Review Article

The Use of Forensic DNA Phenotyping in Predicting Appearance and Biogeographic Ancestry

Peter M. Schneider, Barbara Prainsack, Manfred Kayser

Summary

Background: Persons whose identifying DNA profile (STR profile) is not yet known to the investigating authorities cannot be identified by standard forensic DNA analysis (STR profiling) as it is now practiced. In view of the current public debate, particularly in Germany, on the legalization of so-called forensic DNA phenotyping, we present its scientific basis, societal aspects, and forensic applications and describe the analytic techniques that are now available.

Methods: This review is based on pertinent publications that were retrieved by a selective search in PubMed and in public media, and on the authors' own research.

Results: Forensically validated DNA test systems are available for the categorization of eye, hair, and skin color and the inference of continental biogeographic ancestry. As for statistical measures of test accuracy, the AUC (area under the curve) values lie in the range 0.74–0.99 for eye color, 0.64–0.94 for hair color, and 0.72–0.99 for skin color, depending on the predictive model and color category used. The corresponding positive predictive values (PPV) are lower. Empirical social-scientific research on forensic DNA phenotyping has shown that preserving privacy and protecting against discrimination are major ethical and regulatory considerations.

Conclusion: All three methods of forensic DNA phenotyping—the prediction of externally visible characteristics, biogeographic ancestry, and the estimation of age from crime scene DNA—require a proper regulatory framework and should be used in conjunction with each other. Before forensic DNA phenotyping can be implemented in forensic practice, steps must be taken to minimize the risks of violation of privacy and of ethnic discrimination and to ensure that these methods are used transparently and proportionately.

Cite this as:

Schneider PM, Prainsack B, Kayser M: The use of forensic DNA phenotyping in predicting appearance and biogeographic ancestry. *Dtsch Arztebl Int* 2019; 116: 873–80. DOI: 10.3238/arztebl.2019.0873

The legalization of forensic DNA phenotyping has been publicly debated in Germany and Switzerland in the past few years while the technology has already been implemented in a few other countries. Forensic DNA phenotyping goes beyond standard forensic DNA profiling, where short tandem repeat (STR) polymorphisms are used to identify individuals from DNA obtained at crime scenes and other DNA markers are used for sex determination. In more novel forensic DNA phenotyping, however, information from crime scene DNA about externally visible characteristics and biogeographic ancestry (i.e., the inferred geographic region(s) of origin of a person's biological ancestors) are collected. Estimating a person's age from DNA found at the crime scene is also part of forensic DNA phenotyping and is currently debated as well, but will only be peripherally considered in this review, as it differs from the other two types of forensic DNA phenotyping in its scientific, including molecular, basis, analytical methods, and sample requirements (*Box*). Genetic diseases and their predispositions are excluded from forensic DNA phenotyping, as it is held that their forensic use would disproportionately violate privacy.

The public debate about the legalisation of forensic DNA phenotyping arose after many years of research into the genetic basis of physical appearance and the genetic correlates of biogeographic ancestry, followed by the development of suitable analytical techniques that could be used forensically for the DNA-based prediction from human biological samples obtained at crime scenes (1, 2, 3). Such DNA test systems are needed because of the general limitations of forensic STR-profiling. The latter can only yield an individual identification by a direct comparison of the crime-scene STR-profile with the STR-profile of a suspect in the case, or of a person whose STR-profile is stored in a nationwide forensic DNA database (e.g., the DNA database of the Federal Criminal Police Office [*Bundeskriminalamt*], the German counterpart to the F.B.I.). Positive identification is possible only if there is a match, i.e., absolute agreement of the STR-profile of the crime scene sample with that of a known person.

In criminal cases without a STR-profile match, forensic DNA phenotyping can be a helpful

Institute of Legal Medicine, University Hospital of Cologne, University of Cologne, Germany: Prof. Dr. rer. nat. Peter M. Schneider

Department of Political Science, University of Vienna, Austria; Department of Global Health & Social Medicine, King's College London, United Kingdom: Prof. Dr. phil. Barbara Prainsack

Department of Genetic Identification, Erasmus MC University Medical Center Rotterdam, Netherlands: Prof. Dr. rer. nat. Manfred Kayser

TABLE 1

Currently available forensic DNA test systems*¹ for eye, hair, and skin color

Test system/ Reference	Feature	DNA markers* ²	DNA technology / forensic test validation* ³	Statistical model	Test accuracy* ⁴	Remarks
Non-commercial						
IrisPlex (7, 22)	Eye color	6 aSNPs	1x PCR+SBE+CE Yes	IrisPlex model; multinomial logistic regression; current version; N = 9466; freely available https://hirisplex.erasmusmc.nl/	Current version of IrisPlex model for eye color (6 DNA markers), AUC; PPV; NPV: – blue 0.94; 0.90; 0.90 – brown 0.95; 0.77; 0.96 – intermediate 0.74; 0.09; 0.96 https://hirisplex.erasmusmc.nl/	Eyes that are neither brown nor blue are often reflected in test results not with the highest probab- ility assigned in the intermedi- ate category, but rather with similarly high probabilities in the blue and brown eye color categories
SHEP 1, 2 (23)	Eye color	23 (13) aSNPs	2x PCR+SBE+CE No	Snipper; N = 416; freely available http://mathgene.usc.es/snipper/	AUC (13 DNA markers): – blue 0.999 – brown 0.99 – green-hazel 0.82 PPV; NPV (8 DNA markers): – blue 0.8; 0.99 – brown 0.51; 0.96 – green-hazel 0.55; 0.93	The test assays contain additional DNA markers that are not included in the model; relatively small model data set
HlirisPlex (6, 24)	Hair, eye color	24 aSNPs	1x PCR+SBE+CE Yes	HlirisPlex model for hair color; multinomial logistic regression; current version; N = 1878; freely available https://hirisplex.erasmusmc.nl/	Current version of HlirisPlex model for hair color (22 DNA markers), AUC; PPV; NPV: – red 0.92; 0.73; 0.97 – black 0.83; 0.70; 0.91 – blond 0.80; 0.63; 0.79 – brown 0.72; 0.58; 0.72 https://hirisplex.erasmusmc.nl/ (for eye color, see IrisPlex)	Contains all 6 IrisPlex DNA markers
SHEP 1, 2 (25)	Hair color	12 aSNPs	2x PCR+SBE+CE No	Snipper; N = 605; freely available http://mathgene.usc.es/snipper/	AUC (12 DNA markers): – red 0.94 – blond 0.86 – black 0.84 – brown 0.64	The test assays contain additional DNA markers that are not included in the model; relatively small model data set
SHEP 1, 2, 4 (26)	Skin color	29 (10) aSNPs	3x PCR+SBE+CE No	Snipper; N = 118; freely available http://mathgene.usc.es/snipper/	AUC (10 DNA markers) – white 0.99 – black 0.97 – intermediate 0.80	The test assays contain addi- tional DNA markers that are not included in the model; small model data set
HlirisPlex-S (8, 27, 28)	Skin, hair, eye color	41 aSNPs	2x PCR+SBE+CE; or 1x targeted MPS Yes	HlirisPlex-S model for skin color; multinomial logistic regression; current version; N = 1423; freely available https://hirisplex.erasmusmc.nl/	Current version of HlirisPlex-S model for skin color (36 DNA markers), AUC; PPV; NPV: – very light 0.74; 0.40; 0.94 – light 0.72; 0.60; 0.72 – intermediate 0.73; 0.60; 0.73 – dark 0.88; 0.34; 0.98 – dark to black 0.96; 0.81; 0.99 https://hirisplex.erasmusmc.nl/ (for hair color, see HlirisPlex; for eye color, see IrisPlex)	Contains all 24 HlirisPlex DNA markers; higher test accuracy for skin color than SHEP 1, 2, 4 when tested on the same subjects (27)
Commercial						
ForenSeq™ DNA Signature Prep Kit (Verogen, USA) (31, 32)	Eye, hair color	24 aSNPs	1x targeted MPS yes	Unknown Presumably initial versions of the IrisPlex and HlirisPlex models; available via commercial ForenSeq™ Universal Analysis software (Verogen)	Initial version of IrisPlex model (29) for eye color, AUC; PPV; NPV: – blue 0.91; 0.90; 0.84 – brown 0.93; 0.67; 0.97 – intermediate 0.73; 0.25; 0.90 Initial version of HlirisPlex model (30) for hair color, AUC; PPV; NPV: – red 0.9; 0.77; 0.91 – black 0.78; 0.45; 0.90 – blond 0.75; 0.67; 0.72 – brown 0.72; 0.21; 0.91	The test assay contains addi- tional DNA markers for other purposes. The basis of the statistical analysis is not precisely known. Information in the literature (32) suggests that the parameters of the initial versions of the IrisPlex and HlirisPlex models (29, 30) were used, which, however, have since been replaced by updated model versions based on far larger data sets.

Test system/ Reference	Feature	DNA markers ^{*2}	DNA technology / forensic test validation ^{*3}	Statistical model	Test accuracy ^{*4}	Remarks
Identify (Identitas, USA) (33)	Eye, hair color Skin color: unknown	Unknown	1x SNP microarray Unknown	Unknown Unavailable	Unknown	Service analysis, contains additional DNA markers for other purposes. No in- formation available on DNA markers, statistical model, test accuracy, forensic validation; no publications in scientific journals
Parabon™ Snapshot™ (Parabon NanoLabs, USA) (34)	Eye, hair, skin color	Unknown	1x SNP microarray Unknown	Unknown Unavailable	Unknown	Service analysis; test assay contains additional DNA markers for other purposes. No information available on DNA markers used, test accuracy, statistical model, forensic validation; no publi- cations in scientific journals

^{*1} DNA test systems in the sense of multiplex DNA tests that analyze the necessary number of DNA markers with a single test assay or a small number of test assays.

^{*2} Many of the DNA markers used overlap in the various DNA tests; cf. original publications and review in (2); the number in parentheses is the number of markers used for prediction in tests containing more markers than the ones that were actually used.

^{*3} Forensic test validation in the sense of forensic validation studies of a laboratory test, with reporting of sensitivity, specificity, degradation, population, and concordance analyses, analysis of mock casework samples, and other aspects; these are generally considered a prerequisite to the practical forensic use of a laboratory test.

^{*4} Various statistical parameters are used to express test accuracy, including, for example, AUC, the area under the receiver operating characteristic (ROC) curve, which is calculated from the sensitivity and specificity; PPV, positive predictive value; NPV, negative predictive value; see also (5) for a discussion of PPV/NPV versus sensitivity/specificity. Test accuracy is the accuracy with which a DNA test can predict a particular feature. The test accuracies provided here are not to be confused with the accuracy of individual test results, as shown in the *eFigure*.

aSNPs, autosomal single-nucleotide polymorphisms; CE, capillary electrophoresis; MPS, targeted massively parallel sequencing (also called next generation sequencing, NGS); in targeted MPS, only certain DNA markers are sequenced, which are first amplified in targeted fashion; PCR, polymerase chain reaction; SBE, single base extension; SNP microarrays or DNA chips for the simultaneous analysis of hundreds of thousands of single-nucleotide polymorphisms (SNPs) utilizing DNA hybridization, which requires DNA in a quantity and quality that often cannot be obtained from crime samples.

component of a targeted police investigation to find the unknown person who left DNA traces behind at a scene of crime. The goal of forensic DNA phenotyping is to narrow down the number of potential crime scene trace donors in such cases to a smaller group of persons who most likely have the externally visible characteristics and biogeographic ancestry that were inferred from the crime-scene DNA. Ultimately, however, only forensic STR-profiling can enable the individual identification of a (known) person (exception: monozygotic twins), thus providing evidence presentable in court. Forensic DNA phenotyping is therefore an investigative tool rather than an instrument to identify a specific person.

As of December 2019, forensic DNA phenotyping is explicitly regulated and permitted by law in two EU member states (the Netherlands and Slovakia), and practiced in compliance with existing laws in seven more (the United Kingdom, Poland, the Czech Republic, Sweden, Hungary, Austria, and Spain) (*eTable*) (4). In Switzerland, it is forbidden under current law, but the legalization of forensic DNA phenotyping is currently being considered. In Germany, in November 2019, the Bundestag (Parliament) and Bundesrat (Federal Council) approved a change in the law to permit forensic DNA phenotyping (with the exception of the DNA-based inference of biogeographic ancestry; cf. *eTable* for details).

DNA-based prediction of externally visible characteristics

Of all externally visible characteristics, eye, hair, and skin color can currently be predicted from crime-scene DNA on the level of commonly used color categories (*Table 1*). This is now possible because genes and predictive DNA markers (single-nucleotide polymorphisms, SNPs) have been identified that are either causal for these features or associated with them. Forensic DNA tests for the analysis of these DNA markers have been developed and validated, along with suitable statistical predictive models (*Table 1*).

The development of any statistical predictive model, e.g., for eye color, based on a model reference dataset consisting of genotypes and associated phenotypes is followed by statistical validation of the model, yielding parameters of test accuracy. Test accuracy reflects the average accuracy with which a DNA test can predict a particular externally visible characteristic, e.g., blue eyes, and can be expressed by a variety of statistical parameters (*Table 1*). As recommended by Caliebe et al. (5), test accuracy estimates in forensic DNA phenotyping, rather than for medical diagnostic tests, should be expressed in positive and negative predictive values (PPV and NPV). Differences in test accuracy (*Table 1*) reflect differences in the information content and number of DNA markers employed in the DNA tests used, as

TABLE 2

A selection of available DNA test systems^{*1} for continental biogeographic ancestry

DNA test / Reference	Geographical regions	DNA markers	DNA technology / forensic test validation ^{*2}	Statistical model / availability	Accuracy ^{*3}	Remarks
Non-commercial						
34-plex (35, 36)	3 – AFR, EAS, EUR	34 aSNPs	1x PCR+SBE+CE Yes	Snipper: naive Bayes classification; N = 492, freely available http://mathgene.usc.edu/snipper/	Bayes classification successful for all three regions, >99.9% ^{*4}	No distinction between Europe, Southwest Asia, and South Asia; relatively small reference data set
InDel 46-plex (37, 38)	4 – AFR, EAS, EUR, AME	46 Indels	1x PCR+CE Yes	Snipper: naive Bayes classification; N = 556, freely available http://mathgene.usc.edu/snipper/	Bayes classification successful for all four regions, >99.9% ^{*4}	Relatively small reference data set
Global AIMS (11)	6 – AFR, EUR, EAS, SAS, OCE, AME	127 aSNPs	1x targeted MPS Yes	Snipper: naive Bayes classification; N = 2099, freely available http://mathgene.usc.edu/snipper/	Bayes classification successful for all six regions, >99.9% ^{*4}	
Commercial						
ForenSeq™ DNA Signature Prep Kit (Verogen, USA) (31, 32)	4 – EUR, AFR, AME, EAS	56 aSNPs	1x targeted MPS Yes	Unknown Available via commercial ForenSeq™ Universal Analysis software (Verogen)	Unknown	The test assay contains additional DNA markers for other purposes; no information available on statistical model, reference data set, or accuracy
Precision ID Ancestry Panel (ThermoFisher Scientific, USA) (39, 40)	7 – AFR, AME, EAS, EUR, SAS, SWA, OCE	165 aSNPs	1x targeted MPS Yes	Unknown available via commercial HID Genotyper software (Thermo Fisher Scientific)	Unknown	No information available on statistical model, reference data set, or accuracy

^{*1} DNA test systems in the sense of multiplex DNA tests that analyze the necessary number of DNA markers with a single test assay

^{*2} These accuracy figures are not comparable, in terms of their content or methodology, to the AUC, PPV, and NPV values for externally visible characteristics as given in Table 1, nor are they to be confused with the accuracy of individual test results, as listed in Table 3.

^{*3} Forensic test validation in the sense of forensic validation studies for a laboratory test (see footnote to Table 1).

^{*4} The probability that, in cross-validation, an unknown DNA sample of unambiguous continental origin will be assigned to the correct continent of origin.

aSNPs, autosomal single nucleotide polymorphisms; CE, capillary electrophoresis; Indels, insertion-deletion polymorphisms; MPS, targeted massively parallel sequencing; PCR, polymerase chain reaction; SBE, single base extension; AFR, sub-Saharan Africa; AME, the Americas; EAS, East Asia; EUR, Europe; OCE, Oceania; SAS, South Asia; SWA, Southwest Asia (the indigenous populations of each region)

well as in the underlying reference data used in the statistical predictive models.

Test accuracy (Table 1) should not be confused with the accuracy of individual test results (eFigure). The latter are obtained for individual persons in a specific case by applying a particular DNA test and statistical model, and are generally expressed as probability values (5) (eFigure). What probability of a particular externally visible characteristic suffices for its useful inclusion in a police investigation is not a scientific question but, rather, an operative decision of the investigating authority: what is the individual probability threshold that would justify including this feature in a police investigation when all relevant aspects of the case at hand have been considered, and what significance should be attached to this information during the investigation?

As can be seen from the test accuracy values provided in Table 1, the currently available DNA tests and statistical models permit some categories of eye, hair, and skin

color to be more accurately predicted, on average, than others: blue and brown eyes are more accurately predictable than eyes that are neither brown nor blue, red and black hair more accurately than blond and brown hair, and dark skin colors more accurately than light ones. There are multiple reasons for this. For instance, the lesser accuracy in predictability of blond and brown hair results from the fact that blond hair may darken during childhood and adolescence. Some brown-haired adults who had blond hair as children display a high individual probability value for brown hair, others for blond hair (6). Thus, any test result with a high probability of blond hair allows either of two interpretations: the person is either a blond adult or a brown-haired adult who had blond hair as a child. Persons whose eyes are neither blue nor brown only rarely have a high individual probability value for their true, intermediate eye-color category; typically, they have similar individual probability values for both the blue eye and the brown eye color categories.

TABLE 3

Individual test results from the DNA-based inference of continental biogeographic ancestry. An illustrative example with eight test subjects chosen according to actual region of origin ^{*1}

Actual region of origin	Predicted region of origin	Predicted LR ^{*2} compared to an alternative region of origin	Origin ^{*3} Europe %	Origin ^{*3} South Asia %	Origin ^{*3} Africa %	Origin ^{*3} East Asia %	Origin ^{*3} America %	Origin ^{*3} Oceania %
United Kingdom	EUR	> 10 ⁹ /SAS	100	0	0	0	0	0
Pakistan	SAS	> 10 ⁹ /EUR	0	100	0	0	0	0
Gambia	AFR	> 10 ⁹ /SAS	0	0	100	0	0	0
Vietnam	EAS	> 10 ⁹ /AME	0	0	0	100	0	0
Peru	AME	> 10 ⁹ /SAS	0	0	0	0	100	0
Somalia	AFR ^{*4}	263/SAS	0	0.4	99.6	0	0	0
Afghanistan	SAS ^{*4}	1.55/EUR	39.2	60.8	0	0	0	0
Iraq	EUR ^{*4}	3/SAS	75	25	0	0	0	0

The reference data set used here consisted of data from 2099 individuals (cf. Global AIMS, Table 2): AFR, sub-Saharan Africa (n = 504); AME, the Americas (n = 85); EAS, east Asia (n = 504); EUR, Europe (n = 503); OCE, Oceania (n = 14); SAS, South Asia (n = 489) (indigenous populations).

^{*1} The individual test results, displayed as likelihood ratios, were calculated from the analysis of DNA samples from the eight subjects with the Global AIMS Panel DNA test system, on the basis of a classification by naive Bayes analysis (11).

^{*2} LR – likelihood ratio indicating how many more times likely it is that the individual has the inferred biogeographic ancestry, compared to the alternative one indicated.

^{*3} Probabilistic estimate of the admixture proportion, i.e. the fraction of the individual's ancestry that is derived from the region in question (for genetic admixtures, see [http://mathgene.usc.es/snippet/](http://mathgene.usc.es/snipper/)).

^{*4} Results indicating a mixed biogeographic ancestry or one that cannot be inferred with sufficient accuracy based on the available data. The individual test results shown here are not to be confused with test accuracy values, as listed in Table 2.

In general, genetic predictive models for pigmentation traits need not be assumed to be population-dependent (5); however, the degree of population-independence does depend on the causality of the DNA markers used and on the extent and population distribution of contributory non-genetic factors (5). Current DNA tests for pigmentation traits (Table 1) include both causal and non-causal DNA markers. Heritability estimates of all pigmentation traits are very high; environmental influences such as age (eye and, especially, hair color) and solar radiation (hair and, especially, skin color) do exist, but their contribution is relatively minor. To minimize potential population effects, data from multiple populations are combined e.g. in the IrisPlex, HirisPlex, and HirisPlex-S models (6–8) (<https://hirisplex.erasmusmc.nl/>) (Table 1).

Forensic DNA tests for the more fine-grained prediction of eye, hair, and skin color are not yet available, nor is there any such test for age-related loss of hair color (2). DNA-based prediction of other externally visible characteristics (e.g., hair structure, hair loss, and body height) is not yet possible with test accuracies comparable to those achieved for eye, hair, and skin color (2, 9, 10), and such tests will, therefore, not be discussed any further in this review.

DNA-based inference of biogeographic ancestry

Biogeographic ancestry does not in any way correspond to such concepts as ethnic origin or “race”; ethnicity and “race” are shaped by a multitude of factors that are not genetic. For the same reason, biogeographic ancestry cannot be equated with language, religion, or other manifestations of culture or tradition. It solely concerns the geo-

graphical region(s) from which a person's biological ancestors originated.

The DNA-based inference of a person's biogeographic ancestry is based on the genetic features that person has inherited from his or her biological ancestors. The farther apart the geographic regions of origin of two persons lie, the greater the genetic differences between them. These differences are due to mutations, migrations over the course of human history, local selection, genetic isolation, other effects (including random ones), and heredity. This is why there are DNA markers that are seen only in population groups from particular geographical regions, or that are very common in one geographical region and rare in others; these are called ancestry-informative DNA markers (3).

Ancestry-informative DNA markers can be passed on from one generation to the next in three different ways, which means that they display biogeographic ancestry in three different ways. Markers located on the autosomes are inherited from both parents and thus reflect the geographic origins of both. Markers located on the Y-chromosome are passed on only from father to son and thus exclusively reflect the geographical origin of a (male) person's ancestors in the purely male (paternal) lineage. Mitochondrial DNA markers are passed on only from mother to child and thus exclusively reflect the geographical origin of a (male or female) person's ancestors in the purely female (maternal) lineage. The autosomal DNA is reassorted in each generation through the processes of DNA recombination, meiosis and fertilization, with the result that only half of the total parental autosomal DNA, and thus only half of the autosomal DNA markers, are still present in each offspring. In contrast, the Y-chromosomal

BOX

DNA-based age estimation by methylation analysis

The molecular estimation of age is based on the fact that the activity of certain genes changes as a person ages. One type of age-dependent gene regulation is the increasing or decreasing degree of methylation of cytosine in CpG dinucleotides in the promoter regions of certain genes. The methods of analyzing age-dependent DNA methylation that have become established to date, relying mainly on the study of blood samples from healthy, living people, permit the estimation of age with an average error of ± 4 –5 years (21). The correlation is valid for persons aged 20–60, as biological processes relating to growth in younger persons, and disease in older ones, lead to a greater spread of methylation patterns, rendering age predictions in such persons imprecise. Moreover, a distinction must be drawn between DNA-based age estimation in known persons, e.g., to determine whether a suspect without valid proof of identity should be treated by the judicial system as a minor or an adult (cf. [20]), and DNA-based age estimation from human biological samples found at the scene of a crime, i.e., age estimation as an aspect of forensic DNA phenotyping. Particularly in the latter case, it should be borne in mind that, because of tissue specificity and the quantitative and qualitative requirements for crime-scene DNA (about 100 times more DNA is needed for this purpose than for the genotyping of the established STR markers), only a limited number of specific CpG markers can be used for each type of tissue sample (blood, saliva, semen), and the method of analyzing them must be forensically validated. Differences between analytical methods are far greater in the laboratory analysis of DNA methylation markers than in the analysis of genetic markers. Therefore, in age estimation based on DNA methylation, the same laboratory method should be used to generate both the model data and the data from biological crime scene samples.

(Y-) and mitochondrial (mt-) DNA markers are passed along essentially unchanged over many generations in the male and female lineages, respectively.

DNA-based inference of biogeographic ancestry should, therefore, include ancestry-informative DNA markers of all three kinds—autosomal, Y-chromosomal (in case of males), and mitochondrial. If all of a person's biological ancestors came from the same geographical region, then these three kinds of DNA markers will, in the ideal case, all lead to the same conclusion as to that person's biogeographic ancestry. For a person whose biological ancestors came from different regions, the autosomal DNA markers can be used to make quantitative inferences about that person's mixed biogeographic ancestry, while the Y- and mt-DNA markers separately and uniquely reflect the biographic ancestry of that person in the male and female lineages, respectively. Nonetheless, the currently available forensic DNA test systems often enable only limited inference of a mixing of ancestors from different geographic regions if such mixing occurred many generations ago.

Ancestry-informative DNA markers selected from population genetic studies around the world (mostly SNPs) have been used to develop and validate numerous forensic DNA test systems; these are highly reliable because of the number and diversity of the DNA markers used with respect to the populations that are to be distinguished from one another (Table 2). With the aid of these test systems, crime scene samples can be used to determine the biogeographic ancestry of any person at the level of detail of the continental regions Europe, sub-Saharan Africa, East Asia, South Asia, Oceania, and the Americas (where the indigenous populations of each region are meant) (Tables 2 and 3).

Such DNA tests may be less reliable, however, in the determination of a person's subcontinental ancestry, as the effects of migration within continents markedly lessen the ancestry informativeness of DNA markers.

Increasing the number of ancestry-informative DNA markers to be analyzed in samples from crime scenes, which is now possible with a method called “targeted massively parallel sequencing” (MPS) (Table 2), can make the determination of subcontinental ancestry possible (11), as can certain Y or mtDNA markers (5, 12).

The quality of DNA-based biogeographic ancestry inference depends not only on the geographical informativeness of the DNA markers that are employed, but also in large measure on the population genetic data used as a reference. Thus, DNA-based ancestry inference can only accurately describe the geographical spectrum of ancestry of a tested person if that spectrum of ancestry is well represented in the population genetic reference data set used. It follows that, whenever such test results are reported, the reference data used should be described as well.

Ethical aspects of forensic DNA phenotyping

A number of publications on the ethical, societal, and regulatory aspects of forensic DNA analysis in general (13), and of forensic DNA phenotyping in particular (14), provide valuable assistance for the ethical assessment of these technological developments and for considerations of the appropriate regulatory framework. An analysis of the social-scientific, judicial, and ethical literature as part of the VISAGE research project (www.visage-h2020.eu), as well as interviews with experts and members of civil society organizations (15), have led to the identification of the following major concerns about forensic DNA phenotyping: discrimination against minority groups, invasion of privacy, conflict with data protection (confidentiality) laws, and exaggerated expectations on the part of users and the general public.

Many experts see the greatest problem in the risk that forensic DNA phenotyping will be used in a way that discriminates against minority groups, particularly in societies where racism and xenophobia are now on the rise.

Even those who support the use of forensic DNA phenotyping in specific cases emphasise that these methods should not be used in police investigations until appropriate measures have been taken to ensure that they are used transparently and proportionately. The training of forensic DNA experts and investigators is essential to ensure that the laboratory findings are correctly generated, interpreted, documented, and transmitted to the investigating authorities, and that the authorities understand them correctly and use them properly.

Practical use of forensic DNA phenotyping

Forensic DNA phenotyping can be used even in criminal cases where there are no eyewitnesses; a further advantage over eyewitness testimony is that the estimated individual probability values are always accompanied by case-specific, individual error estimates that the police can take into account in their investigation. For example, if there is an estimated 95% probability of blue eyes (5% error rate), an investigator can consider this figure more reliable than one with an 80% probability (20% error rate) in another case. In contrast, the error rate of an eyewitness report in any particular case is impossible to assess. It is well known that eyewitnesses can testify falsely for a variety of reasons. For example, in the U.S.A., the Innocence Project revealed that 70% of the 350 erroneous verdicts retrospectively identified by STR profile analysis had been reached because of false eyewitness reports (16).

DNA-based predictions of externally visible characteristics and the biographical inference of ancestries that are rare in the area where a crime was committed will be more helpful to police investigations than predictions of features and origins that are common there. At the same time, DNA-based predictions of regionally common features can, indeed, provide useful clues while removing unfounded suspicion from minority groups, as happened in the investigation of the rape and murder of Marianne Vastra in the Netherlands in 1999 (17). Only rarely do the authorities reveal publicly, after the conclusion of a case, that forensic DNA phenotyping was successfully used; this was done in another Dutch rape and murder case, that of Milica van Doorn of 1992 (18).

As STR profile analysis has a relatively high chance of directly identifying a culprit who has already left genetic traces behind elsewhere, an STR profile analysis is generally performed first in every case. For example, in Germany, every third or fourth search for matches with DNA traces from a crime scene in the national forensic DNA database finds a match (19). In cases where STR profile analysis yields no match, an adequate amount of DNA must be available from the scene of the crime for forensic DNA phenotyping to be performed. However, many samples from crime scenes, such as the now commonly obtained skin contact traces, which contain very small quantities of DNA, are consumed for STR-profiling. Moreover, there must be an unequivocal connection between the DNA traces found at the scene of the crime and the crime itself. Thus, the DNA trace is required to have been obtained directly from the victim or from an object used to commit the crime, both to justify the expense of

Key messages

- Standard forensic DNA analysis (STR profiling) does not enable the identification of unknown persons whose STR profiles have not already been obtained by the investigating authorities.
- Forensic DNA phenotyping i.e. obtaining information on externally visible characteristics, biogeographic ancestry, and age from crime scene DNA can provide important investigative clues to help track down unknown perpetrators of crime.
- Currently available analytical techniques can be used to generate probabilistic predictions of eye-, hair-, and skin-color categories, and of continental biogeographic ancestry, on the basis of DNA samples obtained from crime scenes.
- The combination of predicting externally visible characteristics, inferring biogeographic ancestry, and estimating age from crime scene DNA increases the informational yield of criminal investigations.
- For forensic DNA phenotyping to be introduced into routine casework, measures must be taken to ensure its transparent and proportionate use.

forensic DNA phenotyping and to assure its proportionality as well as that of any subsequent DNA mass screening (as per section 81h StPO [Code of Criminal Procedure] in German law). Lastly, forensic DNA phenotyping makes use mainly of DNA traces from crime scenes that are derived from a single individual (which can be determined from the preceding STR-profiling results). Mixed DNA traces derived from two or more persons often do not enable the unequivocal identification of phenotypic features.

Certain externally visible characteristics depend on biogeographic ancestry. People with blond hair, blue eyes, and light skin are always at least partly of European ancestry. People with brown eyes, black hair, and skin of intermediate lightness are found among the indigenous populations of Europe, Asia, and the Americas; these three continental regions of origin can be reliably distinguished from each other by the DNA-based inference of biogeographic ancestry. Moreover, information on geographic origin is of investigative value in itself. It follows that a combined DNA analysis of externally visible characteristics and biogeographic ancestry can increase the informational yield of a criminal investigation. An additionally performed DNA-based estimation of age (20) (*Box*) further increases this yield, both because information on age is of investigative value in itself, and because some of the individual's externally visible characteristics may be affected by age (e.g., hair color or hair loss). It is therefore strongly recommended that all three techniques of forensic DNA phenotyping, i.e., the prediction of externally visible characteristics, biogeographic ancestry and the estimation of age from DNA should be used together in forensic practice, within an appropriate legal regulatory framework.

Acknowledgements

We thank Arwin Ralf and Theresa Gross for their help in the creation of the figures and tables, Walther Parson, Ingo Bastisch, Martina Unterländer, Richard Scheithauer, and Markus Rothschild for helpful comments, and Markus Rothschild for his initiative and practical support. The authors' scientific work related to the topic of this article was financially supported by

the University Hospital of Cologne (PMS), King's College London (BP), the University of Vienna (BP), Erasmus MC University Medical Center Rotterdam (MK), and the European Union Research and Innovation Program HORIZON 2020, contract no. 740580 (VIS-AGE).

Conflict of interest statement

Prof. Schneider has appeared as an invited speaker at scientific meetings, with reimbursement of meeting participation fees and of travel and accommodation expenses by Thermo Fisher Scientific and Promega. He serves in criminal proceedings as a court-appointed expert on the analysis of DNA traces.

Prof. Kayser is a co-inventor of patent no. EP2195448A1 ("Method to predict iris color") but receives no license fees or royalties from this. He has appeared as an invited speaker at scientific meetings, with reimbursement of meeting participation fees and of travel and accommodation expenses by Thermo Fisher Scientific, Promega, and the Wenner-Gren Foundation. He is commissioned to provide medicolegal expert reports on DNA trace analysis for the investigative authorities in multiple countries.

Prof. Prainsack and Prof. Kayser served until April 2018 as members of the Scientific Advisory Board of Identitas Inc., without receiving honoraria or other payments for this activity.

Manuscript received on 24 May 2019; revised version accepted on 19 November 2019.

Translated from the original German by Ethan Taub, M.D.

References

- Kayser M, de Knijff P: Improving human forensics through advances in genetics, genomics and molecular biology. *Nat Rev Genet* 2011; 12: 179–92.
- Kayser M: Forensic DNA phenotyping: predicting human appearance from crime scene material for investigative purposes. *Forensic Sci Int Genet* 2015; 18: 33–48.
- Phillips C: Forensic genetic analysis of bio-geographical ancestry. *Forensic Sci Int Genet* 2015; 18: 49–65.
- VISAGE Consortium: (2018) The regulatory landscape of forensic DNA phenotyping in Europe. www.visage-h2020.eu/Report_regulatory_landscape_FDP_in_Europe2.pdf (last accessed 15 May 2019).
- Caliebe A, Walsh S, Liu F, Kayser M, Krawczak M: Likelihood ratio and posterior odds in forensic genetics: two sides of the same coin. *Forensic Sci Int Genet* 2017; 28: 203–10.
- Walsh S, Liu F, Wollstein A, et al.: The HlrPlex system for simultaneous prediction of hair and eye colour from DNA. *Forensic Sci Int Genet* 2013; 7: 98–115.
- Walsh S, Liu F, Ballantyne KN, van Oven M, Lao O, Kayser M: IrisPlex: A sensitive DNA tool for accurate prediction of blue and brown eye colour in the absence of ancestry information. *Forensic Sci Int Genet* 2011; 5: 170–80.
- Chaitanya L, Breslin K, Zuniga S, et al.: The HlrPlex-S system for eye, hair and skin colour prediction from DNA: introduction and forensic developmental validation. *Forensic Sci Int Genet* 2018; 35: 123–35.
- Pospiech E, Chen Y, Kukla-Bartoszek M, et al.: Towards broadening forensic DNA phenotyping beyond pigmentation: improving the prediction of head hair shape from DNA. *Forensic Sci Int Genet* 2018; 37: 241–51.
- Liu F, Zhong K, Jing X, et al.: Update on the predictability of tall stature from DNA markers in Europeans. *Forensic Sci Int Genet* 2019; 42: 8–13.
- Eduardoff M, Gross TE, Santos C, 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* 2016; 23: 178–89.
- Chaitanya L, van Oven M, Weiler N, et al.: Developmental validation of mitochondrial DNA genotyping assays for adept matrilineal inference of biogeographic ancestry at a continental level. *Forensic Sci Int Genet* 2014; 11: 39–51.
- Wienroth M, Morling N, Williams R: Technological innovations in forensic genetics: social, legal and ethical aspects. *Recent Adv DNA Gene Seq* 2014; 8: 98–103.
- Samuel G, Prainsack B: Forensic DNA phenotyping in Europe: views "on the ground" from those who have a professional stake in the technology. *New Genet Soc* 2019; 38: 119–41.
- Samuel G, Prainsack B: Civil society stakeholder views on forensic DNA phenotyping: balancing risks and benefits. *Forensic Sci Int Genet* 2019; 43: 102157.
- Innocence Project: Eyewitness Identification Reform. www.innocenceproject.org/causes/eyewitness-misidentification/ (last accessed on 15 May 2019).
- Kayser M: Forensic use of Y-chromosome DNA: a general overview. *Hum Genet* 2017; 136: 621–35.
- Frankfurter Allgemeine Zeitung: Nach 25 Jahren des Mordes überführt. www.faz.net/aktuell/gesellschaft/kriminalitaet/moerder-wird-nach-25-jahren-durch-dna-analyse-ueberfuehrt-15346930.html (last accessed on 11 November 2019).
- Bundeskriminalamt. www.bka.de/DE/UnsereAufgaben/Ermittlungsunterstuetzung/DNA-Analyse/DNAstatistik/dnaStatistik_node.html (last accessed on 15 March 2019).
- Ritz-Timme S, Schneider PM, Mahlke NS, Koop BE, Eickhoff SB: Altersschätzung auf Basis der DNA-Methylierung. *Rechtsmedizin* 2018; 28: 202–7.
- Vidali A, Kayser M: Recent progress, methods and perspectives in forensic epigenetics. *Forensic Sci Int Genet* 2018; 37: 180–95.
- Walsh S, Lindenberg A, Zuniga SB, et al.: Developmental validation of the IrisPlex system: determination of blue and brown iris colour for forensic intelligence. *Forensic Sci Int Genet* 2011; 5: 464–71.
- Ruiz Y, Phillips C, Gomez-Tato A, et al.: Further development of forensic eye color predictive tests. *Forensic Sci Int Genet* 2013; 7: 28–40.
- Walsh S, Chaitanya L, Clarisse L, et al.: Developmental validation of the HlrPlex system: DNA-based eye and hair colour prediction for forensic and anthropological usage. *Forensic Sci Int Genet* 2014; 9: 150–61.
- Sochtig J, Phillips C, Maronas O, et al.: Exploration of SNP variants affecting hair colour prediction in Europeans. *Int J Legal Med* 2015; 129: 963–75.
- Maronas O, Phillips C, Sochtig J, et al.: Development of a forensic skin colour predictive test. *Forensic Sci Int Genet* 2014; 13: 34–44.
- Walsh S, Chaitanya L, Breslin K, et al.: Global skin colour prediction from DNA. *Hum Genet* 2017; 136: 847–63.
- Breslin K, Wills B, Ralf A, et al.: HlrPlex-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* 2019; 43: 102152.
- Liu F, van Duijn K, Vingerling JR, et al.: Eye color and the prediction of complex phenotypes from genotypes. *Curr Biol* 2009; 19: R192–3.
- Branicki W, Liu F, van Duijn K, et al.: Model-based prediction of human hair color using DNA variants. *Hum Genet* 2011; 129: 443–54.
- Verogen: Focused Forensic Power. verogen.com/products/ (last accessed on 15 May 2019).
- Jager AC, Alvarez ML, Davis CP, 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* 2017; 28: 52–70.
- <http://wyndhamforensic.ca/forensic-testing/> (last accessed on 2 December 2019).
- Parabon Nanolabs: The Snapshot DNA Phenotyping Service. <https://snapshot.parabon-nanolabs.com/phenotyping> (last accessed 15 March 2019).
- Phillips C, Salas A, Sanchez JJ, et al.: Inferring ancestral origin using a single multiplex assay of ancestry-informative marker SNPs. *Forensic Sci Int Genet* 2007; 1: 273–80.
- Fondevila M, Phillips C, Santos C, et al.: Revision of the SNPforID 34-plex forensic ancestry test: assay enhancements, standard reference sample genotypes and extended population studies. *Forensic Sci Int Genet* 2013; 7: 63–74.
- Pereira R, Phillips C, Pinto N, et al.: Straightforward inference of ancestry and admixture proportions through ancestry-informative insertion deletion multiplexing. *PLoS One* 2012; 7: e29684.
- Santos C, Phillips C, Oldoni F, et al.: Completion of a worldwide reference panel of samples for an ancestry informative Indel assay. *Forensic Sci Int Genet* 2015; 17: 75–80.
- Thermo Fisher Scientific: Precision ID Ancestry Panel. www.thermofisher.com/order/catalog/product/A25642 (last accessed on 15 May 2019).
- Jin S, Chase M, Henry M, et al.: Implementing a biogeographic ancestry inference service for forensic casework. *Electrophoresis* 2018; 39: 2757–65.

Corresponding author

Prof. Dr. rer. nat. Manfred Kayser
Department of Genetic Identification
Erasmus MC University Medical Center Rotterdam
PO Box 2040
3000 CA Rotterdam, the Netherlands
m.kayser@erasmusmc.nl

Cite this as:

Schneider PM, Prainsack B, Kayser M: The use of forensic DNA phenotyping in predicting appearance and biogeographic ancestry. *Dtsch Arztebl Int* 2019; 116: 873–80. DOI: 10.3238/arztebl.2019.0873

► Supplementary material

eTable and eFigure:
www.aerzteblatt-international.de/19m0873

Supplementary material to:

The Use of Forensic DNA Phenotyping in Predicting Appearance and Biogeographic Ancestry

by Peter M. Schneider, Barbara Prainsack, and Manfred Kayser

Dtsch Arztebl Int 2019; 116: 873–80. DOI: 10.3238/arztebl.2019.0873

Illustration of individual test results from DNA-based prediction of eye, hair, and skin color. These are illustrative examples of six test subjects who were chosen for their diverse eye, skin, and hair color phenotypes, as shown in the photographs. A DNA sample from each subject was analyzed with the HirisPlex-S DNA testing system (8). On the basis of the individual genetic data thus obtained, and through the use of the current IrisPlex, HirisPlex, and HirisPlex-S models (<https://hirisplex.erasmusmc.nl/>), individual probability values (test results) for three categories of eye color, four categories of hair color, and five categories of skin color were calculated and are shown to the right of each phenotype photograph. The individual test results shown here are not to be confused with test accuracy values, which are listed in *Table 1*.

eTABLE

Overview of the legal situation concerning forensic DNA phenotyping in Germany, Austria, Switzerland, and other EU member states for which definitive information is available (updated and revised from [4]; current as of December 2019)

Country	Is forensic DNA phenotyping explicitly legally regulated?	Are there legal norms that implicitly forbid forensic DNA phenotyping?	Can forensic DNA phenotyping be considered to be allowed, for practical purposes?*	Remarks
Germany	Nationwide: yes, since revision of the Code of Criminal Procedure (StPO) in November 2019 Bavaria: yes, but only to avert danger	No, since revision of the Code of Criminal Procedure (StPO) in November 2019	Nationwide: yes, since revision of the Code of Criminal Procedure (StPO) in November 2019 Bavaria: yes, but only to avert danger	A change in the StPO to permit forensic DNA analysis for the determination of eye, hair, and skin color and of age was passed by the Bundestag on 15 November 2019 and by the Bundesrat on 29 November 2019. The new law still does not permit the analysis of biogeographic ancestry. In Bavaria, molecular genetic studies of tissue samples of unknown origin, taken from crime scenes, have been allowed since May 2018 for the determination of eye, hair, and skin color, age, and biogeographic ancestry, but solely for the purpose of averting danger (not for criminal investigations per se), according to the provisions of the newly revised Police Responsibilities Act.
Austria	No	No	Yes	Forensic DNA phenotyping has been allowed since 2018, when the Security Police Law was revised for compatibility with the European Union's General Data Protection Regulation.
Switzerland	Currently, no, but a revision in the law is forthcoming	Currently, yes, but no longer after the expected revision in the law	No	A proposed revision of the law that would permit forensic DNA analysis for eye, hair, and skin color, biogeographic ancestry, and age is now under consultation as of December 2019.
Belgium	No	No	Debated/undetermined	Legal norms now forbid the use of coding markers for the identification of persons; it is unclear, however, whether the prohibition applies to forensic DNA phenotyping, which may or may not be considered a technique for the identification of persons in the sense of these legal norms.
Finland	No	No	Debated/undetermined	Legalization is being considered.
France	No	Yes	Debated/undetermined	Current norms can be interpreted as forbidding forensic DNA phenotyping, but a recent judicial decision permits DNA analysis for the inference of "morphological features" of suspected criminal offenders. It is now performed in France on the basis of this decision.
Greece, Ireland	No	Yes	No	
Italy, Croatia, Cyprus	No	No	Debated/undetermined	
Lithuania, Malta, Portugal, Slovenia	No	No	No	
Luxembourg	No	Yes	Debated/undetermined	
The Netherlands	Yes	No	Yes	Forensic DNA analysis to determine biogeographic ancestry and externally visible physical features has been allowed, in principle, since a revision of the law in 2003, but the inclusion of any specific externally visible characteristics requires a decision of the lower house of Parliament and a royal decree; these were obtained for eye color in 2012 and for hair color in 2017. This regulatory procedure was initiated for skin color in 2018, but has not yet been concluded as of December 2019. It remains unclear as of December 2019 whether forensic DNA analysis for age is permitted.
Poland, the Czech Republic, Sweden, Hungary	No	No	Yes	
Slovakia	Yes		Yes	Act no. 417/2002, § 2, a, f, § 4 in force since May 2018.
Spain	No	No	Yes	Legalization is being considered.
United Kingdom	No	No	Yes	Forensic DNA phenotyping is not explicitly forbidden by law and is considered to be allowed.

* For countries where forensic DNA phenotyping is not explicitly legally regulated and uncertainty remains as to whether it is permitted even after all potentially relevant laws have been considered, the question of its legality has been answered on the basis of interviews with legal experts, forensic scientists, and other experts in the individual countries in question. The method and further details are described in the source document, which was created in the setting of the VISAGE project and can be accessed via the VISAGE website at www.visage-h2020.eu