, extendable to STRUCTURE-based genetic cluster analysis using the same population data input. Applying LR, PCA and STRUCTURE analyses simultaneously and to the same chosen reference population datasets is advocated as this enables complementary approaches to ancestry inference and analysis of admixture in individuals with co-ancestry. The VISAGE-ET-AA panel has a more closely integrated statistical analysis package designed to combine the age and appearance predictions with those for BGA in a single workflow, the VISAGE Software.

The combined BGA-pigmentation phenotyping MPS test from Bulbul and Filoglu [27] combined two ancestry panels – 55 of the Kiddlab AIMs with 65 additional 'SWA' AIMs previously developed to differentiate populations from Southwest Asia (here considered equivalent to the Middle East region) [93] with the HIrisPlex-S SNPs for eye, hair and skin color prediction [13]. The authors tested the panel for MPS performance gauging sequence coverage, genotype concordance and ability to analyze mixtures. The panel provided improved ancestry predictive performance when differentiating Southwest Asia (Middle East), European and South Asian populations compared with using 55 AIMs alone.

The AIM selection for MAPlex focused on variation in Asia-Pacific populations, while preserving the differentiation power across all main global population groups obtained with the markers of gAIMs. It also extended gAIMs' geographic scope by adding South Asia as a targeted sub-continental population group. The original remit of MAPlex was to enhance the sub-continental differentiation within East Asia in addition to neighboring Oceania and South Asia regions. Although MAPlex did this successfully [88], the ability to differentiate well-defined populations in all these regions has been the main goal of studies assembling smaller dedicated panels [93–95] exemplified by the Japaneseplex panel [96]. Although Japaneseplex is small in scale and lacks a specific genotyping assay design, this study demonstrated the potential for compiling AIMs dedicated to differentiating one population from others in the same continental group. Such custom assays can potentially be applied in a 'nested' approach, i.e., where an initial continental-wide inference is refined by follow-up tests targeting a specific region or population.

The EUROFORGEN NAME panel is a stand-alone MPS tool developed to enhance the gAIMs panel for improved differentiation of European individuals from those of North African, Middle East, and South Asian populations [89]. Therefore, it extends the existing gAIMs ancestry test which did not originally aim to differentiate these neighboring populations from Europe. Although the success NAME obtained in differentiating North African/Middle East populations from other closely related populations was limited, when gAIMs and NAME markers are combined, Middle East populations, but not North Africans, were distinguishable from Europeans and South Asians (see Fig. 4 of [88]). Studies to select the best AIMs for specific population differentiations, were until recently, hindered by being restricted to a pool of 650,000 candidate SNPs from the Stanford analyses using 650 K microarray SNP genotyping of the HGDP-CEPH diversity panel (detailed in [8]). As discussed below, 929 unrelated HGDP-CEPH panel samples have now undergone whole-genome-sequencing, which identifies several million SNPs and Indels per individual [97], making it more straightforward in the future to identify and compile AIMs informative for HGDP-CEPH populations additional to those of 1000 Genomes (namely, Native American, Oceanian, Middle East population groups). One other benefit of the NAME panel initiative was to find the most informative and best performing loci, in MPS sequencing terms. Consequently, 29 SNPs from the NAME panel were adopted for the VISAGE-ET-AA tool [92], representing almost 28 % of the total autosomal AIMs assembled in that assay.

The PhenoTrivium panel, developed by Diepenbroek et al. in 2020 [90], combines 163 of the 165 PIAP ancestry SNPs and the 41 HIrisPlex-S eye, hair, skin color SNPs into a single AmpliSeq-based panel, but importantly, also includes 120 lineage-specific Y-SNPs. The paper reporting the development of PhenoTrivium also evaluated forensic performance by applying sequence thresholds of a minimum 100 reads (50 for Y-SNPs); and 95 %/5 % homozygote or 65 %/35 % heterozygote allele read frequency balance. The forensic sensitivity of PhenoTrivium was assessed with dilution series DNA and post-mortem blood and bone samples. Full concordant SNP profiles were obtained down to 125 pg input DNA, with just SNP rs1470608 failing to reach sequence coverage thresholds. The downstream ancestry analyses using the Converge software (likelihood calculations and bootstrapping runs with the admixture analysis algorithm) of samples with known admixture patterns (self-declared parental and grandparental combinations), representing a comprehensive evaluation of the analysis of co-ancestry with this software. Converge-based bootstrapping co-ancestry analyses were also compared to those of Snipper and FROG-kb, indicating Converge handles admixture component prediction well.

The AIM panel of the VISAGE-BT-AA tool [91] was designed to match the MPS multiplex scales of the gAIMs, MPSplex and PIAP panels to ensure good sequencing performance with forensic DNA, but more

importantly, to begin the process of combining FDP and ancestry SNPs in one MPS genotyping assay. The combination of 115 AIMs with the 41 pigmentation predictive SNPs of HIrisPlex-S (with SNPs rs16891982; rs1426654; rs12913832 shared for BGA and pigmentation prediction purposes) – means the VISAGE-BT-AA tool has the smallest number of AIMs of any forensic MPS ancestry test. As it was developed to differentiate South Asians in addition to the five main continental population groups, only the most differentiating markers were retained. A small proportion of previously used AIMs were excluded to preserve space for optimum South Asian-informative SNPs taken from the previous Eurasiaplex [98] and NAME [89] panel development work. Therefore, the final AIM panel of VISAGE-BT-AA tool consisted of 57 gAIMs, 7 Kiddlab 56 SNPs, 12 NAME SNPs, 5 SNPs from PIAP not in the 56 Kiddlab panel, and 19 South Asian informative SNPs from Eurasiaplex [91]. As well as 3 tri-allelic SNPs from gAIMs, a further 12 tri-allelic SNPs were added to the VISAGE-BT-AA AIM panel to provide scope for mixed DNA analysis, and to improve the differentiation of six of the seven target population groups [91]. The careful selection process for all the above AIMs is underlined by the ancestry prediction success of VISAGE-BT-AA, where cross-validation of the reference population samples gives 100 % classification success for all groups apart from Middle East, which has 80.6 % success, with the biggest proportion of error (17.2 %) coming from samples misclassified as European [91]. Selecting AIMs for the VISAGE-BT-AA that are not only the most ancestry informative but have proven performance in MPS has benefited the sensitivity and reliability of the VISAGE-BT-AA tool based on different MPS chemistries and instruments [28–30]. For instance, Xavier et al. [28] reported for VISAGE-BT-AA tool 100 % genotype call rates with 0.1 ng input DNA and in DNA with up to 240 min of sonication degradation, with good MPS performance across the whole range of SNPs in the test. Although the primer pool for the VISAGE-BT-AA is available as a community panel from TFS, the entire VISAGE-BT-AA MPS tool is not commercially available as of yet.

Given the MPS performance success of the VISAGE-BT-AA tool, the fact that multiplex space could be shared between AIMs and phenotype predictive SNPs without reducing ancestry informativeness, and because a total of 153 SNPs is likely to be well below the multiplex limits of targeted MPS while preserving the necessary sensitivity and reliability, for the VISAGE- ET-AA tool it was decided to increase the multiplex size almost three-fold relative to VISAGE-BT-AA. Although the VISAGE-ET-AA tool [14] has a markedly increased number of appearance predictive SNPs for 7 traits (see chapter 1), it also has a more complex combination of AIM SNPs from autosomes, X-chromosome, Y- chromosome, and autosomal microhaplotypes in order to improve analysis of co-ancestry patterns in individuals with admixed backgrounds. Despite a reduced total number of autosomal AIM SNPs (115 in VISAGE-BT-AA to 104 in VISAGE-ET-AA), the augmented autosomal SNP panel of VISAGE-ET-AA tool more than doubles the Middle East informative markers [92] originally chosen for the VISAGE-BT-AA tool (12−28). Consequently, it more efficiently distinguishes Middle East populations from those of Europe, Africa and South Asia, with the benefit that this differentiation power extends to the majority of samples from North African and East African populations when such comparisons are made. The VISAGE-ET-AA tool additionally includes 87 Y-SNPs and 16 X-SNPs to provide a distinct method for obtaining extra detail about co-ancestry patterns identified in males with admixed backgrounds. To test how effectively Y- and X-SNPs can do this, a supplementary analysis system was developed for the X-SNP data and tested on genotypes compiled from six admixed populations of 1000 Genomes, plus a cohort of urban and rural Brazilian samples sequenced with the VISAGE-ET-AA tool [92]. Y-SNP data from the male Brazilian test samples were analyzed using haplotype designations based on a core 859 Y-SNP dataset defining 640 haplogroups [77]. In all eight admixed populations, X-SNP data was used to evaluate the possible ancestry of each male sample's X-chromosome, by co-analysing these samples in PCA with African, European and Native American male reference population data for the 16 X-SNPs only. Clearly separated PCA clusters were formed by the reference samples from each of the admixture contributor populations, and when the test points were positioned in these reference clusters, the X chromosome ancestry was inferred to be the same. The four 1000 Genomes admixed American populations gave varied patterns ranging from 10 % to 50 % unassigned (i.e., positioned between clusters), but the two 1000 Genomes African American populations gave more clearly lineated patterns with 64 % of X-chromosomes African in ASW, 87.5 % in ACB; 20 % European in ASW, 8 % in ACB; and only 16 % unassigned in ASW, 5 % in ACB. The two Brazilian samples from urban and rural regions had the benefit of Y-SNP pattern comparisons (1000 Genomes Y data was incomplete), which allowed analysis of sex biased admixture patterns that contrasted between the two samples. Urban Brazilians had a high proportion of European X-chromosomes (62 %) and few African (19 %) whereas rural Brazilians had the inverse pattern: 17 % European X-chromosomes and 61 % African. Applying an independent X-SNP and Y-SNP ancestry inference regime in parallel to that of autosomal SNP genotype analysis allows a degree of extra detail to be obtained for individuals in which co-ancestry has been detected. A detailed description of the AIM set of the VISAGE ET-AA is currently underway [92].

The gAIMs, MAPlex, VISAGE-BT-AA tool and VISAGE-ET-AA tool AIM panels were all built on the principle of balancing cumulative population-specific Divergence values (termed $I_{n\ \mathbf{POP}}$ in [87] across all the population differentiations the panels were designed to make (Fig. 4 in [87]; Fig. 4 in [88]). This principle was extensively explored in the 2015 review [8] and such AIM selection and balancing steps applied to the forensic AIM panels were based on methods developed for the LACE ancestry panel used in genomics studies [99]. The benefits from balancing an AIM set in this way include i) more efficient analysis of co-ancestry in admixed individuals; ii) equilibrated likelihood statistics in most populations reducing the bias towards particular populations; and iii) more robustness to the statistical effects of missing genotypes. As more worldwide population groups are differentiated, the balancing process becomes more difficult to achieve and it will not be possible to ensure balanced I_n for comparisons of more closely related groups, e.g., Europeans vs South Asians vs Middle East populations. Nevertheless, balancing $I_{n\ \mathbf{POP}}$ values for the five major continental population groups of Africans; Europeans; East Asians; Americans and Oceanians represents a key step in the assembly of panels differentiating these groups.

To evaluate the benefit of attempting to balance I_n as much as possible, direct comparisons of co-ancestry patterns in samples from the six admixed 1000 Genomes populations, using STRUCTURE analysis, were made for both VISAGE Appearance and Ancestry tools. Inferred co-ancestry proportions in these samples were compared between those estimated using the Affymetrix Human Origins array (comprising >572,000 SNPs) and the VISAGE-BT-AA tool, with sample-to-sample correlations giving r^2 values above 0.8 in all populations and co-ancestries, apart from African/American co-ancestry in Puerto Ricans (Fig. 4 of [28]). The same analyses were made for the VISAGE-ET-AA tool [92], but the study compared the co-ancestry patterns to those from the full genome-wide SNP datasets obtained directly from 1000 Genomes (comprising several million SNPs, parsed to remove <0.05 minor allele frequency variants) [100]. The VISAGE-ET-AA tool correlation analyses showed very similar values to those obtained with the VISAGE-BT-AA tool – an important finding, given the 26 % reduction in the autosomal AIM SNPs in VISAGE-ET (from 103 in BT to 76 markers in ET, excluding those extra SNPs added to ET to differentiate Middle East populations). Additional independent evaluations of the ability of forensic MPS ancestry panels to differentiate populations in comparison to much larger SNP sets were made by Resutik et al. [101] in a study that analyzed a larger dataset of CEPH, SGDP and EGDP genotypes combined with those of 1000 Genomes, for the SNPs of VISAGE-BT-AA tool, MAPlex, and PIAP, compared to a 100,000-SNP dataset from the Affymetrix Human Origins array. Computing the G′ similarity score and area under the precision-recall curve (*AUC-PR*) to assess similarities in STRUCTURE cluster memberships between the forensic AIMs panels and

the 100,000 SNP set showed all three forensic panels had similar performance, with VISAGE-BT-AA tool giving a marginally better match to the large-scale SNP set STRUCTURE patterns at inferred cluster values above six.

3.1.3. Autosomal microhaplotypes as emerging forensic AIMs

MAPlex was the first forensic assay to combine binary SNPs, multiple-allele SNPs i.e., tri-allelic SNPs catalogued in [82]; tetra-allelic SNPs in [102] and microhaplotypes (MH) based on closely spaced autosomal SNPs [103]. A large proportion of autosomal MHs show sufficient population differentiation to be viable AIMs in their own right. The use of MHs in forensic BGA assays was reviewed elsewhere [104] and the population differentiation capabilities of a 65-MH panel were explored in [105] (a subset of 130 Kidd lab MH loci with high I_n values, but lacking MPS designs). The study by Chen et al. [106] indicated African-European vs East Asian differentiation can be achieved with only ten MHs. Therefore, given their widely recorded capacity to efficiently detect mixed-source DNA, MHs justify inclusion in MPS-based BGA panels and tools. MAPlex [88] incorporated 22 MHs; 13 shortened versions of loci from 130 MHs the Kiddlab identified [103], while the VISAGE-ET-AA tool includes 21 ancestry-informative MHs [14].

Although not all autosomal MHs make suitable AIMs, the ability to detect mixed-source DNA more efficiently than binary SNPs justifies their inclusion in MPS panels for forensic use. The same rationale applies to multiple-allele SNPs, although a detailed study of relative I_n values (amongst four divergence metrics) comparing MHs, binary and tri-allelic AIM SNPs [107] indicated that binary SNPs were more informative than tri-allelic SNPs marker-for-marker. When MPS panels of autosomal MHs are very large, collective ancestry informativeness can be very high, even if loci are primarily selected for individual identification (i.e., with low population divergence of haplotype frequencies in each MH marker), due to the summarized effect of many markers. For instance, an MPS panel designed for individual identification based on 113 MHs [108] had sufficient ancestry informativeness to allow differentiation of Middle East, North African, European, and South Asian populations with population subset STRUCTURE analyses in follow-up studies [109].

Initial explorations of the haplotype frequencies of the 113 autosomal MHs developed by de la Puente et al. [108] demonstrated the potential of these markers to infer the BGA of contributors in simple 2-way mixed-source DNA [109]. This potential to move beyond identification of simple-mixture DNA contributors from their haplotype patterns, to additionally inferring their BGA, was studied in a limited but promising pilot study of the 21 MH loci in the VISAGE-ET-AA tool [92]. Simple mixtures of two control DNAs with African and European ancestries were made at 1:1, 3:1 and 9:1 ratios, and the sequence patterns from just the 21 MHs were collected and analyzed to attempt de-convolution from contrasting sequences and read coverage levels [92]. The contrasts in haplotype frequencies in Africans and Europeans in the MHs meant that 11/21 loci showed three-haplotype combinations and 5/21 four-haplotypes, which in the 3:1 and 9:1 ratios were sufficiently distinct in sequence reads to be assignable to the major and minor contributors. When the identified haplotypes assigned to each contributor were run in STRUCTURE (which easily handles multi-allele markers), applying 1000 Genomes African, European and East Asian reference population data for the 21 MHs, BGA was successfully inferred in the 3:1 and 9:1 mixtures, where sequence ratios were sufficiently contrasted. Although a limited pilot study and based on deconvolution intended to be restricted to the three main population groups of African, European, and East Asian individuals only, these findings indicate that MHs can be successfully used to add inferred BGA to the information obtained from sequence analysis of simple 2-way mixed DNA [92]. It is also a reasonably secure inference to expect the majority of the 21 MH loci used to have multiple-haplotype patterns. Using known haplotype frequencies, simulated haplotype patterns that could be expected from mixed DNA can establish the likelihood of being able to de-convolute mixtures of individuals from the three population groups, but this pilot focused on the ability to interpret the sequence output of the 21 MH markers in VISAGE-ET-AA, which simulations are unable to model.

3.2. Expansion of population reference data for forensic BGA inference

Population frequency data on the AIM SNPs used in forensic BGA tools are required for two purposes, i) as a marker selection dataset for the initial identification and validation of the AIM SNPs by selecting SNPs with large allele frequency differences between different worldwide populations, and ii) as a reference dataset for obtaining the final BGA inference outcome for the DNA sample in question based on the results of the BGA genotyping tool applied to the case sample. In 2016, Soundararajan et al. [110] argued for establishing a common set of forensic AIMs that could be agreed upon and developed within a collaborative framework. The main motivation for suggesting this initiative was the poor geographic coverage of population diversity sample sets used as reference data for BGA inference such as the HGDP-CEPH panel (the most widely used sample set, covered in [8]) and, using the study authors' description, a largely 'empty matrix' of non-overlapping AIMs selections amongst 21 forensic panels. Agreement on a universal set of AIMs for forensic BGA DNA testing would bring the benefits of broadened regional coverage by sampling many geographic gaps (e.g., Remote Oceanian, Native American, etc.) and more detailed population inferences. Although this initiative did not progress, it is beneficial to exclude redundant, i.e., non-overlapping AIMs, but also SNPs that duplicate divergence within single genomic regions. It is noteworthy that Kidd lab's 56-SNP panel has four SNP pairs closely sited in genes: *EDAR*, *HCLS1-GOLGB1*, *ADH1B* and *ALDH2* (rs3827760-rs260690; rs1919550-rs12498138; rs1229984-rs3811801; rs2238151-rs671, respectively); despite LR calculations for this panel in FROG-kb and Snipper assuming allelic independence. Some regional variation remains uncharted and not properly represented. For instance, much of Remote Oceania is known to have quite different allelic diversity patterns to Near Oceania, due to differences in human population history of both regions, but Remote Oceania lacks population data completely, while for Near Oceania it is very sparse. Native Americans from North, Middle and South America are similarly under-represented in the available population data, as well as populations from Middle East regions. Another major factor is that AIMs highly informative for a particular population comparison may not amplify efficiently in an MPS multiplex assay or show complex flanking region sequences hampering reliable alignment. Therefore, the forensic community remains reliant on an accumulating bank of genome-wide variant data from a large number of worldwide populations.

The advantage of public genomic data based on whole-genome-sequencing data is that they establish catalogs of *all* SNPs detected, making selection of future AIMs open and flexible. For some time, SNPs considered as AIMs have centered on 1000 Genomes Phase-III data (summarized in Fig. 2 of [8]), which were generated with low coverage (2–3X) whole-genome-sequencing studies [100,111]. In recent years, the total number of compiled SNP catalogs (VCF files) from public WGS data has increased markedly. The Simons Foundation Genome Diversity Project (SGDP [112] and the Estonian Genome Diversity Project (EGDP [113]) both address the under-sampling of certain regions but rely on a strategy of 2–3 samples per population. Although such data can provide a regional/global overview, they are not useful as population data because of insufficient SNP allele frequency estimation (both projects' data scope is reviewed in [114]. Hence, such sources are not useful as population reference data for (forensic) BGA DNA testing and are also of limited value for selecting AIM SNPs in the first case. In contrast, the most extensive human variant catalog assembled to date, gnomAD (genome aggregation database) [115] has large sample sizes of several thousand individuals, but lacks detailed population definitions with loosely defined population identifiers such as Hispanic, non-Finnish European and 'Other', so is of limited use as a reference population database. Unlike 1000 Genomes, the parallel gnomAD project database

holds summary allele frequencies in each population, so it is impractical to use this data for reference purposes with current forensic ancestry inference statistical analyses that rely on individual SNP genotype data, unless SNP allele frequencies can be placed directly into LR calculations (possible in Snipper at: http://mathgene.usc.es/snipper/frequencies_new.html).

Two recent initiatives at 1000 Genomes Project have been particularly useful here. First, the completion of whole-genome sequencing of the HGDP-CEPH panel samples [97] adds 929 complete genomes to the 2504 Phase-III genomes already sequenced. Snipper has compiled these HGDP-CEPH data for the gAIMs, FDSPK, PIAP, MAPlex and VISAGE-BT BGA tools. The HGDP-CEPH whole-genome variant catalogs expand reference population data for regions not previously covered for these tools, notably (Native) America, Middle East, and Oceania (although just two small population samples from Papua New Guinea). Snipper provides a flexible method to adjust which reference population data are compared by user-defined grouping selected to best match the observed patterns. For instance, an indication of European-Middle East-South Asian co-ancestry can focus on just these reference populations in a PCA, which expands the 2D space reducing the overlap between reference clouds. In the second 1000 Genomes project, the original 2504 Phase-III samples have been re-sequenced at higher coverage levels (an average 30X) by the New York Genome Center [116]. The data generated from higher average sequence coverage markedly improves the quality of many of the original variant calls, which often changes the genotypes for a significant proportion of 1000 Genomes samples in certain SNPs. The importance of reliable variant calling is illustrated here by the example of SNP rs3857620, reported by Zhao et al. [117] to have a South Asian-indicative A-allele found at much lower frequencies in other population groups and therefore suggested to be highly informative for South-Asian-European differentiations. The 1000 Genomes Phase-III data listed online (http://www.ensembl.org/Homo_sapiens/Info/Index) has an A-allele frequency of 0.461 in South Asians and 0.003 in Europeans at this SNP; but A-allele frequencies from the 30X high sequence coverage analysis are < 0.001 in *all* populations, including South Asians – so when more accurate SNP genotype calls are made from ten-fold increased levels of sequencing coverage, the rs3857620 SNP variation is uninformative for all population comparisons. In the revised population reference data sets in Snipper, all 1000 Genomes data now compiles the 30X high coverage variant calls released in 2020, with the added advantage that some AIMs now have data not previously available in Phase-III listings – e.g., PIAP AIM rs10954737.

It is important to re-emphasize the point made in a previous discussion paper [2] and in the introduction section, that the ancestry inferences made of an unknown DNA donor from forensic BGA testing is entirely dependent on the reference population data available to the BGA analysis of the forensic sample. As mentioned before, this notion applies to all three components of FDP, but is especially important for BGA prediction where population reference data are used directly in BGA prediction tools, while in appearance and age prediction reference data are applied to the prediction models implemented in the prediction tools. Therefore, in BGA prediction the reference data and their limitations impact the prediction outcome more directly than in appearance and age prediction. It is thus important to communicate the reference population dataset used for such inferences and its limitations in the final report sent to the investigating authorities.

3.3. Statistical approaches used by MPS-based forensic BGA tools

Most forensic BGA tools rely on likelihood-based methods to predict a donor's most probable geographic region of biological i.e., genetic ancestry [118,119]. In individuals with significant levels of co-ancestry from families or populations with admixed genetic backgrounds, the likelihood approach tends to break down, as alleles indicative of multiple ancestries are present. Varied proportions of indicative alleles markedly reduce the LR comparing the two most likely origins, as both divisor and numerator will have relatively high probabilities [119]. For this reason, model-based genetic clustering methods such as STRUCTURE [120,121] and ADMIXTURE [[122] are often used to identify and examine the distribution of genetic clusters in an individual's DNA. Matching genetic clusters of unknown and reference samples provides a direct pointer to a person's co-ancestry from the ratio of multiple clusters when present at high proportions. Such patterns can indicate historical population-scale admixture, which is typically found in continental margin regions, e.g., North Africa, instead of recent admixture in the individual's direct family. Apart from the PCA-only ancestry analysis module in the FDSPK UAS software, Bayes likelihood comparison models have predominated in the downstream analysis of AIM SNP genotypes from PIAP, as well as gAIMs, MAPlex and VISAGE-BT-AA custom panels. Usefully, FROG-kb allows input of 55 of the 56 FDSPK AIMs to compute likelihood-based predictions of ancestry [85]. UAS and FROG-kb were formally compared by Sharma et al. [84] finding a 12 % error rate in UAS (32/266 donors with self-declared ancestry) vs 9.8 % in FROG-kb; with most UAS error in South Asians – a population without reference data in the UAS PCA module. For the VISAGE-ET-AA, a probability-based approach without LR implementation was applied and implemented in the VISAGE Software, which provides probability estimates for all geographic regions considered.

3.3.1. Snipper and PIAP HID SNP Genotyper likelihood analyses

As described above and in detail previously [1], likelihood analysis forms the core statistical approach used in Snipper [118]. The commonly-used 'multiple profiles' Snipper option (http://mathgene.usc.es/snipper/analysismultipleprofiles.html) provides the most flexibility – where reference/unknown SNP profiles are distinguished by '1′/'0′ end-column labels and profiles can be re-arranged into multiple input data worksheets in Excel. This portal allows adjustment for non-independence (option: 'Hardy-Weinberg principle not applicable') and generates 2D PCA plots for principal components (PCs): PC1 vs PC2; PC1 vs PC3; and PC2 vs PC3, with black 'unknown' points linked to each prediction likelihood for casework SNP profiles. For every new forensic MPS ancestry panel published, Snipper compiles full 30X high-coverage 1000 Genomes/CEPH population data for use with the multiple profiles portal. This data currently consists of a representative population or, to best match numbers, a set of populations, for each continental group (Phase III: African Yoruba; CEPH European from Utah; East Asian Han Chinese; South Asian Gujarati from Houston, plus HGDP-CEPH Oceanian Papua New Guinean; Middle East Emirati, Saudi and Yemeni populations from the whole-genome-sequencing studies of Almarri et al. [123]; HGDP-CEPH American, comprising five native populations). A separate dataset of SNP profiles lists the above Europeans, South and East Asians with HGDP-CEPH Middle East (four Israeli Arab populations) and HGDP-CEPH North Africans (Mozabite Algerian), to enable LR analysis and PCA of Eurasian population sub-sets. Combined 1000 Genomes and HGDP-CEPH samples total 2535, from 69 non-overlapping populations (Japanese, Tuscan and Nigerian Yoruba populations are common to both sets but have different samples). A further 486 admixed population samples of the 1000 Genomes are listed in a separate worksheet, as well as 130 SGDP samples and 402 EGDP samples in the final worksheet, to enable 'test profiles' to be introduced in customized analyses.

The HID SNP Genotyper ancestry plug-in of the Torrent Suite™ Software from Thermo Fisher Scientific (HSG-TSS) analyses PIAP genotypes and provides LR values based on 51 widely distributed populations from seven 'root populations' comprising: Africa; America; East Asia; Europe; South Asia; Southwest Asia (i.e., Middle East); and Oceania; used for admixture analysis. The population likelihood calculations create a ranked list of values using LRs of all population comparisons, so a typical example could be that an East Asian individual has highest likelihoods for Ami (Taiwanese Aboriginal), Japanese-HapMap, Japanese, Korean, as the four highest LRs. A confidence value is given to predictions based on likelihoods obtained, allowing the

user to exercise caution when relatively low likelihoods are returned. The extent to which users adopt one or more of the population-specific LRs to assign ancestry more precisely than root population inference is not known. Several studies have specifically discussed this interpretative choice [86,93,118,119,124] and as a sensible rule-of-thumb the best interpretation is to consider an individual originates from any listed population showing an LR of 10 or less. The HSG-TSS admixture analysis algorithm uses a bootstrapping system to estimate co-ancestry proportions in individuals with admixture. Admixture proportions are estimated based on a maximum likelihood approach comparing the seven root populations with bootstrapping replication runs analyzing a different subset of PIAP SNPs in each replication to capture uncertainty in the estimations. Co-ancestry estimates use the average of the bootstrapping replications for each population which are presented as a percentage of each contributing population with the corresponding likelihood.

The informative study by Jin et al. [124] focused on the admixture analyses made by the HSG-TSS algorithm, using 34 admixed donors with self-declared co-ancestries, plus 648 single-ancestry donors. Results indicated generalized root population inferences for single-ancestry donors had ~99 % reliability (643 predictions matched self-declarations), but admixed donors had more inaccurate predictions when the co-ancestry patterns in these individuals were complex, or their co-ancestries were from closely related populations, e.g., Europe vs Southwest Asia. In a similar HSG-TSS study with fewer test samples, Al-Asfi et al. [125] found concordant co-ancestry predictions for 4/11 admixed donors, whereas 7/11 were given more co-ancestry components by HSG-TSS than was known or declared. This study looked in more detail at the individual population inferences based on ranked LRs and found 22/36 single-ancestry donors had a prediction matching their true ancestry amongst the top five population-specific inferences. Both studies advised caution with forensic casework analyses that to give data in the report which goes beyond a simple single-ancestry inference with an accompanying high confidence score, although simple parental co-ancestry predictions (i.e., at or around 50:50 ratios) for divergent root populations were the most reliable.

3.3.2. GenoGeographer

GenoGeographer is a data analysis software developed by Tvedebrink et al. [126] to gauge the similarity or dissimilarity of MPS AIM SNP profiles and the reference population data used to make BGA inferences. GenoGeographer delivers a likelihood ratio test (LRT, distinct from LR) recording the absolute concordance between an AIM SNP profile and a population (computing a z-score) rather than a relative measure of the profile's likelihood in two populations, (as in LR analyses discussed above). LRT analysis adjusts for the possibility in forensic casework that no reference population is appropriate for the donor's true origin, so all null hypotheses are rejected when no relevant reference data can make a reliable inference. Mogensen et al. [127] gauged the efficiency of GenoGeographer using the PIAP and adding Greenlander and Somali population reference data to the widely used continental population groups of Africa, North Africa, Middle East, Europe, South/Central Asia, and East Asia obtained from FROG-kb, treating Greenlander/Somali profiles as unknowns. GenoGeographer marked 22.4 % of test profiles as not assignable to any reference population used, while ~84 % of the remaining 77.6 % were correctly assigned to either Greenlander or Somali. In contrast, conventional LR analyses gave 78.1 % correct and 21.9 % incorrect assignments. Overall, GenoGeographer was able to reduce the error rate three-fold by using the z-score to exercise caution when its value was above a certain level and prevented erroneous assignments being made.

In a similar, but simplified manner, Snipper allows the cross-validation of the population reference datasets compiled for forensic MPS BGA tools, and when the likelihoods generated are ranked in a series of 'LR plots' the unknown profile position can be superimposed to gauge whether its LR is within the value ranges observed or is an outlier (i.e., reliable vs unreliable LRs). The lowest values in each reference population, or the LR values seen in incorrect assignments allows an inference threshold to be set to reduce the risk of incorrect assignments from below-threshold values. Fig. 2 of [128] shows an example of such plots, applying a universal LR threshold of 1000 (i.e., no assignment reported for LRs below '1000-times more likely one population than another'). Population-specific thresholds can also be set when some populations are less divergent than others in the reference data used.

3.3.3. Combining different types of AIMs in BGA prediction frameworks and tools

The introduction of MHs as AIMs has highlighted the problem of collecting suitable population reference data for multiple-allele markers with many low frequency alleles/haplotypes, therefore requiring much larger sample sizes to properly estimate frequencies. Since sampling of Oceanian and American populations remains scant, it is difficult to use MH loci without significant inaccuracy in the calculation of likelihoods for under-represented population groups using these markers. Furthermore, haplotypes need to be counted rather than estimating allele frequencies from single SNP genotypes, so is dependent on large sample sizes to be representative of the actual haplotype frequencies in the population. Consequently, Snipper has so far failed to adapt LR analysis of panels combining SNPs and MHs, as the underlying variation needs to be estimated in different ways. Nevertheless, as with STR data, STRUCTURE can accept both types of markers as joint input applying suitable numerical transforms to haplotypes (e.g., AAA=111; ACA=121, etc.). It should be remembered that PCA analysis only works with binary SNP data, precluding tri-allelic and MH variation, and it is recommended to apply more appropriate methods of multi-dimensional scaling able to handle binary/multiple-allele SNP and MH data. R-scripts are available to execute principal coordinates analysis (PcoA [129]) and neighbor-joining tree plots (see Fig. 8 C in [82] and Figs. 3–4 in [109]) to better represent genetic distances in 2D space using all the variation in a panel. Lastly, experiences with MAPlex and MPSplex indicates LRs are often reduced when MH likelihoods are combined with SNP likelihoods. This suggests that MH data in such mixed panels are best reserved for monitoring and de-convoluting mixed DNA, with the potential to infer each contributor's BGA [109].

3.4. The continuing challenge of assessing BGA in cases of genetic admixture and co-ancestry

GenoGeographer is designed primarily to address the problem of genetically admixed individuals analyzed with single-ancestry reference data [126]. Genetically admixed individuals with co-ancestry from different geographic regions, present a problem for likelihood-based ancestry tests producing reduced values [83], which are then not reliable enough indicators of ancestry for reporting to investigators. Pereira et al. used STRUCTURE rather than GenoGeographer to study Brazilian population samples and families (sibs have similar co-ancestry ratios and predictably combine their parent's ancestral backgrounds), as a testbed for population and individual admixture measurement, respectively [130]. Brazilian populations are known to have co-ancestry from European, African, and Native American populations. As would be expected, the larger the AIMs panel (ranging from 46 Indels to 164 PIAP SNPs/210 AIMs combined), the less variable the co-ancestry estimates obtained; with a more marked effect on consistency of individual co-ancestry estimates than population level estimates. For an individual suspect in a crime case, accurately estimating co-ancestry for genetically admixed individuals is more difficult if: i) the admixture contributors have lower than average divergence; ii) the AIMs panel is small; and iii) there is insufficiently balanced $I_{n\ \mathbf{POP}}$ levels, given the AIM set and reference populations used (see Section 3.1.2.).

The above points suggest that reliable analysis of complex genetic admixture remains difficult to achieve with the current forensic BGA tools, even those with expanded autosomal SNP numbers currently

available for tools based on targeted MPS. However, there are several factors that can help both the interpretation of allelic variation detected in the tested individual and the design of future AIM sets, which were considered by the VISAGE Consortium when designing the AIM panel of the VISAGE-ET-AA. First, with existing MPS panels it should be possible to report co-ancestry in the SNP profile analyzed and infer contributing population groups when simple, balanced parental admixture is detected. The study by de la Puente et al. [109] assessing the capabilities of a large-scale MH panel for ancestry inference even suggests that the co-ancestry components of a mixed DNA contributor with an admixed background are detectable, without the expected confounding effects of 'double mixtures' (Fig. 5 in [109]). Second, when an individual originating from continental margin population has identical patterns of admixture to those of individuals with admixed family history (e.g., North Africans have the same patterns as European-African admixed individuals), it is important to consider both possibilities in the ancestry report. A set of reasonable and robust guidelines for reporting individual co-ancestry has been proposed by Jin et al. [124] and readers are encouraged to consider these as a framework for building their own reporting guidelines, based on the AIMs panel used and its limitations for measuring co-ancestry (admixed 1000 Genomes population data for all published MPS BGA panels and tools are available in Snipper for this purpose). The third factor is the most important to consider as BGA MPS panels continue to expand in the number of AIM SNPs and become more sophisticated. It should be increasingly easy to include non-autosomal marker data either from parallel MPS assays or within the same enlarged MPS multiplex. In this way, the paternal and maternal lineages can be compared with the autosomal data in admixed males. The FDSPK A-B multiplex combines 56 AIM SNPs and 24 Y-STRs, but to our current knowledge this data has not been combined in any forensic ancestry studies. Likewise, the TFS Precision ID Identity Panel has 34 upper Y-Clade SNPs, which although a separate MPS multiplex, adds patrilineal analysis to PIAP data. A separate multiplex MPS tool for simultaneous analysis of 859 Y-SNPs allowing the inference of 640 Y-haplogroups has been published [77] and is available for high-resolution Y-haplogrouping, allowing detailed paternal ancestry inference. As the increased copy-number of mtDNA makes the balancing of mitochondrial sequence with autosomal and/or Y-chromosomal sequences with the same MPS tool challenging, future efforts need to find out if such combined MPS tool can be developed for forensic applications. In any case, separate whole mitogenome MPS tools are available for separate analysis of maternal ancestry [78]], including a commercial solution [131]. As described above, the VISAGE-ET-AA tool already combines autosomal AIM SNPs with Y-SNPs for simultaneous bi-parental and paternal ancestry inference, although genotyping 87 Y-SNPs limits the level of paternal BGA inference and will need to be increased in future FDP tools. Similarly, X-SNPs provide a reasonably informative substitute for mtDNA sequence analysis for the inference of X-chromosome ancestry in a large proportion of males with recent admixture.

3.5. *Ancestry inference with genealogy-scale SNP genotyping*

Although MPS has been the key technological development in forensic DNA analysis since 2015, arguably the recent rapid evolution of investigative genetic genealogy methods based on genomic data obtained with SNP microarrays, or whole-genome sequencing has also changed our attitude to what could be possible in the future [132]. SNP microarray-level genotyping is generally not suitable for forensic material [133], although there can be crime scene samples that contain DNA of high enough quantity and quality to even allow successful SNP analysis using whole-genome sequencing [134]. Hence, a middle ground between hundreds of thousands of genome-wide SNPs present on microarrays and current MPS tools containing up to hundreds of AIM SNPs is a potential future way to address many of the issues covered above – particularly improving complex co-ancestry analysis. The medium-scale MPS-based Kintelligence SNP tests from Verogen intended for long-range familial searching using dedicated portions of the GEDmatch database (termed GEDmatch PRO), could also provide data-rich ancestry information when combined with the 1000 Genomes/HGDP-CEPH reference populations that can now be readily compiled in Snipper. In combination with evolving data analysis regimes, such panels of several thousand SNPs could potentially provide greater detail, a wider range of sub-continental population differentiations and more refined co-ancestry analysis. In this way, an unidentified contact trace could be simultaneously analyzed for relatives in GEDmatch and have detailed ancestry inferences made. Dedicated tools with many thousands of AIM and appearance SNPs may be developed in the future, for which targeted MPS methods involving capture enrichment appear more promising than amplification based targeted MPS methods, especially for degraded DNA samples. In 2021, the FORCE panel based on capture MPS was published [31], designed to be an all-in-one tool for 5,422-SNPs for investigative genetic genealogy and other forensic purposes. The scale of the FORCE panel allows the combination of several thousands of SNPs with several distinct forensic purposes with 4069 kinship/identity SNPs, 241 BGA SNPs, 41 HIrisPlex-S SNPs for eye, hair, and skin color (4 overlapping BGA SNPs), 246 X-SNPs, and 829 Y-chromosome SNPs. The BGA SNPs in FORCE were compiled from the merged PIAP and VISAGE-BT-AA panels. While in their paper, the authors describe ancestry resolution with the FORCE panel based on 5 continental groups, the enlarged number of autosomal BGA SNPs in combination with 829 of the Y-SNPs from Ralf et al. [77] may allow FORCE to provide enhanced BGA inference compared to the smaller MPS-based BGA tools described in this review, which, however, remains to be formally assessed.

4. Recent progress in predicting age from crime scene DNA

Obtaining investigative leads from crime scene DNA within the concept of FDP gains extra power when the prediction of age is included together with appearance and ancestry. This not only is because age per se allows the characterization of a person, but also due to the fact that the expression of certain appearance traits depend on age, and for some appearance traits, age is used as predictor in the genetic prediction models (Table 1). While various approaches for age estimation have been considered in forensic molecular biology, only DNA methylation (DNAm) analysis has proved to be sufficiently accurate to provide a practical solution for forensic applications. The first papers mentioning the forensic usefulness of epigenetic age estimation appeared in 2011 and presented the first DNAm markers potentially useful for this purpose [135,136]. Since then, numerous studies have been published on epigenetic age prediction, many of them containing results relevant to the forensic field. Progress on the implementation of DNA methylation markers for forensic age prediction has been summarized in several previous review articles [9,10,137]. In this part of the review, we focus on summarizing the most recent advances in forensically relevant epigenetic age prediction published since these last review article on this topic in 2016–17.

In recent years, several DNA methylation-based age estimation tools suitable for forensic applications have emerged, as discussed below and summarized in Table 3, in addition to those published earlier and summarized elsewhere [10]. Notably, early research showed that age estimated using DNA methylation markers has reduced accuracy in the elderly, which is related with a global decrease in the stability of DNA methylation with advanced age [10]. This problem is a consequence of inter-individual variability in the rate of aging, which is more evident in the elderly and can be due to both hereditary DNA variants, environmental factors influencing the rate of DNAm progression, and stochastic effects. The inclusion of additional markers may reduce this problem. A study by Cho et al. suggested that the accuracy of age estimation based on DNA methylation in *ELOVL2*, *C1orf132*, *TRIM59*, *KLF14*, and *FHL2* can be increased in elderly people by adding information beyond

Table 3
MPS-based tools for predicting epigenetic age from crime scene DNA.

Name	Tissue	Number of CpGs	MPS technology	Forensic validation	Prediction model, reference data, prediction error	Prediction tool	References
Vidaki et al.	Blood	16	Illumina	No	Model building N = 1156 (microarray data) MAE= 3.3; model validation N = 46, MAE= 7.5	No	[205]
Naue et al.	Blood	13	Illumina	No	Model building N = 208, MAE= 3.21; model validation N = 104, MAE= 3.16; Reduced model (4 CpGs): MAE= 3.24 and 3.64, respectively	No	[181]
Hong et al.	Saliva	7	Illumina	No	Model building N = 154, MAE= 3.09; model validation N = 36, MAE= 3.19	No	[162]
VISAGE Basic Tool for age from blood	Blood	32	Illumina	Limited	No	No	[159]
VISAGE Enhanced Tool for age from somatic tissues	Blood, buccal cells, bones	44	Illumina	Yes	3 models for 3 tissues: Blood (6 CpGs): model building N = 112, MAE= 2.2; model validation N = 48, MAE= 3.2; Buccal cells (5 CpGs): model building N = 112, MAE= 2.5; model validation N = 48, MAE= 3.7; Bones (6 CpGs): model building N = 112, MAE= 3.3; model validation N = 49, MAE= 3.4	VISAGE Software	[15]
VISAGE Enhanced Tool for age from semen	Semen	13	Illumina	Yes	Reduced model (6 CpGs): model building N = 125, MAE= 4.3; model validation N = 54, MAE= 5.1	VISAGE Software	[16,163]
Aliferi et al.	Blood	18	Illumina	Limited	Model building N = 77, MAE= 3.6; model validation N = 35, MAE= 3.3	No	[152]

MPS-tools are listed in time-wise order of publication, for those published earlier than 2017, see Freire-Aradas et al. [10], MAE: mean absolute error: e.g., an MAE of 4 means that in the studied samples the mean absolute difference between predicted age and calendar age was 4 years.

methylation status from sjTREC DNA analysis [138]. Another problem is the under-representation of the young in epigenetic age prediction research. To start to fill this gap, some studies aimed to identify appropriate age predictors in children [139,140]. Freire-Aradas et al. identified six genes with CpGs highly correlated with age in the young and proposed a model based on these predictors dedicated for epigenetic age prediction in children. *KCNAB3* was found to be a particularly informative marker for measuring methylation progression during childhood and adolescence [139]. The same research group proposed prediction model trained using individuals in a wide age range from childhood to old age [141]. The quantile regression neural network model used included CpG sites from *ELOVL2*, *ASPA*, *PDE4C*, *FHL2*, *CCDC102B*, *MIR29B2CHG* and chr16:85395429 (GRCh38) and predicted age in a test set with a median absolute error (MAE) of 3.32 [141]. McEwen et al. [140] developed an epigenetic clock based on 94 CpGs that can predict age in young individuals with an MAE of only 0.35 years, using a test set consisted of 689 individuals.

Moreover, it is known that DNA methylation, depending on the CpG site, is genetically controlled, but can also be influenced by environmental factors. A detailed review of this topic is beyond the scope of this article, but can be found elsewhere [142]. Some studies consistently indicate that age correlated CpGs may have different sensitivities to confounding external factors such as disease. For instance, Spólnicka et al. studied age-related DNAm changes in groups of patients with late-onset and early-onset Alzheimer's disease, as well as in patients with Graves' disease [143]. This study reported an unchanged age prediction accuracy for DNA markers *ELOVL2* and *MIR29B2CHG* in the patient group, while DNAm levels in *TRIM59*, *FHL2* and *KLF14* were modified in the patient group, influencing the results of age prediction in patients suffering from these diseases. Also, using the same model, unreliable results for age prediction were reported in patients with chronic lymphocytic leukemia [144] (mean epigenetic age of patients was 10.7 years higher than the mean true age in this group). Assuming such DNAm alterations can serve as a diagnostic biomarker of the disease, the authors used the DNAm data to develop a calculator that allowed the prediction of this disease [144]. In a different study, Spolnicka et al. [145] found that intense physical exercise may modify methylation in some age markers. The study observed elevated DNAm levels in *TRIM59* and *KLF14* and explained this modification as a methylation response to stress associated with the lifestyle of elite athletes. The only marker that remained without detected influence from the tested external factors in these studies was *ELOVL2*, which further emphasizes its importance in predicting chronological age in forensic tests, as originally suggested by Zbieć-Piekarska et al. in 2015 [146]. Analysis of 104 children with different growth disorders showed DNA methylation may be impaired in some CpGs and, consequently, the precision of age prediction may be reduced in children when using these predictors [147]. The VISAGE Consortium and Project also addressed this issue and investigated 8 DNAm age predictors (*ELOVL2*, *MIR29B2CHG*, *TRIM59*, *KLF14*, *FHL2*, *EDARADD*, *PDE4C,* and *ASPA*) in a cohort of individuals with severe alcohol dependency [148] by using the VISAGE Enhanced Tool for age prediction in somatic tissues [15]. Of the 8 markers analysed, altered DNAm was only observed in *MIR29B2CHG*, but the impact on age prediction was small [148].

It is important to note that because crime scene investigation deals with challenging biological materials often containing very limited amounts of DNA which typically is degraded, the DNAm analysis technology used for epigenetic age prediction sets the limits within FDP. Although DNAm microarray technologies enable the simultaneous analysis of large numbers of CpGs, in the same way SNP arrays do for a large number of SNPs, hybridization-based microarray technology is generally not suitable for typical crime scene DNA. Foox et al. recently proposed whole genome bisulfite sequencing as a possible solution to analyze large number of CpGs in challenging forensic material [149], but genomic approaches may not be practical for implementation in most forensic DNA laboratories. Therefore, most forensically motivated studies have focused on developing age prediction models based on relatively small numbers of carefully selected age-predictive DNAm markers that can be analyzed with DNA technologies suitable for crime scene DNA analysis. DNAm-based age prediction has several technological hurdles that do not exist in SNP-based appearance and ancestry prediction. First, DNAm marker analysis requires technologies providing fully quantitative outcomes, instead of qualitative analysis of SNPs. Second, all technologies for DNAm analysis currently used in forensics rely on a bisulfite conversion step prior to the actual DNAm analysis step. Bisulfite conversion requires more DNA than needed for SNP-typing and degrades the DNA during the conversion procedure; hence, typically degraded forensic DNA gets further destroyed. Third, technical errors in DNAm analysis vary between analysis technologies (see method-to-method bias below). Some studies indicate that reliable DNAm analysis is possible with as little as 2–5 ng of bisulfite-converted DNA [150–152]. Very good sensitivity results for DNA methylation

testing were obtained using MPS in the work of Aliferi et al. [152]. Analysis of 11 age markers was possible with 5 ng of initial DNA, which is ~1 ng PCR input. The prediction accuracy of MAE = 3.3 years at such low DNA concentrations represents a significant improvement of existing levels of sensitivity in forensic tests. The algorithm was validated using an independent cohort and obtained an accuracy of MAE = 3.8, and for those under 55 years of age, an MAE = 2.6 [152].

In general, there is a risk that age prediction in samples with bisulfite-converted DNA will have amounts of converted DNA that are below the empirical sensitivity threshold of the analyzing lab, which will lead to increased age prediction error. An *in-silico* analysis conducted by Naue et al. [153] demonstrated, as can be expected, that in samples with low DNA amounts, stochastic effects increase and can lead to significant fluctuations in DNAm levels, which affect the age prediction outcome. This was also observed empirically outside the forensic field by Smith et al. [154], who reported that samples of low DNA concentration show higher variability in their measures of methylation level than samples of high DNA concentration. In forensic traces, low levels of input DNA are often degraded, and even more so after bisulfite conversion. However, a recent study testing blood of deceased persons confirmed previous conclusions that DNA methylation is a stable chemical modification not strongly affected by tissue decay, suggesting that the reliability of age predictions from crime scene samples is mostly determined by low DNA quantity [155].

Regarding DNAm marker analysis for age prediction, various technologies have been used thus far, including pyrosequencing, Sequenom, MassARRAY, MALDI-TOF, SNaPshot, single base extension using locked nucleic acid primer modification, droplet digital PCR, and targeted MPS [156–159]. Each of these methods may produce different technical errors in DNAm analysis (see method-to-method bias below). Richards et al. [160] noted that there is no agreed methodology for the targeted detection and analysis of DNAm markers in forensic studies yet and considered the benefits of using MPS for this purpose. Fleckhaus and Schneider [161] showed that multiplexing of several CpG targets is difficult but possible with simple pyrosequencing and suggested that multiplex pyrosequencing can be a viable alternative to other technologies with reproducibility equal to a singleplex pyrosequencing. However, pyrosequencing has much less multiplexing capacity than targeted MPS.

The multiplexing of DNAm markers is further limited after bisulfite conversion of DNA. With high degrees of DNA degradation, reduced DNA sequence complexity and lack of DNA strand complementarity, which are all characteristic of bisulfite-treated DNA, the design of PCR primers becomes much more challenging. In particular, regions rich in CpG sites are problematic to analyse. Balancing the amplification efficiency of individual PCR reaction targets is also not straightforward for multiple DNAm markers. Some studies have shown that small-size DNAm multiplexes can be optimized [151,162]. Targeted MPS is particularly effective as a tool for hundreds (or thousands) of SNPs from forensic DNA, while multiplexing capacity for DNAm markers is lower because of the necessity of quantitative analysis. From a practical perspective, it is noteworthy that MPS technologies are nowadays available in most forensic DNA laboratories, facilitating the implementation of new methods based on targeted MPS, including for DNAm analysis. In line with that, the VISAGE Consortium developed and validated three tools for age prediction based on targeted MPS, i.e., i) the VISAGE Basic Tool for age estimation in blood [159], ii) the VISAGE Enhanced Tool for age estimation in somatic tissues [15] and iii) the VISAGE Enhanced Tool for age estimation in semen [163], which are further described below. Beyond the activities of VISAGE, targeted MPS as a method of DNA methylation analysis in the forensic field is also under continuous development [152].

An important technology-driven issue in quantitative DNAm analysis highly relevant for age prediction is the method-to-method bias i.e., differences in technical variation of DNAm quantification between different genotyping methods used for the different steps of DNAm marker discovery, model data establishment, and final casework application. Hence, the application of different DNAm analysis methods for the different steps lead to artificially increased age prediction errors [162,164]. This is often seen when the reference data used in prediction modelling were established with DNAm microarrays, easily applicable to many individuals, while different technology is used for the forensic tool applied to crime scene DNA. Feng et al. proposed a solution based on Z-score transformation to correct for differences in DNAm data between model reference data generated with one analysis technology applied to actual casework data produced with another technology. These authors demonstrated that Z-score transformation was successful in using a prediction model developed from data produced with Epityper microarrays to predict age in samples analyzed with pyrosequencing [164]. The model was then successfully validated to predict age in blood deposited on FTA cards or gauze, and stored under room conditions [165]. However, when the method-to-method differences in DNAm data become too large, simple mathematical transformation methods cannot compensate any longer. In one study, investigation of four widely used methods for targeted DNA methylation analysis showed comparable DNA methylation and subsequent age prediction values obtained using EpiTYPER®, pyrosequencing, and MPS (Illumina) technologies and the biggest differences reflected in higher prediction errors for SNaPshot [156]. The SNaPshot assays provides a semi-quantitative measurement of DNA methylation, and when using this technology, differences in DNA methylation measures were noted even with different capillary electrophoresis instruments [166,167]. To avoid potential method-to-method bias, it is strongly recommended to use the same DNAm analysis technology for model data generation for both, model building and model validation, and for the final forensic casework applications [162]. The VISAGE Consortium followed this rationale and developed DNAm sequencing tools based on targeted MPS for future casework application as well as applying them for establishing the reference data used for prediction model building and validation, thereby avoiding any method-to-method bias.

Not only DNAm markers and technology used for DNAm analysis, but also the statistical method applied for DNAm-based age prediction impact on the accuracy of epigenetic age prediction. Several studies demonstrated advantages of using machine learning (ML) methods for DNAm-based age prediction [168,169]. Aliferi et al. [151] tested 17 ML methods, detected differences in their predictive outcomes and found the best performance of a support vector machine with a polynomial function. It is known that ML methods are sensitive to data size and this topic was further investigated by the same group in the context of forensic age estimation [170]. Notably, ML advantage over standard methods such as regression analysis was not seen in appearance prediction on the example of eye, hair, and skin color [171].

Finally, a very important factor in DNAm-based age prediction in the FDP context, which does not apply to SNP-based prediction of appearance and ancestry, are differences in DNAm patterns between different tissues or cell-types from which crime scene DNA is extracted. As crime scene samples can derive from different tissues, this is highly important in forensic age prediction. Developing a universal epigenetic predictive system that would sensitively and accurately predict age across the range of human tissues and body fluids for forensic applications appears not feasible (see below), at least thus far with a limited number of CpGs that can be analyzed with forensically suitable DNA technology such as targeted MPS. Horvath's universal epigenetic clock can predict age in different cell types because it uses a large number of CpGs typically generated with DNA methylation microarray technology, unsuitable for crime scene DNA analysis. However, even with 353 CpGs, the analysis error noted for some tissues was higher than for others. For example, in muscle tissue, the median absolute difference between epigenetically predicted age and true chronological age was 18 years, while it was 3.7 years for whole blood based on the same DNAm markers [172]. Notably, the apparent universality of the model was made possible by using a training set consisting of methylation data for many thousands of

samples from different cell types. It can be expected that with lower numbers of CpG predictors, tissue specificity effects increase. However, it has been shown that some DNAm markers provide age information more universally across tissue types, despite different rates of DNAm change, leading to strong age correlation across different tissues [173,174]. However, most of the age-related DNA methylation sites have tissue-specific effects [173]. This has led to the notion that age prediction from DNA of different tissue sources requires different DNAm prediction tools based on different sets of DNAm markers selected in a tissue-specific way. In some cases, particularly when analyzing somatic tissues that overlap in their DNA methylation patterns, the same DNAm markers and thus genotyping tools can be used, but prediction models for each tissue based on tissue-specific data is required to achieve optimum prediction accuracy. This idea was adopted by the VISAGE Consortium, which developed two Enhanced MPS tools for age prediction, one for somatic tissues [15] and one for semen [163], and used four different prediction models, one for each of four tissues [15,163]. Given the tissue dependency particularly affecting forensic age estimation based on small sets of DNAm markers, in the following sections, we describe recent advances in predicting age from forensic DNA separately for the major tissue types commonly confronted with in crime scene investigations.

4.1. Epigenetic age estimation from DNA in blood

Blood is a common source of DNA in forensic analyses, very often collected at the crime scenes in the form of bloodstains that typically indicate violent crime. Therefore, the practical value of predicting age from blood-derived DNA is obvious, and many studies that have considered the forensic importance of epigenetic age prediction focused on blood as the source of DNA [10]. The first analyses to select DNA methylation age predictors based on the Infinium HumanMethylation27 BeadChip DNA methylation microarray data failed to identify *ELOVL2*, which is included in the updated Infinium HumanMethylation450 BeadChip version and turned out to be the most important forensic age predictor. CpGs from *ELOVL2* were suggested as age predictors for the first time in 2012 by Garagnani et al. [175], and were included in the early model proposed by Hannum et al. [176]. The usefulness of *ELOVL2* in forensic epigenetic age prediction has been confirmed since by many studies and CpGs from this gene got included in most forensic epigenetic age prediction tools.

Differences in DNA methylation levels between human populations have been reported [177], and this phenomenon is equally applicable for age predictive DNAm markers. For instance, although the 5-CpG model developed by Zbiec-Piekarska et al. [178] based on Polish data has been successfully validated in Korean data, Cho et al. demonstrated that remodeling based on Korean data improved the prediction accuracy considerably i.e., an MAE from 4.18 to 3.29 [138]. The same age markers were also studied in Italian and Portuguese population samples. In Italians, a final predictive model included eight CpGs from *ELOVL2*, *FHL2*, *MIR29B2CHG* and *TRIM59*, thus excluding *KLF14*, and giving a predicted age with MAE = 4.5 years [179]. The same four markers were also confirmed in Portuguese, achieving an MAE = 4.97 years [180]. It is noteworthy that DNA methylation in *KLF14* varies over a person's lifetime within a very narrow range (e.g., ~0–10 % for a cohort at age range 1–75 years) [15], which may be difficult to test with some technologies such as semi-quantitative SNaPshot.

Other recently published age prediction models based on DNA from blood further confirmed some well-known age DNAm markers and suggested potential additional age predictors. Pan et al. (2020) studied seven previously known DNAm markers for blood including CpGs from *ASPA, EDARADD, CCDC102B, ZNF423, ITGA2B, KLF14*, and *FHL2* in a Han Chinese population, and confirmed the age correlation for all of them [169]. Feng et al. (2018) studied 390 Chinese at 21 age-related differentially methylated regions and confirmed the highest suitability for epigenetic age estimation of *ELOVL2*, *TRIM59*, *MIR29B2CHG*, *PDE4C*, *CCDC102B*, *RASSF5*, and a DMR on chr10:22334463/65, giving an MAE of 2.49 years. Naue et al. [181] proposed two age prediction models developed using amplicon bisulfite MPS data and a random forest regression prediction method. Their basic model contained 13 CpGs from *DDO, ELOVL2, F5, GRM2, HOXC4, KLF14, LDB2, MEIS1-AS3, NKIRAS2, RPA2, SAMD10, TRIM59, ZYG11A* and predicted age with a MAE of 3.16 years. A reduced model based on the strongest CpG predictors from *ELOVL2, F5, KLF14*, and *TRIM59*, resulted in age estimation with only a slightly higher error (MAE = 3.24). Lau and Fung [182] attempted to find an optimal set of DNAm markers for blood by re-analyzing the HumanMethylation450 BeadChip data for blood and used several methods for marker selection and prediction modelling. This study finally recommended a simple forward selection and multiple linear regression to produce the best predictive performance with the most accurate model restricted to 16 CpG predictors and included well established markers from *ELOVL2*, *FHL2, KLF14*, and *TRIM59* with an MAE = 3.76 years. The VISAGE Consortium implemented the five markers previously proposed by Zbiec-Piekarska et al. [178] i.e., *ELOVL2, MIR29B2CHG, FHL2, TRIM59* and *KLF14* for developing the VISAGE Basic Tool for Age estimation in blood based on a total of 32 CpGs [159]. The VISAGE Enhanced Tool for Age estimation from somatic tissues analyses 44 CpGs from eight genes *ELOVL2, FHL2, KLF14, MIR29B2CHG, TRIM59, EADARADD, PDE4C* and *ASPA*. It comes with a prediction model for blood that includes 6 CpG sites from *ELOVL2, MIR29B2C, KLF14, FHL2, TRIM59* and *PDE4C* and predicted age in a test set of 48 blood samples with an MAE of 3.2 years [15] (for more information on the VISAGE Enhanced Tool for Age from somatic tissues, see below).

While the above-mentioned studies all explored known DNAm markers previously identified for age estimation in blood, three studies proposed entirely novel age DNAm markers for blood. Mawlood et al. [183] reported an inverse correlation with age for two CpG sites in mitochondrial DNA and the developed model predicted age with MAE = 9.3 years. Xin et al. [184] analyzed DNAm in the human telomerase reverse transcriptase gene and reported a model based on this single gene that was able to predict the age in the test set of (a very limited number of) 30 samples with MAE of 5.19. Although the prediction errors obtained from these newly proposed DNAm age markers are higher than those achieved with models based on previously established DNAm age markers, in the future these novel markers shall be tested together with previously established ones to evaluate their usefulness for improving age prediction from blood-derived DNA.

Recently, Vidaki et al. [185] studied DNAm-based age prediction in blood from males using CpGs located on the non-recombining part of the Y-chromosome. In blood samples of 1057 males aged 15–87, they identified 75 Y-CpGs correlated with age, of which a subset of 19 CpGs was able to predict age with an MAE = 8.46 years with an age-independent prediction accuracy. An age prediction model and tool based on male-specific Y-chromosomal CpGs is forensically relevant as it may allow predicting age of male perpetrators from analyzing male-female DNA mixtures [185]. For a summary of MPS-based epigenetic age prediction tools from blood, see Table 3.

4.2. Epigenetic age estimation from DNA in saliva and buccal cells

Saliva is also frequently found at crime scenes such as in form of cigarette butts, chewing gum, or bottles; hence, it is a highly relevant source of forensic DNA. It is worth mentioning the interesting, but still unsolved, problem whether whole saliva or buccal cells present in saliva, provide the better source of DNA for developing the most practical epigenetic age model for forensic investigations. Since the early work of Bocklandt et al. [136] introducing a model for saliva based on just three CpGs, several studies have explored this topic, which resulted in several age prediction models for saliva or buccal cells. Hamano et al. [186] reported a model for age estimation in saliva based on CpGs from *ELOVL2* and *EDARADD* that predicted age with an MAE = 6.25 years.

The study confirmed the usefulness of the model for age prediction in cigarette butts with a median absolute deviation (MAD) of 7.65 years, but this requires further confirmation, as the number of samples was small. As in the earlier research on epigenetic age prediction in blood [187], the group used the methylation sensitive high-resolution melting (MS-HRM) method to measure DNAm levels for these saliva studies. Both models, for blood and saliva, involve analysis of *ELOVL2*. Hong et al. [188] proposed an alternative model for epigenetic age prediction in saliva that was based on six age-correlated CpG predictors from *SST, CNGA3, KLF14, TSSK6, TBR1*, and *SLC12A5*, and predicted age with an MAD = 3.15 years. This model was developed based on a discovery analysis using data from 54 HumanMethylation450 BeadChip array datasets and replication in 226 samples using a SNaPshot protocol. The developed assay also included a CpG site in *PTPN7* able to differentiate between leukocytes and buccal epithelial cells, an approach that had previously been used in another method for epigenetic age prediction in buccal cells [189]. In a follow-up study, Hong et al. [162] analyzed the same marker set in a subset of 95 samples using MPS technology and used the data obtained to develop a platform-independent model for age prediction in saliva. Koop et al. [190] evaluated the possibility of using DNA methylation levels in *PDE4C* to predict the age of deceased people from buccal swabs and concluded that this material is a good source of DNA for epigenetic age assessment provided that sufficient DNA can be extracted. This study provided further evidence that DNAm is stable in post-mortem biological samples [190]. The VISAGE Enhanced tool for age estimation from somatic tissues comes with a prediction model for buccal cells that includes five CpGs from *PDE4C*, *MIR29B2CHG*, *ELOVL2*, *KLF14* and *EDARADD* and predicted age in a test set of 48 samples with an MAE = 3.7 years [15] (for more information on the VISAGE Enhanced Tool for Age from somatic tissues, see below). Ambroa-Conde et al. developed a method that included a selection of seven DNA methylation markers for saliva and buccal cells (cg10501210, *LHFPL4, ELOVL2, PDE4C, HOXC4, OTUD7A* and *EDARADD*). The combined model based on multivariate quantile regression allowed prediction in a test set with an MAE = 3.66 years (93 saliva and 91 buccal cell samples). The developed models can be used with the Snipper application suite [191]. As with blood studies, some inter-population differences in DNA methylation patterns were also found in saliva and buccal cells [192,193], although thus far DNAm variation between populations has been studied less in saliva than in blood, which needs to be improved. For a summary of MPS-based epigenetic age prediction tools from saliva and buccal cell DNA, see Table 3.

4.3. Epigenetic age estimation from DNA in other somatic tissues

Next to blood and saliva / buccal cells, various other somatic tissues including muscle, cartilage and skeletal bone and teeth samples, are routinely analyzed in forensic DNA laboratories for the purpose of human identification. In some of these cases, age estimation may provide additional information for facilitating the human identification process, while more rarely, crime scene investigation may involve such tissues. Naue et al. [174] tested 13 DNAm markers previously selected for blood and attempted to predict age in brain, bone, muscle, buccal epithelial cells, and blood obtained from 29 deceased people aged 0–87 years using a model previously developed for blood. They found 7 CPGs to correlate with age in all 5 tissues. The authors suggested that CpGs from *DDO, ELOVL2, KLF14, NKIRAS, RPA2, TRIM59*, and *ZYG11* could provide the core loci for a multi-tissue age predictor in forensics [174]. Two prediction models that were reported for teeth, albeit based on a small number of samples [10]. Bekaert et al. [194] analyzed 29 teeth and Giuliani et al. [195] analyzed 21 teeth for epigenetic age estimation. Marquez-Ruiz et al. [196] examined 65 teeth of people aged 15–85, which confirmed the usefulness of CpGs from *ELOVL2* and *PDE4C* for epigenetic prediction of age from teeth, while a CpG from *ASPA* was not correlated with age in this sample set. Moreover, the study showed that telomere length, although also correlated with age, only slightly increased the prediction accuracy in a combined model with DNAm markers [196]. Lee et al. [197] analyzed 32 skeletal samples using the Infinium MethylationEPIC BeadChip array and identified 19 potential DNAm age predictors for bone samples. In addition, this study investigated five predictors for blood as well as two new markers (*TMEM51* and *EPHA6*). These, plus two previously known blood DNAm markers from *TRIM59* and *ELOVL2*, were selected as the most promising candidates for epigenetic age estimation in bones [197]. The VISAGE Enhanced Tool for Age estimation from somatic tissues comes with a prediction model for bones that includes six CpGs from *ELOVL2, KLF14, PDE4C* and *ASPA* and predicted age in a test set of 49 bone samples with an MAE = 3.4 years [15] (for more information on the VISAGE Enhanced Tool for Age from somatic tissues, see below). Hao et al. [198] reported a model that included 10 CpG sites from the *LAG3*, *SCGN, ELOVL2, KLF14, C1orf132, SLC12A5, GRIA2*, and *PDE4C* genes, which predicted age from hair with median absolute deviation MAD = 3.68 years. Recently, Fokias et al. [199] reported epigenetic age prediction from finger- and toenails based on pyrosequencing with MAE of 5.48–9.36. For a summary of MPS-based epigenetic age prediction tools from somatic tissues, see Table 3.

4.4. Epigenetic age estimation from DNA of semen

Semen traces form the most important crime scene evidence in sexual assault cases. However, compared to somatic cells, DNA methylation patterns in sperm cells are completely different [200]. Not surprisingly therefore, the universal age predictor developed by Horvath [172] lacked a significant age correlation in sperm and it predicted age from sperm with a significantly lower estimated age than the true age of the sperm donors. The first model specifically developed for predicting age from sperm DNA achieved an MAE = 4.2 years using an assay based on SNaPshot technology [201]. It was subjected to forensic validation in a study involving two small sets of samples, 12 male donors aged 24–57 years and 19 forensic casework samples including semen stains from rape and sexual assault cases. The MAE achieved in this independent validation was 4.8 years for the donors and 5.2 years for forensic casework samples. The study attempted to estimate the sensitivity of the method by analyzing the approximate amounts of bisulfite converted DNA after the initial conversion of 100 ng of genomic DNA. The study indicated the sensitivity at the level of a minimum 5 ng of bisulfite converted DNA [150]. Two of the markers selected, cg06979108 in the *NOX4* gene and cg12837463 (no gene identified), were further validated in a small sample set from China [202]. In a large study, Jenkins et al. investigated 329 donors using the Infinium HumanMethylation450 BeadChip arrays and selected 51 differentially methylated regions that included 261 CpGs which were significantly correlated with age. Based on these data, the authors developed a predictive model that relied on mean DNA methylation of multiple CpGs in these 51 genomic regions. The assessed MAE was as low as 2.37 years [200]. Because of the limitation of DNAm markers for age prediction from semen, the VISAGE Consortium has conducted a discovery analysis for age-correlated CpGs in semen using Illumina's InfiniumEPIC BeadChip array and revealed new potential age predictors suitable for age prediction in sperm cells [163]. This study also involved validation of the newly discovered markers together with the three predictors reported by Lee et al. [201] and development of an epigenetic age model for semen, which was validated. The final VISAGE Enhanced Tool for Age estimation from semen comprised four novel and one previously known DNAm markers and predicted age with an MAE = 5.1 years [163]. The VISAGE Consortium developed two MPS assays that enable collection of DNA methylation data for thirteen markers initially considered for age estimation in semen and six markers from five genes finally included in the model. Both assays have proved to be suitable for data collection for the purpose of age estimation from semen in forensic investigations [163]. The prediction model for age estimation in semen developed by the

VISAGE Consortium is implemented in the VISAGE Software. For MPS-based epigenetic age prediction tools from semen, also see Table 3.

4.5. Towards universal age estimation tools in forensics

The idea of a universal set of DNAm markers that can predict age across forensically relevant tissues appears to be a practical goal for forensics but is challenged by the increase of tissue-related methylation differences with decreasing number of CpGs in the model, as discussed above. Thus far, this universal approach in forensically motivated age prediction has been tested primarily for the simultaneous age prediction of different somatic tissues, because DNAm patterns largely overlap between somatic tissues in contrast to semen. Prior to the study by Naue et al. [163] already mentioned above, Alghanim et al. [203] investigated 27 CpG sites from *SCGN*, *DLX5* and *KLF14* in blood and saliva, and validated CpGs from *SCGN* and *KLF14* as age predictors for blood and saliva, proposing several predictive models that were able to predict age with an MAE = 7.1 – 10.3 in independent testing datasets. It has been shown that five DNAm markers initially validated for age estimation in blood from *ELOVL2, FHL2, KLF14, MIR29B2CHG* and *TRIM59*, respectively, were able to predict age in saliva and buccal cells with similar accuracies. In this study, CpG markers from *ELOVL2, KLF14* and *TRIM59* were strongly correlated with age in all three body fluids. Two different prediction modelling approaches were tested in this study. The first included three separate models based on data established in blood, saliva, and buccal swabs that predicted the age with an MAE = 3.478, 3.552, and 4.293 years in blood, saliva, and buccal swabs, respectively. The second was based on a single model trained with data obtained from blood, saliva and buccal cell samples and gave a predicted age with a slightly larger MAE = 3.8 years in a testing set that also included data from all three types of samples [204]. The markers selected by Vidaki et al. [205] were used by Aliferi et al. [151] for developing two multiplex MPS assays optimized for forensic use. The study evaluated various machine learning approaches to reach the most accurate prediction algorithm based on the twelve predictors. The best-performing predictive algorithm using the support vector machine approach provided a prediction accuracy of MAE = 4.1 years in blood. The method also showed potential usefulness for predicting age in saliva [151]. An independent validation of the age prediction models for blood and buccal cells based on the CpG markers from *ASPA, EDARADD, PDE4C* and *ELOVL2* [194] confirmed the expected performance of both models, but showed lower accuracy parameters compared to the original studies, which the authors explained were due to technical problems and the need to recalibrate the predictive models [206].

The available literature supports the idea that a relatively small number of DNAm markers can predict age in various somatic tissues. As mentioned in the previous sections, this was achieved with the VISAGE Enhanced Tool for Age from somatic tissues, which consists of one MPS tool for 44 CpGs from 8 genes and three tissue-specific prediction models for blood, buccal cells, and bones, respectively, based on subsets of markers included in the MPS tool. The study showed that while a small number of CpG predictors is sufficient to estimate age in different somatic tissues, accurate prediction requires the use of separate models for each tissue based on tissue-specific model data [15]. The three models were developed based on training sets of 112 DNA samples obtained from blood, buccal cells and bones, respectively and validated in independent sets of 48 blood and buccal, and 49 bone samples, respectively. The three models, as well as the separate model for semen used by the VISAGE Enhanced Tool for Age estimation from semen [163], are included as prediction tools in the VISAGE Software. The VISAGE Consortium also tested the potential applicability of these models for age prediction in tissues selected from 24 deceased individuals that included blood, cartilage, and muscle samples. As may be expected, only age prediction from blood provided satisfactory outcomes (MAE=3.1). The study concluded that the markers may be suitable for age estimation in the two remaining tissues but after remodeling using data from such tissue, which were not available in this study [15]. Age prediction in different tissues using a single small set of CpG markers is also being explored by others [207].

4.6. Other issues with epigenetic age estimation in forensics

The forensic community has recognized the need for standardization in DNAm-based age estimation, as emphasized by Naue and Lee (2018), particularly regarding DNA extraction protocols, which arise from differences in the cell composition of body fluids. It is important to ensure that measures are in place to properly check the quantity of DNA before and after bisulfite conversion, to control the efficiency of this conversion process, the PCR and DNAm quantification, due to known variation in the performance of various kits and assays and in the interpretation and reporting of DNAm analysis results [208]. Appropriate methods are currently lacking but expected to become available soon.

Ethical and legal issues relating to epigenetic age estimation are also important. It has been noted that although age is not considered to constitute sensitive personal information, the use of DNA methylation markers for age estimation may reveal a wider range of information regarding a donor's health [209]. There is also the important discussion issue whether age is considered an externally visible characteristic or not, which is relevant in countries where the legalization of FDP is based on laws focusing on externally visible characteristics and where age is not particularly mentioned in legal frameworks, such as in the Netherlands; contrasting with Germany, where age estimation is specifically mentioned in the FDP legislation.

Finally, it is worth noting that applications of epigenetic age estimation in forensics goes beyond crime scene investigation within the concept of FDP as discussed above. Outside FDP, it has the potential to be used to obtain age from analyzing biological reference samples of known living persons for whom age is uncertain, for instance in legal cases involving persons lacking identification documentation to demonstrate age. Age estimation of the living has received increasing attention because of recent refugee migration events, such as those currently taking place in Europe. While various non-molecular approaches have been used for age estimation in the living [210], the application of epigenetic age prediction has also been suggested for this purpose but was regarded by some authors not to be ready for practical use [210]. While epigenetic age prediction from crime scene samples is challenged by the compromised DNA often obtained from crime scene traces, these issues do not apply when predicting epigenetic age from living persons as DNA from reference material is of high-enough quality and quantity to allow reliable DNAm microarray analysis. Issues of epigenetic age prediction in the living in court are mainly linked to the very high expectations for the prediction accuracy of such analyses. The major forensic application of epigenetic age prediction in the living is to differentiate if a perpetrator is adolescent or adult because there are different laws pertaining to these two age groups. Because the legal threshold is a specific age, for instance 18 or 21 years depending on the country, in order to be suitable, epigenetic age prediction for such legal applications has to be extremely accurate to provide useful information. Currently, available epigenetic age tests and models have not been properly validated to achieve such high accuracy. For instance, a model based on 391 CpGs available from Infinium MethylationEPIC microarray analysis achieved an MAE= 2 years in buccal cells [211]. More research is needed to better understand the age prediction error and the causes of predictive outliers that do not match age estimation expectations. Moreover, it is important to investigate the significance of inter-population epigenetic variation as well as the impact of environmental factors and other confounding factors on the accuracy of epigenetic age prediction [212]. Combining DNAm markers with other molecular approaches such as D-aspartic acid and the accumulation of pentosidine [213], or with non-molecular approaches such as skeletal or dental X-ray measurements [210], potentially allows further reduction of age estimation inaccuracy in the living in the future.

Declaration of Competing Interest

All authors were part of the EU Horizon2020-funded VISAGE Consortium and Project (http://www.visage-h2020.eu) that worked on improving, integrating, implementing, disseminating, and socially and ethically assessing Forensic DNA Phenotyping on appearance, biogeographic ancestry and age. The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. MK is a co-inventor of the patent EP2195448A1 "Method to predict iris color" but has not received license fees or royalties from this.

Acknowledgements

The authors thank all their colleagues from the VISAGE Consortium for their enthusiastic work that led to successfully fulfilling all objectives of the VISAGE Project within the project time 2017–2021 (http://www.visage-h2020.eu). The authors' work on DNA-based appearance, ancestry and age prediction was supported in part by the European Union's Horizon 2020 Research and Innovation programme under grant agreement No. 740580 within the framework of the Visible Attributes through Genomics (VISAGE) Project and Consortium. WB's work was additionally supported by the Polish National Centre for Research and Development [grant number DOB-BIO7/17/01/2015].

References

[1] M. Kayser, P. de Knijff, Improving human forensics through advances in genetics, genomics and molecular biology, Nat. Rev. Genet. 12 (3) (2011) 179–192.
[2] P.M. Schneider, B. Prainsack, M. Kayser, The use of forensic DNA phenotyping in predicting appearance and biogeographic ancestry, Dtsch. Arztebl. Int. 51–52 (51–52) (2019) 873–880.
[3] G. Samuel, B. Prainsack, 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.
[4] M. Kayser, Forensic use of Y-chromosome DNA: a general overview, Hum. Genet 136 (5) (2017) 621–635.
[5] Milicia van Doorn. [cited 2023; Available from: ⟨https://nltimes.nl/2018/12/11/twenty-years-prison-1992-zaandam-murder⟩.
[6] T.D. Albright, Why eyewitnesses fail, Proc. Natl. Acad. Sci. USA 114 (30) (2017) 7758–7764.
[7] M. Kayser, Forensic DNA phenotyping: predicting human appearance from crime scene material for investigative purposes, Forensic Sci. Int Genet 18 (2015) 33–48.
[8] C. Phillips, Forensic genetic analysis of bio-geographical ancestry, Forensic Sci. Int Genet 18 (2015) 49–65.
[9] H.Y. Lee, S.D. Lee, K.J. Shin, Forensic DNA methylation profiling from evidence material for investigative leads, BMB Rep. 49 (7) (2016) 359–369.
[10] A. Freire-Aradas, C. Phillips, M.V. Lareu, Forensic individual age estimation with DNA: from initial approaches to methylation tests, Forensic Sci. Rev. 29 (2) (2017) 121–144.
[11] D. Ballard, J. Winkler-Galicki, J. Wesoly, Massive parallel sequencing in forensics: advantages, issues, technicalities, and prospects, Int J. Leg. Med 134 (4) (2020) 1291–1303.
[12] S. Walsh, et al., Global skin colour prediction from DNA, Hum. Genet 136 (7) (2017) 847–863.
[13] L. Chaitanya, et al., The HIrisPlex-S system for eye, hair and skin colour prediction from DNA: Introduction and forensic developmental validation, Forensic Sci. Int Genet 35 (2018) 123–135.
[14] C. Xavier, et al., Development and inter-laboratory evaluation of the VISAGE Enhanced Tool for Appearance and Ancestry inference from DNA, Forensic Sci. Int Genet 61 (2022), 102779.
[15] A. Wozniak, et al., Development of the VISAGE enhanced tool and statistical models for epigenetic age estimation in blood, buccal cells and bones, Aging (Albany NY) 13 (5) (2021) 6459–6484.
[16] A. Heidegger, et al., Development and inter-laboratory validation of the VISAGE enhanced tool for age estimation from semen using quantitative DNA methylation analysis, Forensic Sci. Int Genet 56 (2022), 102596.
[17] S. Walsh, et al., IrisPlex: a sensitive DNA tool for accurate prediction of blue and brown eye colour in the absence of ancestry information, Forensic Sci. Int Genet 5 (3) (2011) 170–180.
[18] S. Walsh, et al., Developmental validation of the HIrisPlex system: DNA-based eye and hair colour prediction for forensic and anthropological usage, Forensic Sci. Int Genet 9 (2014) 150–161.
[19] S. Walsh, et al., The HIrisPlex system for simultaneous prediction of hair and eye colour from DNA, Forensic Sci. Int Genet 7 (1) (2013) 98–115.
[20] S. Rauf, et al., Unveiling forensically relevant biogeographic, phenotype and Y-chromosome SNP variation in Pakistani ethnic groups using a customized hybridisation enrichment forensic intelligence panel, PLoS One 17 (2) (2022), e0264125.
[21] A.C. Jager, et al., Developmental validation of the MiSeq FGx forensic genomics system for targeted next generation sequencing in forensic DNA casework and database laboratories, Forensic Sci. Int Genet 28 (2017) 52–70.
[22] P.G. Hysi, et al., Genome-wide association meta-analysis of individuals of European ancestry identifies new loci explaining a substantial fraction of hair color variation and heritability, Nat. Genet 50 (5) (2018) 652–656.
[23] M. Simcoe, et al., Genome-wide association study in almost 195,000 individuals identifies 50 previously unidentified genetic loci for eye color, Sci. Adv. 7 (2021) 11.
[24] M. Kukla-Bartoszek, et al., Searching for improvements in predicting human eye colour from DNA, Int J. Leg. Med. 135 (6) (2021) 2175–2187.
[25] O. Maronas, et al., Development of a forensic skin colour predictive test, Forensic Sci. Int. Genet. 13 (2014) 34–44.
[26] K. Breslin, 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 Sci. Int. Genet. 43 (2019), 102152.
[27] O. Bulbul, G. Filoglu, Development of a SNP panel for predicting biogeographical ancestry and phenotype using massively parallel sequencing, Electrophoresis 39 (21) (2018) 2743–2751.
[28] C. Xavier, 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.
[29] C. Xavier, et al., Evaluation of the VISAGE basic tool for appearance and ancestry inference using ForenSeq(R) chemistry on the MiSeq FGx(R) system, Forensic Sci. Int Genet 58 (2022), 102675.
[30] L. Palencia-Madrid, et al., Evaluation of the VISAGE basic tool for appearance and ancestry prediction using powerseq chemistry on the MiSeq FGx system, Genes (Basel) 11 (6) (2020).
[31] A. Tillmar, et al., The FORCE Panel: an all-in-one SNP marker set for confirming investigative genetic genealogy leads and for general forensic applications, Genes (Basel) (2021) **12**.
[32] A. Visconti, et al., Genome-wide association study in 176,678 Europeans reveals genetic loci for tanning response to sun exposure, Nat. Commun. 9 (1) (2018) 1684.
[33] N.G. Crawford, et al., Loci associated with skin pigmentation identified in African populations, Science 358 (2017) 6365.
[34] M. Jonnalagadda, et al., A genome-wide association study of skin and iris pigmentation among individuals of south Asian ancestry, Genome Biol. Evol. 11 (4) (2019) 1066–1076.
[35] F. Peng, et al., Genome-wide association studies identify multiple genetic loci influencing eyebrow color variation in Europeans, J. Invest Dermatol. 139 (7) (2019) 1601–1605.
[36] B. Hernando, et al., Genetic determinants of freckle occurrence in the Spanish population: Towards ephelides prediction from human DNA samples, Forensic Sci. Int Genet 33 (2018) 38–47.
[37] M. Kukla-Bartoszek, et al., DNA-based predictive models for the presence of freckles, Forensic Sci. Int Genet 42 (2019) 252–259.
[38] K. Adhikari, et al., A genome-wide association scan in admixed Latin Americans identifies loci influencing facial and scalp hair features, Nat. Commun. 7 (2016) 10815.
[39] E. Pospiech, et al., Exploring the possibility of predicting human head hair greying from DNA using whole-exome and targeted NGS data, BMC Genom. 21 (1) (2020) 538.
[40] E. Pospiech, et al., Evaluation of the predictive capacity of DNA variants associated with straight hair in Europeans, Forensic Sci. Int Genet 19 (2015) 280–288.
[41] F. Liu, et al., Meta-analysis of genome-wide association studies identifies 8 novel loci involved in shape variation of human head hair, Hum. Mol. Genet 27 (3) (2018) 559–575.
[42] E. Pospiech, et al., Towards broadening Forensic DNA Phenotyping beyond pigmentation: Improving the prediction of head hair shape from DNA, Forensic Sci. Int Genet 37 (2018) 241–251.
[43] M. Marcinska, et al., Evaluation of DNA variants associated with androgenetic alopecia and their potential to predict male pattern baldness, PLoS One 10 (5) (2015), e0127852.
[44] F. Liu, et al., Prediction of male-pattern baldness from genotypes, Eur. J. Hum. Genet 24 (6) (2016) 895–902.
[45] S.P. Hagenaars, et al., Genetic prediction of male pattern baldness, PLoS Genet 13 (2) (2017), e1006594.
[46] Y. Chen, et al., Genetic prediction of male pattern baldness based on large independent datasets, Eur. J. Hum. Genet (2022).
[47] C.X. Yap, et al., Dissection of genetic variation and evidence for pleiotropy in male pattern baldness, Nat. Commun. 9 (1) (2018) 5407.
[48] F. Liu, et al., Common DNA variants predict tall stature in Europeans, Hum. Genet 133 (5) (2014) 587–597.
[49] F. Liu, et al., Update on the predictability of tall stature from DNA markers in Europeans, Forensic Sci. Int Genet 42 (2019) 8–13.
[50] A.R. Wood, et al., Defining the role of common variation in the genomic and biological architecture of adult human height, Nat. Genet 46 (11) (2014) 1173–1186.
[51] L. Lello, et al., Accurate Genomic Prediction of Human Height, Genetics 210 (2) (2018) 477–497.
[52] L. Yengo, et al., A saturated map of common genetic variants associated with human height, Nature 610 (7933) (2022) 704–712.

[53] E.M. Gorden, K. Sturk-Andreaggi, C. Marshall, Capture enrichment and massively parallel sequencing for human identification, Forensic Sci. Int Genet 53 (2021), 102496.
[54] K. Adhikari, et al., A genome-wide association study identifies multiple loci for variation in human ear morphology, Nat. Commun. 6 (2015) 7500.
[55] J.R. Shaffer, et al., Multiethnic GWAS reveals polygenic architecture of earlobe attachment, Am. J. Hum. Genet 101 (6) (2017) 913–924.
[56] S. Wu, et al., Genome-wide association studies and CRISPR/Cas9-mediated gene editing identify regulatory variants influencing eyebrow thickness in humans, PLoS Genet 14 (9) (2018), e1007640.
[57] E. Pospiech, et al., Overlapping association signals in the genetics of hair-related phenotypes in humans and their relevance to predictive DNA analysis, Forensic Sci. Int Genet 59 (2022), 102693.
[58] F. Liu, et al., A genome-wide association study identifies five loci influencing facial morphology in Europeans, PLoS Genet 8 (9) (2012), e1002932.
[59] L. Paternoster, et al., Genome-wide association study of three-dimensional facial morphology identifies a variant in PAX3 associated with nasion position, Am. J. Hum. Genet 90 (3) (2012) 478–485.
[60] K. Adhikari, et al., A genome-wide association scan implicates DCHS2, RUNX2, GLI3, PAX1 and EDAR in human facial variation, Nat. Commun. 7 (2016) 11616.
[61] J.B. Cole, et al., Genomewide association study of African children identifies association of SCHIP1 and PDE8A with facial size and shape, PLoS Genet 12 (8) (2016), e1006174.
[62] J.R. Shaffer, et al., Genome-wide association study reveals multiple loci influencing normal human facial morphology, PLoS Genet 12 (8) (2016), e1006149.
[63] P. Claes, et al., Genome-wide mapping of global-to-local genetic effects on human facial shape, Nat. Genet 50 (3) (2018) 414–423.
[64] J.D. White, et al., Insights into the genetic architecture of the human face, Nat. Genet 53 (1) (2021) 45–53.
[65] Z. Xiong, et al., Novel genetic loci affecting facial shape variation in humans, Elife (2019) 8.
[66] B. Bonfante, et al., A GWAS in Latin Americans identifies novel face shape loci, implicating VPS13B and a Denisovan introgressed region in facial variation, Sci. Adv. 7 (2021) 6.
[67] M. Zhang, et al., Genetic variants underlying differences in facial morphology in East Asian and European populations, Nat. Genet 54 (4) (2022) 403–411.
[68] Z. Xiong, et al., Combining genome-wide association studies highlight novel loci involved in human facial variation, Nat. Commun. 13 (1) (2022) 7832.
[69] P. Claes, et al., Modeling 3D facial shape from DNA, PLoS Genet 10 (3) (2014), e1004224.
[70] J. Fagertun, et al., Predicting facial characteristics from complex polygenic variations, Forensic Sci. Int Genet 19 (2015) 263–268.
[71] C. Lippert, et al., Identification of individuals by trait prediction using whole-genome sequencing data, Proc. Natl. Acad. Sci. USA 114 (38) (2017) 10166–10171.
[72] B. Hallgrimsson, et al., Let's face it–complex traits are just not that simple, PLoS Genet 10 (11) (2014), e1004724.
[73] C. Arnold, The controversial company using DNA to sketch the faces of criminals, Nature 585 (7824) (2020) 178–181.
[74] A. Vidaki, M. Kayser, From forensic epigenetics to forensic epigenomics: broadening DNA investigative intelligence, Genome Biol. 18 (1) (2017) 238.
[75] S.C.E. Maas, et al., Validated inference of smoking habits from blood with a finite DNA methylation marker set, Eur. J. Epidemiol. 34 (11) (2019) 1055–1074.
[76] S.C.E. Maas, et al., Validating biomarkers and models for epigenetic inference of alcohol consumption from blood, Clin. Epigenet. 13 (1) (2021) 198.
[77] A. Ralf, et al., Forensic Y-SNP analysis beyond SNaPshot: high-resolution Y-chromosomal haplogrouping from low quality and quantity DNA using Ion AmpliSeq and targeted massively parallel sequencing, Forensic Sci. Int Genet 41 (2019) 93–106.
[78] L. Chaitanya, et al., Simultaneous whole mitochondrial genome sequencing with short overlapping amplicons suitable for degraded DNA using the ion torrent personal genome machine, Hum. Mutat. 36 (12) (2015) 1236–1247.
[79] C. Phillips, et al., Global patterns of STR sequence variation: sequencing the CEPH human genome diversity panel for 58 forensic STRs using the Illumina ForenSeq DNA Signature Prep Kit, Electrophoresis 39 (21) (2018) 2708–2724.
[80] A.J. Pakstis, et al., Increasing the reference populations for the 55 AISNP panel: the need and benefits, Int J. Leg. Med. 131 (4) (2017) 913–917.
[81] A.J. Pakstis, et al., Genetic relationships of European, Mediterranean, and SW Asian populations using a panel of 55 AISNPs, Eur. J. Hum. Genet 27 (12) (2019) 1885–1893.
[82] C. Phillips, et al., A compilation of tri-allelic SNPs from 1000 Genomes and use of the most polymorphic loci for a large-scale human identification panel, Forensic Sci. Int. Genet. 46 (2020), 102232.
[83] J.D. Churchill, et al., Population and performance analyses of four major populations with Illumina's FGx Forensic Genomics System, Forensic Sci. Int Genet 30 (2017) 81–92.
[84] V. Sharma, et al., Evaluation of ForenSeq Signature Prep Kit B on predicting eye and hair coloration as well as biogeographical ancestry by using Universal Analysis Software (UAS) and available web-tools, Electrophoresis 40 (9) (2019) 1353–1364.
[85] H. Rajeevan, et al., FrogAncestryCalc: a standalone batch likelihood computation tool for ancestry inference panels catalogued in FROG-kb, Forensic Sci. Int. Genet. 46 (2020), 102237.
[86] A.J. Pakstis, et al., Population relationships based on 170 ancestry SNPs from the combined Kidd and Seldin panels, Sci. Rep. 9 (1) (2019) 18874.
[87] C. Phillips, et al., Building a forensic ancestry panel from the ground up: The EUROFORGEN Global AIM-SNP set, Forensic Sci. Int. Genet. 11 (2014) 13–25.
[88] C. Phillips, et al., MAPlex - a massively parallel sequencing ancestry analysis multiplex for Asia-Pacific populations, Forensic Sci. Int Genet 42 (2019) 213–226.
[89] V. Pereira, et al., Development and validation of the EUROFORGEN NAME (North African and Middle Eastern) ancestry panel, Forensic Sci. Int. Genet. 42 (2019) 260–267.
[90] M. Diepenbroek, et al., Evaluation of the Ion AmpliSeq PhenoTrivium Panel: MPS-Based Assay for Ancestry and Phenotype Predictions Challenged by Casework Samples, Genes (Basel) 11 (12) (2020).
[91] M. de la Puente, et al., Development and evaluation of the ancestry informative marker panel of the VISAGE basic tool, Genes (Basel) 12 (8) (2021).
[92] J. Ruiz-Ramírez, M.d.l. Puente, A. Xavier, J. Ambroa-Conde, A. Álvarez-Dios, A. Freire-Aradas, A. Mosquera-Miguel, A. Ralf, et al., Development and evaluations of the ancestry informative markers of the VISAGE Enhanced Tool for Appearance and Ancestry, Forensic Sci. Int.: Genet. 64 (2023) 102853.
[93] O. Bulbul, et al., Improving ancestry distinctions among Southwest Asian populations, Forensic Sci. Int Genet 35 (2018) 14–20.
[94] C.X. Li, et al., A panel of 74 AISNPs: Improved ancestry inference within Eastern Asia, Forensic Sci. Int Genet 23 (2016) 101–110.
[95] J.Y. Jung, et al., Ancestry informative markers (AIMs) for Korean and other East Asian and South East Asian populations, Int J. Leg. Med. 133 (6) (2019) 1711–1719.
[96] I. Yuasa, et al., Japaneseplex: A forensic SNP assay for identification of Japanese people using Japanese-specific alleles, Leg. Med. (Tokyo) 33 (2018) 17–22.
[97] A. Bergstrom, et al., Insights into human genetic variation and population history from 929 diverse genomes, Science 367 (2020) 6484.
[98] C. Phillips, et al., Eurasiaplex: a forensic SNP assay for differentiating European and South Asian ancestries, Forensic Sci. Int. Genet. 7 (3) (2013) 359–366.
[99] J.M. Galanter, et al., Development of a panel of genome-wide ancestry informative markers to study admixture throughout the Americas, PLoS Genet 8 (3) (2012), e1002554.
[100] C. Genomes Project, et al., A global reference for human genetic variation, Nature 526 (7571) (2015) 68–74.
[101] P. Resutik, S.A. M. Krützen, A. Kratzer, C. Haas, C. Phillips, N. Arora, Comparative evaluation of the MAPlex, Precision ID Ancestry Panel, and VISAGE Basic Tool for biogeographical ancestry inference, Forensic Sci. Int.: Genet. (2023) (submitted).
[102] C. Phillips, et al., Tetra-allelic SNPs: Informative forensic markers compiled from public whole-genome sequence data, Forensic Sci. Int. Genet. 19 (2015) 100–106.
[103] K.K. Kidd, et al., Evaluating 130 microhaplotypes across a global set of 83 populations, Forensic Sci. Int. Genet. 29 (2017) 29–37.
[104] O. Bulbul, et al., Ancestry inference of 96 population samples using microhaplotypes, Int J. Leg. Med. 132 (3) (2018) 703–711.
[105] F. Oldoni, K.K. Kidd, D. Podini, Microhaplotypes in forensic genetics, Forensic Sci. Int. Genet. 38 (2019) 54–69.
[106] P. Chen, et al., Identifying novel microhaplotypes for ancestry inference, Int. J. Leg. Med 133 (4) (2019) 983–988.
[107] E.Y.Y. Cheung, et al., Performance of ancestry-informative SNP and microhaplotype markers, Forensic Sci. Int. Genet. 43 (2019), 102141.
[108] M. de la Puente, et al., Building a custom large-scale panel of novel microhaplotypes for forensic identification using MiSeq and Ion S5 massively parallel sequencing systems, Forensic Sci. Int. Genet. 45 (2020), 102213.
[109] M. de la Puente, et al., Broadening the applicability of a custom multi-platform panel of microhaplotypes: bio-geographical ancestry inference and expanded reference data, Front Genet 11 (2020), 581041.
[110] U. Soundararajan, et al., Minimal SNP overlap among multiple panels of ancestry informative markers argues for more international collaboration, Forensic Sci. Int Genet 23 (2016) 25–32.
[111] An integrated map of genetic variation from 1,092 human genomes, Nature 491 (7422) (2012) 56–65.
[112] S. Mallick, et al., The Simons Genome Diversity Project: 300 genomes from 142 diverse populations, Nature 538 (7624) (2016) 201–206.
[113] L. Pagani, et al., Genomic analyses inform on migration events during the peopling of Eurasia, Nature 538 (7624) (2016) 238–242.
[114] C. Phillips, D.M. M. de la Puente, J. Amigo, M.V. Lareu, Using online whole-genome variation data for forensic ancestry inference tests, in: A.B. Elena Pilli (Ed.), Forensic DNA Analysis: Technological Development and Innovative, CRC Press, 2021.
[115] S. Gudmundsson, et al., Variant interpretation using population databases: Lessons from gnomAD, Hum. Mutat. 43 (8) (2022) 1012–1030.
[116] M. Byrska-Bishop, et al., High-coverage whole-genome sequencing of the expanded 1000 Genomes Project cohort including 602 trios, Cell 185 (18) (2022) 3426–3440, e19.
[117] S. Zhao, et al., AIM-SNPtag: a computationally efficient approach for developing ancestry-informative SNP panels, Forensic Sci. Int Genet 38 (2019) 245–253.
[118] C. Santos, et al., Inference of ancestry in forensic analysis II: analysis of genetic data, Methods Mol. Biol. 1420 (2016) 255–285.
[119] E.Y.Y. Cheung, M.E. Gahan, D. McNevin, Prediction of biogeographical ancestry from genotype: a comparison of classifiers, Int. J. Leg. Med. 131 (4) (2017) 901–912.
[120] J.K. Pritchard, M. Stephens, P. Donnelly, Inference of population structure using multilocus genotype data, Genetics 155 (2) (2000) 945–959.
[121] L. Porras-Hurtado, et al., An overview of STRUCTURE: applications, parameter settings, and supporting software, Front Genet 4 (2013) 98.
[122] D.H. Alexander, J. Novembre, K. Lange, Fast model-based estimation of ancestry in unrelated individuals, Genome Res 19 (9) (2009) 1655–1664.

[123] M.A. Almarri, et al., The genomic history of the Middle East, Cell 184 (18) (2021) 4612–4625, e14.
[124] S. Jin, et al., Implementing a biogeographic ancestry inference service for forensic casework, Electrophoresis 39 (21) (2018) 2757–2765.
[125] M. Al-Asfi, et al., Assessment of the precision ID ancestry panel, Int J. Leg. Med 132 (6) (2018) 1581–1594.
[126] T. Tvedebrink, et al., Weight of the evidence of genetic investigations of ancestry informative markers, Theor. Popul Biol. 120 (2018) 1–10.
[127] H.S. Mogensen, et al., Ancestry prediction efficiency of the software GenoGeographer using a z-score method and the ancestry informative markers in the Precision ID Ancestry Panel, Forensic Sci. Int Genet 44 (2020), 102154.
[128] C. Santos, et al., Forensic ancestry analysis with two capillary electrophoresis ancestry informative marker (AIM) panels: Results of a collaborative EDNAP exercise, Forensic Sci. Int Genet 19 (2015) 56–67.
[129] J.C. Gower, Some distance properties of latent root and vector methods used in multivariate analysis, Biometrika 53 (1966) 325–338.
[130] V. Pereira, et al., Evaluation of the precision of ancestry inferences in south american admixed populations, Front Genet 11 (2020) 966.
[131] C. Strobl, et al., Evaluation of the precision ID whole MtDNA genome panel for forensic analyses, Forensic Sci. Int Genet 35 (2018) 21–25.
[132] D. Kling, et al., Investigative genetic genealogy: current methods, knowledge and practice, Forensic Sci. Int Genet 52 (2021), 102474.
[133] J.H. de Vries, et al., Impact of SNP microarray analysis of compromised DNA on kinship classification success in the context of investigative genetic genealogy, Forensic Sci. Int Genet 56 (2022), 102625.
[134] A. Tillmar, et al., Getting the conclusive lead with investigative genetic genealogy - a successful case study of a 16 year old double murder in Sweden, Forensic Sci. Int Genet 53 (2021), 102525.
[135] C.M. Koch, W. Wagner, Epigenetic-aging-signature to determine age in different tissues, Aging (Albany NY) 3 (10) (2011) 1018–1027.
[136] S. Bocklandt, et al., Epigenetic predictor of age, PLoS One 6 (6) (2011), e14821.
[137] S.E. Jung, K.J. Shin, H.Y. Lee, DNA methylation-based age prediction from various tissues and body fluids, BMB Rep. 50 (11) (2017) 546–553.
[138] S. Cho, et al., Independent validation of DNA-based approaches for age prediction in blood, Forensic Sci. Int Genet 29 (2017) 250–256.
[139] A. Freire-Aradas, et al., Tracking age-correlated DNA methylation markers in the young, Forensic Sci. Int Genet 36 (2018) 50–59.
[140] L.M. McEwen, et al., The PedBE clock accurately estimates DNA methylation age in pediatric buccal cells, Proc. Natl. Acad. Sci. USA 117 (38) (2020) 23329–23335.
[141] A. Freire-Aradas, et al., A common epigenetic clock from childhood to old age, Forensic Sci. Int Genet 60 (2022), 102743.
[142] B.E. Koop, et al., Epigenetic clocks may come out of rhythm-implications for the estimation of chronological age in forensic casework, Int J. Leg. Med 134 (6) (2020) 2215–2228.
[143] M. Spolnicka, et al., DNA methylation in ELOVL2 and C1orf132 correctly predicted chronological age of individuals from three disease groups, Int J. Leg. Med 132 (1) (2018) 1–11.
[144] M. Spolnicka, et al., DNA methylation signature in blood does not predict calendar age in patients with chronic lymphocytic leukemia but may alert to the presence of disease, Forensic Sci. Int Genet 34 (2018) e15–e17.
[145] M. Spolnicka, et al., Modified aging of elite athletes revealed by analysis of epigenetic age markers, Aging (Albany NY) 10 (2) (2018) 241–252.
[146] R. Zbiec-Piekarska, et al., Examination of DNA methylation status of the ELOVL2 marker may be useful for human age prediction in forensic science, Forensic Sci. Int Genet 14 (2015) 161–167.
[147] F. Mayer, et al., Altered DNA methylation at age-associated CpG sites in children with growth disorders: impact on age estimation? Int J. Leg. Med 136 (4) (2022) 987–996.
[148] D. Piniewska-Rog, et al., Impact of excessive alcohol abuse on age prediction using the VISAGE enhanced tool for epigenetic age estimation in blood, Int J. Leg. Med 135 (6) (2021) 2209–2219.
[149] J. Foox, et al., Epigenetic forensics for suspect identification and age prediction, Forensic Genom. 1 (3) (2021) 83–86.
[150] J.W. Lee, et al., A validation study of DNA methylation-based age prediction using semen in forensic casework samples, Leg. Med (Tokyo) 31 (2018) 74–77.
[151] A. Aliferi, et al., DNA methylation-based age prediction using massively parallel sequencing data and multiple machine learning models, Forensic Sci. Int Genet 37 (2018) 215–226.
[152] A. Aliferi, et al., Combining current knowledge on DNA methylation-based age estimation towards the development of a superior forensic DNA intelligence tool, Forensic Sci. Int Genet 57 (2022), 102637.
[153] J. Naue, et al., Forensic DNA methylation profiling from minimal traces: How low can we go? Forensic Sci. Int Genet 33 (2018) 17–23.
[154] R.W. Smith, C. Monroe, D.A. Bolnick, Detection of Cytosine methylation in ancient DNA from five native american populations using bisulfite sequencing, PLoS One 10 (5) (2015), e0125344.
[155] H. Correia Dias, et al., Age estimation based on DNA methylation using blood samples from deceased individuals, J. Forensic Sci. 65 (2) (2020) 465–470.
[156] A. Freire-Aradas, et al., A comparison of forensic age prediction models using data from four DNA methylation technologies, Front Genet 11 (2020) 932.
[157] Y. Takahashi, et al., Age prediction by methylation analysis of small amounts of DNA using locked nucleic acids, J. Forensic Sci. 68 (1) (2023) 267–274.
[158] L. Manco, H.C. Dias, DNA methylation analysis of ELOVL2 gene using droplet digital PCR for age estimation purposes, Forensic Sci. Int 333 (2022), 111206.
[159] A. Heidegger, et al., Development and optimization of the VISAGE basic prototype tool for forensic age estimation, Forensic Sci. Int Genet 48 (2020), 102322.
[160] R. Richards, et al., Evaluation of massively parallel sequencing for forensic DNA methylation profiling, Electrophoresis 39 (21) (2018) 2798–2805.
[161] J. Fleckhaus, P.M. Schneider, Novel multiplex strategy for DNA methylation-based age prediction from small amounts of DNA via Pyrosequencing, Forensic Sci. Int Genet 44 (2020), 102189.
[162] S.R. Hong, et al., Platform-independent models for age prediction using DNA methylation data, Forensic Sci. Int Genet 38 (2019) 39–47.
[163] A. Pisarek, et al., Epigenetic age prediction in semen - marker selection and model development, Aging (Albany NY) 13 (15) (2021) 19145–19164.
[164] L. Feng, et al., Systematic feature selection improves accuracy of methylation-based forensic age estimation in Han Chinese males, Forensic Sci. Int Genet 35 (2018) 38–45.
[165] F. Peng, et al., Validation of methylation-based forensic age estimation in time-series bloodstains on FTA cards and gauze at room temperature conditions, Forensic Sci. Int Genet 40 (2019) 168–174.
[166] M.H. So, H.Y. Lee, Genetic analyzer-dependent DNA methylation detection and its application to existing age prediction models, Electrophoresis 42 (14–15) (2021) 1497–1506.
[167] J.E. Lee, et al., A collaborative exercise on DNA methylation-based age prediction and body fluid typing, Forensic Sci. Int Genet 57 (2022), 102656.
[168] I. Smeers, et al., Evaluation of three statistical prediction models for forensic age prediction based on DNA methylation, Forensic Sci. Int Genet 34 (2018) 128–133.
[169] C. Pan, et al., The evaluation of seven age-related CpGs for forensic purpose in blood from Chinese Han population, Forensic Sci. Int Genet 46 (2020), 102251.
[170] A. Aliferi, D. Ballard, Predicting chronological age from DNA methylation data: a machine learning approach for small datasets and limited predictors, Methods Mol. Biol. 2432 (2022) 187–200.
[171] M.A. Katsara, et al., Evaluation of supervised machine-learning methods for predicting appearance traits from DNA, Forensic Sci. Int Genet 53 (2021), 102507.
[172] S. Horvath, DNA methylation age of human tissues and cell types, Genome Biol. 14 (10) (2013) R115.
[173] R.C. Slieker, et al., Age-related DNA methylation changes are tissue-specific with ELOVL2 promoter methylation as exception, Epigenet. Chrom. 11 (1) (2018) 25.
[174] J. Naue, et al., Proof of concept study of age-dependent DNA methylation markers across different tissues by massive parallel sequencing, Forensic Sci. Int Genet 36 (2018) 152–159.
[175] P. Garagnani, et al., Methylation of ELOVL2 gene as a new epigenetic marker of age, Aging Cell 11 (6) (2012) 1132–1134.
[176] G. Hannum, et al., Genome-wide methylation profiles reveal quantitative views of human aging rates, Mol. Cell 49 (2) (2013) 359–367.
[177] H. Heyn, et al., DNA methylation contributes to natural human variation, Genome Res 23 (9) (2013) 1363–1372.
[178] R. Zbiec-Piekarska, et al., Development of a forensically useful age prediction method based on DNA methylation analysis, Forensic Sci. Int Genet 17 (2015) 173–179.
[179] A. Montesanto, et al., A new robust epigenetic model for forensic age prediction, J. Forensic Sci. 65 (5) (2020) 1424–1431.
[180] H.C. Dias, et al., DNA methylation age estimation in blood samples of living and deceased individuals using a multiplex SNaPshot assay, Forensic Sci. Int 311 (2020), 110267.
[181] J. Naue, et al., Chronological age prediction based on DNA methylation: Massive parallel sequencing and random forest regression, Forensic Sci. Int Genet 31 (2017) 19–28.
[182] P.Y. Lau, W.K. Fung, Evaluation of marker selection methods and statistical models for chronological age prediction based on DNA methylation, Leg. Med (Tokyo) 47 (2020), 101744.
[183] S.K. Mawlood, et al., Quantification of global mitochondrial DNA methylation levels and inverse correlation with age at two CpG sites, Aging (Albany NY) 8 (4) (2016) 636–641.
[184] Y. Xin, et al., Studies of hTERT DNA methylation assays on the human age prediction, Int J. Leg. Med. 133 (5) (2019) 1333–1339.
[185] A. Vidaki, et al., Male-specific age estimation based on Y-chromosomal DNA methylation, Aging (Albany NY) 13 (5) (2021) 6442–6458.
[186] Y. Hamano, et al., Forensic age prediction for saliva samples using methylation-sensitive high resolution melting: exploratory application for cigarette butts, Sci. Rep. 7 (1) (2017) 10444.
[187] Y. Hamano, et al., Forensic age prediction for dead or living samples by use of methylation-sensitive high resolution melting, Leg. Med (Tokyo) 21 (2016) 5–10.
[188] S.R. Hong, et al., DNA methylation-based age prediction from saliva: high age predictability by combination of 7 CpG markers, Forensic Sci. Int Genet 29 (2017) 118–125.
[189] M. Eipel, et al., Epigenetic age predictions based on buccal swabs are more precise in combination with cell type-specific DNA methylation signatures, Aging (Albany NY) 8 (5) (2016) 1034–1048.
[190] B.E. Koop, et al., Postmortem age estimation via DNA methylation analysis in buccal swabs from corpses in different stages of decomposition-a "proof of principle" study, Int J. Leg. Med 135 (1) (2021) 167–173.
[191] A. Ambroa-Conde, et al., Epigenetic age estimation in saliva and in buccal cells, Forensic Sci. Int Genet 61 (2022), 102770.
[192] J. Becker, et al., Evidence for differences in DNA methylation between Germans and Japanese, Int J. Leg. Med 136 (2) (2022) 405–413.

[193] H. Oka, et al., Some tips on age estimation using DNA methylation in saliva samples as an index across the Japanese and Indonesian ethnicities, Leg. Med (Tokyo) 56 (2022), 102042.
[194] B. Bekaert, et al., Improved age determination of blood and teeth samples using a selected set of DNA methylation markers, Epigenetics 10 (10) (2015) 922–930.
[195] C. Giuliani, et al., Inferring chronological age from DNA methylation patterns of human teeth, Am. J. Phys. Anthr. 159 (4) (2016) 585–595.
[196] A.B. Marquez-Ruiz, et al., DNA methylation levels and telomere length in human teeth: usefulness for age estimation, Int J. Leg. Med 134 (2) (2020) 451–459.
[197] H.Y. Lee, et al., Epigenetic age signatures in bones, Forensic Sci. Int Genet 46 (2020), 102261.
[198] T. Hao, et al., Predicting human age by detecting DNA methylation status in hair, Electrophoresis 42 (11) (2021) 1255–1261.
[199] K. Fokias, et al., Age determination through DNA methylation patterns in fingernails and toenails, Forensic Sci. Int.: Genet. (2023). DOI:https://doi.org/10.1016/j.fsigen.2023.102846.
[200] T.G. Jenkins, et al., Paternal germ line aging: DNA methylation age prediction from human sperm, BMC Genom. 19 (1) (2018) 763.
[201] H.Y. Lee, et al., Epigenetic age signatures in the forensically relevant body fluid of semen: a preliminary study, Forensic Sci. Int Genet 19 (2015) 28–34.
[202] L. Li, et al., Methylation-based age prediction using pyrosequencing platform from seminal stains in Han Chinese males, J. Forensic Sci. 65 (2) (2020) 610–619.
[203] H. Alghanim, et al., Detection and evaluation of DNA methylation markers found at SCGN and KLF14 loci to estimate human age, Forensic Sci. Int Genet 31 (2017) 81–88.
[204] S.E. Jung, et al., DNA methylation of the ELOVL2, FHL2, KLF14, C1orf132/MIR29B2C, and TRIM59 genes for age prediction from blood, saliva, and buccal swab samples, Forensic Sci. Int Genet 38 (2019) 1–8.
[205] A. Vidaki, et al., DNA methylation-based forensic age prediction using artificial neural networks and next generation sequencing, Forensic Sci. Int Genet 28 (2017) 225–236.
[206] M. Pfeifer, et al., Inter-laboratory adaption of age estimation models by DNA methylation analysis-problems and solutions, Int J. Leg. Med 134 (3) (2020) 953–961.
[207] H. Correia Dias, et al., A blood-bone-tooth model for age prediction in forensic contexts, Biology (Basel) 10 (12) (2021).
[208] J. Naue, H.Y. Lee, Considerations for the need of recommendations for the research and publication of DNA methylation results, Forensic Sci. Int Genet 37 (2018) e12–e14.
[209] M. Shabani, et al., Forensic epigenetic age estimation and beyond: ethical and legal considerations, Trends Genet 34 (7) (2018) 489–491.
[210] L. Shi, et al., DNA methylation markers in combination with skeletal and dental ages to improve age estimation in children, Forensic Sci. Int Genet 33 (2018) 1–9.
[211] S. Horvath, et al., Epigenetic clock for skin and blood cells applied to Hutchinson Gilford Progeria Syndrome and ex vivo studies, Aging (Albany NY) 10 (7) (2018) 1758–1775.
[212] C.G. Bell, et al., DNA methylation aging clocks: challenges and recommendations, Genome Biol. 20 (1) (2019) 249.
[213] J. Becker, et al., Age estimation based on different molecular clocks in several tissues and a multivariate approach: an explorative study, Int J. Leg. Med. 134 (2) (2020) 721–733.
[214] F. Liu, et al., Eye color and the prediction of complex phenotypes from genotypes, Curr. Biol. 19 (5) (2009) R192–R193.
[215] W. Branicki, et al., Model-based prediction of human hair color using DNA variants, Hum. Genet. 129 (4) (2011) 443–454.
[216] A. Caliebe, et al., Likelihood ratio and posterior odds in forensic genetics: two sides of the same coin, Forensic Sci. Int. Genet. 28 (2017) 203–210.
[217] M. Eduardoff, et al., Inter-laboratory evaluation of the EUROFORGEN Global ancestry-informative SNP panel by massively parallel sequencing using the Ion PGM, Forensic Sci. Int. Genet. 23 (2016) 178–189.
[218] C. Xavier, et al., Forensic evaluation of the Asia Pacific ancestry-informative MAPlex assay, Forensic Sci. Int. Genet. 48 (2020), 102344.