Taylor & Francis
Taylor & Francis Group

Australian Journal of Forensic Sciences

ISSN: 0045-0618 (Print) 1834-562X (Online) Journal homepage: www.tandfonline.com/journals/tajf20

DNA intelligence using sex-chromosome, phenotype-informative and ancestry-informative markers in an Australian population

Jessica L Watson, Kelly Grisedale, Jodie Ward & Dennis McNevin

To cite this article: Jessica L Watson, Kelly Grisedale, Jodie Ward & Dennis McNevin (19 Apr 2025): DNA intelligence using sex-chromosome, phenotype-informative and ancestry-informative markers in an Australian population, Australian Journal of Forensic Sciences, DOI: 10.1080/00450618.2025.2491376

To link to this article: https://doi.org/10.1080/00450618.2025.2491376

AUSTRALIAN JOURNAL OF FORENSIC SCIENCES
https://doi.org/10.1080/00450618.2025.2491376

Taylor & Francis
Taylor & Francis Group

DNA intelligence using sex-chromosome, phenotype-informative and ancestry-informative markers in an Australian population

Jessica L Watson [a,b,c], Kelly Grisedale [a,b], Jodie Ward [b,c] and Dennis McNevin [b,c]

[a]Forensics Command, Australian Federal Police, Canberra, Australia; [b]National DNA Program for Unidentified and Missing Persons, Australian Federal Police, Canberra, Australia; [c]Centre for Forensic Science, School of Mathematical & Physical Science, Faculty of Science, University of Technology Sydney, Sydney, Australia

ABSTRACT

Single nucleotide polymorphism (SNP) data can be used to infer the biological sex, externally visible characteristics (EVCs) and biogeographical ancestry (BGA) of an unknown individual. There are multiple pipelines available that can be used to generate these inferences and provide investigative leads for law enforcement to pursue. It is important for inference pipelines to be evaluated within a population representative of the intended jurisdiction prior to casework implementation. This study presents the performance of several pipelines using an Australian study population with self-declared biological sex, eye colour, hair colour and recent ancestry. The proportion of consistent results for EVC inference was higher for the HIrisPlex online tool using published interpretation guidelines for eye colour (97%) and hair colour (80%) when compared to the MiSeq FGx® Universal Analysis Software (UAS) for eye colour (74%) and hair colour (69%). For inferring BGA, a principle coordinate analysis pipeline produced the most consistent results when compared to self-declared data (86%). This was improved to 90% when inconclusive results obtained from admixed individuals were analysed with Structure. This study highlights the strengths and limitations of multiple inference pipelines to assist in the development of interpretation and reporting guidelines for Australian applications.

ARTICLE HISTORY
Received 3 February 2025
Accepted 23 March 2025

KEYWORDS
Biogeographical ancestry; single nucleotide polymorphism; externally visible characteristics; biological sex; forensic DNA phenotyping

1. Introduction

Forensic genomics can be utilized to generate DNA intelligence when database searches fail to identify a DNA sample of an unknown individual, thus providing new investigative leads[1,2]. Single nucleotide polymorphisms (SNPs) are single base pair variants that can be used to infer the biological sex, externally visible characteristics (EVCs) and biogeographical ancestry (BGA) of an individual. Using massively parallel sequencing (MPS), it is possible to sequence millions of reads from multiple samples in a single run and derive

CONTACT Jessica L Watson Jessica.Watson@student.uts.edu.au
Supplemental data for this article can be accessed online at https://doi.org/10.1080/00450618.2025.2491376.

a single DNA genotype containing various classes of SNPs that can be analysed to generate various intelligence information[3].

For the generation of DNA intelligence, phenotype-informative SNPs (piSNPs) and ancestry-informative SNPs (aiSNPs) are increasingly considered to offer investigative leads for forensic casework applications. For example, by inferring an unknown person's biological sex, eye colour, hair colour and recent ancestry, physical traits such as pigmentation can be incorporated into a craniofacial reconstruction of unidentified human remains, narrowing the pool of potential candidates[4–6]. There are several forensic DNA panels available that assess multiple classes of SNPs, allowing for these inferences to be generated in a single workflow[7–11].

The biological sex of an unknown individual can be inferred through routine forensic DNA techniques such as DNA quantification and short tandem repeat (STR) typing. A DNA quantification workflow typically includes a male-specific target from the Y chromosome to indicate the presence of male DNA[12]. Autosomal STR typing kits incorporate biological sex inference through the amplification of an insertion/deletion (indel) in the Amelogenin gene, which has different variants on the X and Y chromosomes, as well as the inclusion of Y chromosome STRs (Y STRs)[13]. Many MPS panels also include a selection of Y STRs and/or X chromosome STRs (X STRs) to infer the biological sex of the DNA donor[8–10].

In order to infer an individual's BGA, ancestry-informative SNPs (aiSNPs) with low heterozygosity and high population heterogeneity are required[14]. Ideally, these SNPs have alleles that are shared by individuals within a population group but not with other population groups[14,15]. BGA inference requires a panel of these aiSNPs, a reference database consisting of populations with genotypes of individuals who have known ancestry and a prediction algorithm that compares an unknown person's genotype to the reference database. The Kidd Lab Panel of 55 aiSNPs is the most widely used set for BGA inference, and its utility has been demonstrated using a number of prediction algorithms[14,16].

Multidimensional scaling (MDS) or dimensionality reduction methods use eigenvalue decomposition to reduce the genotypes of a collection of individuals to two or three coordinates in a two- or three-dimensional (2D or 3D) space to explain the variance amongst genotypes[17]. Principle component analysis (PCA) requires a numerical representation of genotypes and is limited to biallelic genotypes to preserve the genetic distances between variants[18,19]. The resulting scatter plot visualizes the samples and their genetic distances from one another based on the derived coordinates. The Universal Analysis Software (UAS; Verogen) has an inbuilt PCA algorithm that generates a 2D plot[18]. In this plot, individuals with genetic similarity will cluster together, reflecting different population groups.

However, clusters that cannot be differentiated in a 2D plot, like that produced by the UAS, may become differentiated by plotting a third dimension. Principle coordinate analysis (PCoA) requires an input matrix of genetic distances which can account for tri- and tetra-allelic SNPs, allowing for more flexibility than the PCA method[20]. As for PCA, a 2D or 3D plot can then be generated and individuals with genetic similarity will cluster together.

Another category of BGA prediction algorithms are model-based likelihood estimators which are more appropriate for inferring admixture[17]. Structure is an algorithm that estimates the proportion of genetic contributions to a matrix of genotypes from

K multiple ancestral population groups[21,22]. Using a Bayesian updating framework, Markov chain Monte Carlo (MCMC) simulations adjust model parameters until the likelihood function for the genotype matrix is maximized[21,22]. The estimated ancestral population contributions for the questioned genotype are compared with those for reference genotypes with known ancestry in order to infer the BGA of the questioned genotype[17].

The Forensic Research/Reference on Genetics-knowledge base (FROG-kb) is a forensic application of the Allele FREquency Database (ALFRED) that calculates the probability of observing a genotype within each population in the database, commonly known as a random match probability (RMP)[23,24]. The populations are ranked by RMP, with genotypes more likely to be observed in populations exhibiting higher RMPs.

The inference of externally visible characteristics (EVCs) has similar requirements to BGA: a panel of piSNPs, a reference database consisting of genotypes of individuals with known phenotypes, and a prediction algorithm. For EVC inference, the HIrisPlex panel, which consists of 24 piSNPs, can be used to infer hair and eye colour[25–27]. EVCs are inferred using a multinomial logistic regression (MLR) model that associates categorical phenotype (i.e. hair colour, hair shade and eye colour) with reference genotypes that have known EVCs. The probability of each phenotype is reported as a p-value (not to be confused with statistical significance). The UAS has an inbuilt MLR algorithm with a fixed reference database of individuals with known EVCs and reports p-values for hair and eye colours[18]. Similarly, the HIrisPlex System also uses MLR and a private database to report p-values for hair and eye colours, as well as hair shades[25–27].

In this study, DNA from several individuals with varying eye colour, hair colour and ancestry was genotyped to evaluate the application of these inference pipelines in an Australian population. The inferences were compared against the self-declared information provided by the volunteers to assess the suitability of various inference algorithms for generating reliable and actionable DNA intelligence that could assist in identifying an unknown individual.

2. Methods

2.1. Ethics approval and sample procurement

This research was approved by the University of Technology Sydney (UTS) Human Research Ethics Committee (HREC; UTS HREC NO. ETH21–5821). Volunteers provided self-administered buccal swabs and completed a questionnaire to provide their self-declared biological sex, eye colour, hair colour (at 20 years old) and ancestry (of themselves, their parents and their grandparents; Supplementary Material 1). DNA was extracted following the manufacturer's recommended protocols with the EZ1® DNA Investigator Kit (QIAGEN) on the EZ1® Advanced XL (QIAGEN)[28]. Extracted DNA was subject to DNA quantification, STR profiling and SNP genotyping as indicated in Table S1.

2.2. DNA quantification

All samples were quantified using the Quantifiler™ Trio DNA Quantification Kit (Thermo Fisher Scientific) on the QuantStudio™ 5 Real-Time PCR System (Thermo Fisher Scientific) following the manufacturer's recommended protocols[12,29]. This kit quantified 80 bp small autosomal (SA), 214 bp large autosomal (LA) and 75 bp male-specific targets and a degradation index (DI) was calculated from the ratio of SA and LA target concentrations.

2.3. STR profiling

Some samples underwent STR profiling using the GlobalFiler™ PCR Amplification Kit (Thermo Fisher Scientific; $n = 25$)[13]. Amplification was performed on the Veriti™ 96-Well Fast Thermal Cycler (Thermo Fisher Scientific) with a 29 cycle protocol. The SA target concentration was used to calculate the required input volume of extracted DNA for a 1.0 ng DNA input template amount. Capillary electrophoresis was performed on the 3500×L Genetic Analyser (Thermo Fisher Scientific) and genotyped with GeneMapper™ *ID-X* v1.6[30,31]. The analytical threshold was 225 RFU and the homozygous threshold was 1000 RFU.

2.4. Library preparation and sequencing with the ForenSeq® DNA Signature Prep Kit

The ForenSeq® DNA Signature Prep Kit (QIAGEN) targets 24 piSNPs, 56 aiSNPs (of which two are also piSNPs), 24 Y STRs, 7 X STRs, 27 autosomal STRs, 94 identity-informative SNPs (iiSNPs) and Amelogenin[10]. Samples were diluted based on the SA target concentration to deliver 1.0 ng in 5 µL (0.2 ng/µL) and underwent manual library preparation following the manufacturer's recommended protocol with primer mix B ($n = 57$). Samples were prepared in batches of 14 samples with a positive control (PC; 2800 M Control DNA (Promega)) and a negative control (NC; nuclease-free water). Sequencing was performed on the MiSeq® FGx Sequencing System (QIAGEN) using the MiSeq® FGx Reagent Kit (QIAGEN) with a standard flow cell (SFC)[32,33]. The results were analysed on the UAS v1.3 using the default analytical and interpretation thresholds and exported in the Sample Details Report and Phenotype Estimation Report[34].

2.5. Library preparation and sequencing with the ForenSeq® Kintelligence Kit

The ForenSeq® Kintelligence Kit (QIAGEN) targets the same piSNPs, aiSNPs and iiSNPs as the ForenSeq® DNA Signature Prep Kit, as well as an additional 85 Y SNPs, 106 X SNPs and 9,687 kinship-informative SNPs[9,35]. Samples were diluted based on the SA target concentration to deliver 1.0 ng in 25 µL (0.04 ng/µL) and libraries were prepared following a modified protocol ($n = 16$)[36]. Samples were prepared in batches of up to 12 samples with a PC (NA24385 Control DNA (Coriell Institute)) and NC (nuclease-free water) and sequenced in batches of 3 samples on the MiSeq® FGx Sequencing System with the MiSeq® FGx Reagent Kit

Table 1. Criteria for inferring biological sex for each genotyping assay. The criteria for the ForenSeq® DNA signature prep kit and ForenSeq® kintelligence kit are as defined by the manufacturer[18,34].

Method	Male	Female	Inconclusive
Quantifiler™ Trio DNA Quantification Kit[12]	Y chromosome target concentration above detection threshold	Y chromosome target concentration below detection threshold	All targets below detection threshold
GlobalFiler™ PCR Amplification Kit[13]	Amelogenin typed XY and alleles typed at DYS391 and Y indel	Amelogenin typed XX and no alleles typed at DYS391 and Y indel	No alleles detected or the contributor status is a mixture
ForenSeq® DNA Signature Prep Kit[18]	≥3 X STRs and ≥ 3 Y STRs typed	≥3 X STRs and < 3 Y STRs typed	Negative control, the contributor status is a mixture, < 3 X STRs typed or < 3 Y STRs typed
ForenSeq® Kintelligence Kit[34]	≥10 Y SNPs typed	No Y SNPs typed and call rate ≥ 50%	Negative control, the contributor status is a mixture or male and female criteria are not met

and SFC. The results were exported from the UAS v2.5 in Sample Reports and analysed using published optimized thresholds with a Microsoft Excel macro[36]. The SNP profiles were manually edited on the UAS to be consistent with the genotypes generated with the optimized thresholds and the Phenotype and Ancestry Reports were exported.

2.6. Inference of biological sex

Biological sex was inferred from results obtained from quantification, STR profiling and SNP profiling. Table 1 defines the criteria for inferring whether the DNA donor was biologically male or female, or whether the biological sex was inconclusive.

2.7. Inference of hair and eye colour

EVCs were inferred for each sample using either the ForenSeq® DNA Signature Prep Kit or the ForenSeq® Kintelligence Kit as the same 24 piSNPs are targeted by both kits. The in-built UAS MLR pipeline was assessed using the exported Phenotype Estimation Report (UAS v1.3, ForenSeq® DNA Signature Prep Kit) and Phenotype & Ancestry Report (UAS v2.5, ForenSeq® Kintelligence Kit). The EVC inferences were made with maximum p-value for each category of eye colour (blue, intermediate and brown) and hair colour (red, blond, brown and black). P-values were unable to be generated by the UAS unless all 24 piSNPs were typed.

The piSNPs were uploaded to the HIrisPlex online tool to report p-values and area under the receiver operating characteristic curve (AUC) values to account for information loss in partial profiles[37]. As a result, samples that only yielded partial piSNP profiles were able to be analysed with the HIrisPlex online tool. Hair colour was inferred from the p-values for the hair colours (red, blond, brown and black) and hair shade (light and dark) using the Enhanced Model Version 1 Prediction Guide by Walsh et al.; the final inferences were either red, blond, blond or brown, brown, brown or black or black[27]. If maximum eye colour p-value exceeded 0.9, that eye colour was inferred. However, if the maximum eye colour p-value was less

Table 2. Counts of self-declared hair colours of volunteers.

Hair Colour	Count
Red	2
Blond	5
Dark Blond	11
Brown	32
Dark Brown	20
Black	3

Table 3. Counts of self-declared eye colours of volunteers.

Eye Colour	Count
Blue	30
Grey[a]	2
Green[b]	8
Hazel[b]	9
Brown	24

[a]Categorised as blue.
[b]Categorised as intermediate.

Table 4. Performance metrics used to assess EVC inferences. p = number of volunteers with a particular EVC (positive); N = number of volunteers without a particular EVC (negative); TP = number of P for which EVC was correctly inferred (true positive); TN = number of N for which EVC was correctly inferred (true negative); FN = number of P for which EVC was incorrectly inferred (false negative); FP = number of N for which EVC was incorrectly inferred (false positive).

Metric	Formula
Sensitivity or true positive rate (TPR)	$\frac{TP}{P}$
Specificity or true negative rate (TNR)	$\frac{TN}{N}$
Positive predictive value (PPV)	$\frac{TP}{TP+FP}$
Negative predictive value (NPV)	$\frac{TN}{TN+FN}$
Balanced accuracy	$\frac{1}{2}\left(\frac{TP}{P}+\frac{TN}{N}\right)$

than 0.9, it was inferred that the DNA donor could have that eye colour or an intermediate eye colour.

The EVC inferences were compared to the self-declared hair and eye colours of the volunteers for consistency (Tables 2 and 3). The UAS and HIrisPlex pipelines were assessed by calculating the performance metrics in Table 4.

2.8. Inference of biogeographical ancestry

BGA was inferred for each sample using either the ForenSeq® DNA Signature Prep Kit or ForenSeq® Kintelligence Kit as the same 56 aiSNPs are targeted by both kits.

Table 5. Population groups and inclusion criteria for each biogeographical ancestry (BGA) inference method.

Method	Population Groups	Inclusion Criteria	Inconclusive Criteria
Universal Analysis Software (UAS)	African, East Asian and European[a]	Questioned genotype is closer to the population cluster centroid than at least one other reference genotype from that population	Questioned genotype is further from the population cluster centroid than all other reference genotypes from that population
Principle Coordinate Analysis (PCoA)	Sub-Saharan African, Middle Eastern/North African, Oceanian, American, European, South Asian and East Asian	Questioned genotype is closer to the population cluster centroid than at least one other reference genotype from that population	Questioned genotype is further from the population cluster centroid than all other reference genotypes from that population
Structure	Sub-Saharan African, Middle Eastern/North African, Oceanian, American, European, South Asian and East Asian	A population is inferred as a major (> 50%) or a minor (10–50%) contributor	N/A
Forensic Resource Reference on Genetics – Knowledge Base (FROG-kb)	Sub-Saharan African, Middle Eastern/North African, Oceanian, American, European, South Asian, East Asian, Asian[b]	Populations included until the RMP decreases by at least a factor of 3	N/A
Optimised Pipeline	Sub-Saharan African, Middle Eastern/North African, Oceanian, American, European, South Asian and East Asian	If sample is inconclusive with PCoA, the sample is analysed with Structure	N/A

[a]Admixed American not considered in this study.
[b]The 161 populations on FROG-kb for the KiddLab 55 aiSNPs were organized into eight population groups based on geography and ethnicity. See Table S2 for population groupings.

Inferences were generated using four pipelines: UAS (PCA), PCoA, Structure and FROG-kb (Table 5). The in-built PCA algorithm on the UAS plots the sample in relation to three population groups (African, East Asian and European); however, the Admixed American group was determined not to be relevant to an Australian population and excluded from this study[18]. Two principle coordinates were considered.

For PCoA and Structure, reference population data was compiled from 2,262 individuals with known ancestries from 1000 Genomes, the HGDP-CEPH database and the Simons Genome Diversity Project[38–41]. PCoA was performed using the 'ape: Analyses of Phylogenetics and Evolution' package in R and three principle coordinates were considered[42,43]. The reference and questioned genotypes were imported into the Structure software and analysed with the following parameter settings: 10000 burnin repetitions 10,000 MCMC repetitions after burnin, Admixture Model, Allele Frequencies Correlated and computation of the probability of the data (for estimating K, the number of ancestral populations)[21,22,44]. When running the simulations, K was set to seven and with 10 iterations. Finally, the aiSNPs were uploaded to FROG-kb in the format for the 'KiddLab – Set of 55 AI SNPs' to generate RMP values for each population in the database[23,24].

The population groups and interpretation criteria for each BGA inference method are detailed in Table 4. An 'Optimised Pipeline' was derived where samples were analysed with PCoA and, if the results were inconclusive, the sample was analysed with Structure.

The inferences were compared to the self-declared ancestry of the volunteers and determined to be consistent if all and only self-declared population(s) were inferred, partially consistent if some of the self-declared population(s) were inferred or all self-declared population(s) and additional populations were inferred, inconsistent if no self-declared population(s) were included or inconclusive if an inference could not be generated.

3. Results

3.1. Biological sex

The biological sex inferred by the samples analysed with the Quantifiler™ Trio Quantification Kit ($n = 73$), GlobalFiler™ PCR Amplification Kit ($n = 25$) and ForenSeq® DNA Signature Prep Kit ($n = 57$) were consistent between methods and with the self-declared sex (Table S2). However, the inference was inconclusive for two (12.5%) of the samples analysed with the ForenSeq® Kintelligence Kit ($n = 16$) for which volunteers self-declared as biologically female. Both profiles had two Y SNPs called with the coverage ranging from 22 to 40 reads per SNP. The quantification and STR profiling of these samples did not indicate contamination or a mixture that would result in an inconclusive biological sex inference.

3.2. Eye colour

When analysed with the UAS, 74% of inferences were consistent with the self-declared eye colour of the volunteers, while 23% were inconsistent (Figure 1, Table S3). Of the inconsistencies, 94% were inferred as more likely to have brown or blue eye colour when the individuals self-declared as having intermediate eye colour (Figure 2). The remaining inconsistency was for an individual with self-declared brown eye colour and all of the p-values reported by the UAS were less than 0.5 with blue eye colour having the maximum p-value (blue eye colour $p = 0.41$, intermediate eye colour $p = 0.25$, brown eye colour $p = 0.34$; Figure 2). There were two samples with partial piSNP profiles (call rate of 80%) and the UAS was unable to generate p-values.

The HIrisPlex pipeline generated more consistent inferences than the UAS, with 97% of genotypes resulting in eye colour inferences consistent with the self-declared information (Figure 1, Table S3). Only two inferences were inconsistent: one sample was inferred as brown eye colour by both HIrisPlex and UAS ($p = 0.94$), despite the individual self-declaring intermediate eye colour (Figure 3); the second inconsistency involved an individual with self-declared brown eye colour, where the HIrisPlex pipeline inferred likely blue or intermediate eye colour and the UAS pipeline inferred likely blue eye colour (Figure 3). Notably, 88% of the samples that were inconsistent with the UAS pipeline produced consistent inferences with the HIrisPlex pipeline.

Table 6 shows the calculated sensitivity and specificity for each pipeline by eye colour. For all eye colours, these values were equal to or higher for HIrisPlex than for the UAS, with the exception of specificity for intermediate eye colour.

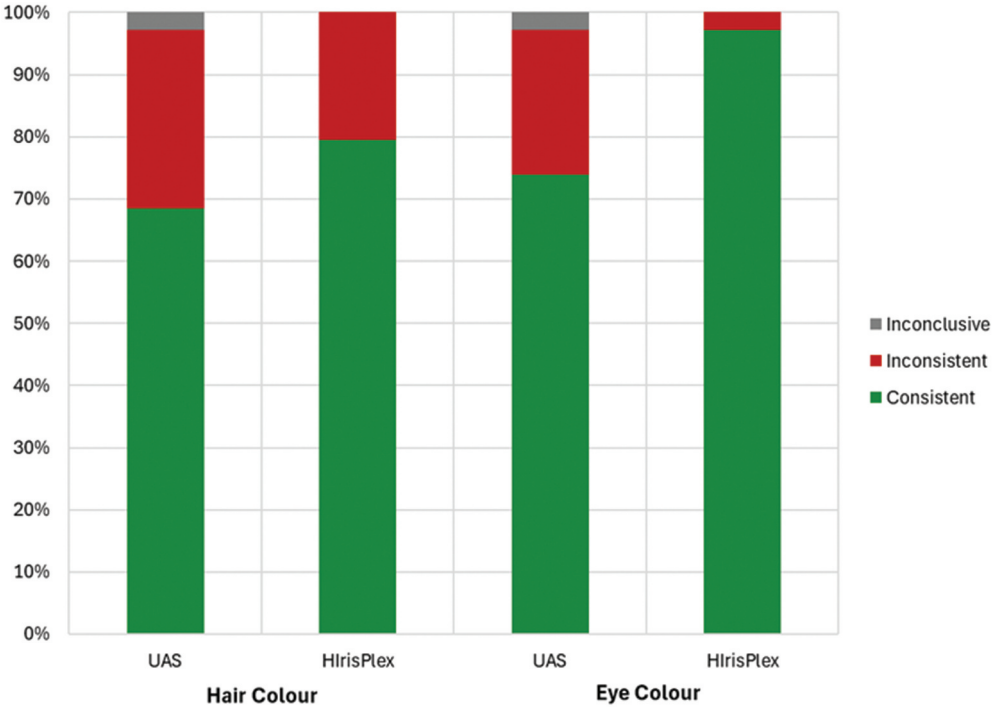

Figure 1. Externally visible characteristic (EVC) inference consistency with self-declared information of the volunteers by method: universal analysis software (UAS) and HIrisPlex.

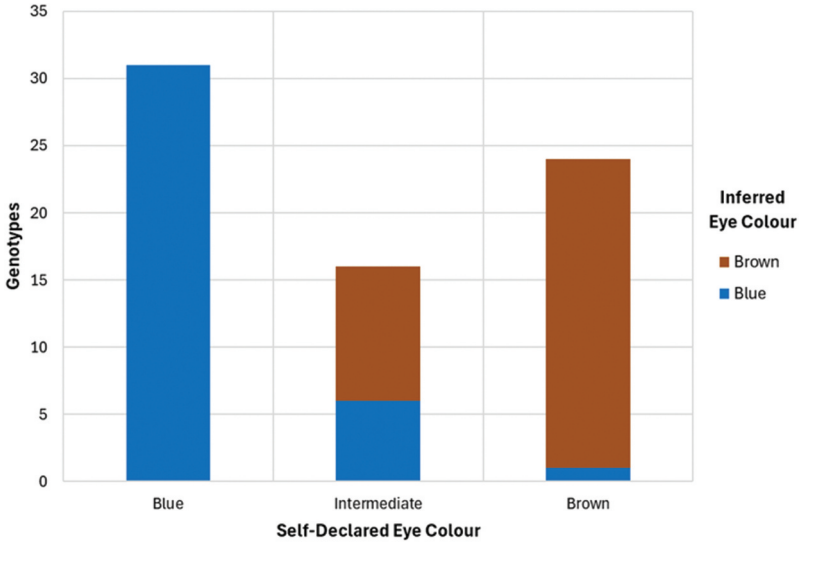

Figure 2. Eye colours inferred using the universal analysis software (UAS) by the self-declared eye colour of the volunteers.

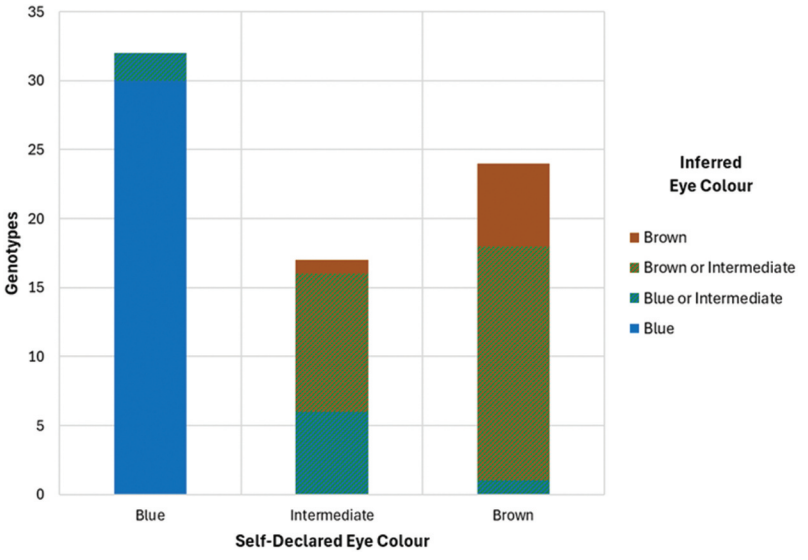

Figure 3. Eye colours inferred using the HIrisPlex by the self-declared eye colour of the volunteers.

Table 6. Performance metrics for universal analysis software (UAS) and HIrisPlex methods by eye colour category.

Eye Colour	Method	Sensitivity (TPR)	Specificity (TNR)	Positive Predictive Value (PPV)	Negative Predictive Value (NPV)	Balanced Accuracy
Blue	UAS	1.00	0.83	0.82	1.00	0.91
	HIrisPlex	1.00	0.97	0.97	1.00	0.99
Intermediate	UAS	0.00	1.00	0.00	0.77	0.39
	HIrisPlex	0.97	0.97	0.97	0.97	0.97
Brown	UAS	0.86	0.79	0.70	0.97	0.84
	HIrisPlex	0.97	0.97	0.97	0.97	0.97

3.3. Hair colour

Hair colour inferences were less consistent than eye colour inferences for both pipelines (Figure 1, Table S4). With the UAS, 69% of inferences were consistent with the self-declared hair colour, while 80% were consistent when analysed with the HIrisPlex pipeline. The two samples with partial piSNP profiles that were unable to generate results in the UAS pipeline and were therefore inconclusive, produced consistent inferences with the HIrisPlex pipeline.

All inconsistent inferences generated by the HIrisPlex pipeline were also inconsistent for the UAS pipeline, of which 10 samples were from volunteers with self-declared brown or dark brown hair colour and five were from those with self-declared blond or dark blond hair colour (Figures 4, 5). An additional six samples had UAS inferences that were inconsistent with the self-declared hair colour, of which five had self-declared brown hair colour and one had self-declared red hair colour (Figure 4). For all hair colours, the sensitivity and specificity for the HIrisPlex method was equal to or higher than the UAS (Table 7).

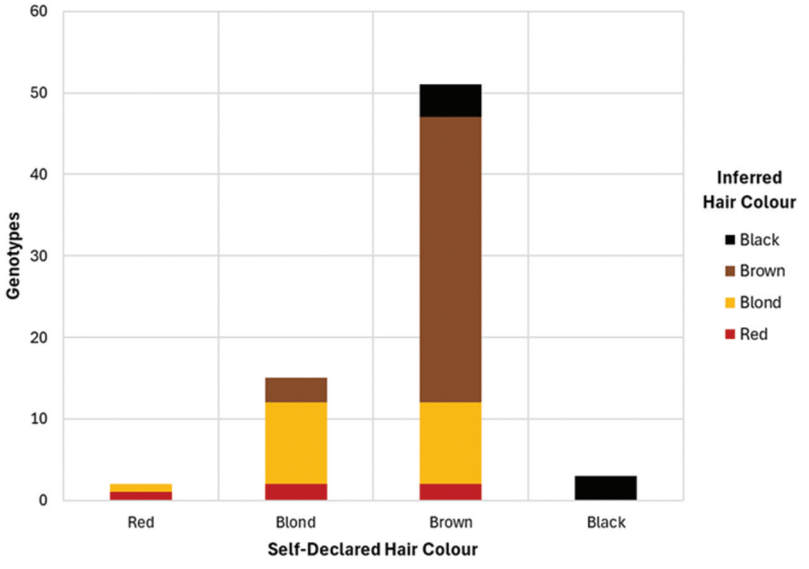

Figure 4. Hair colours inferred using the universal analysis software (UAS) by the self-declared hair colour of the volunteers.

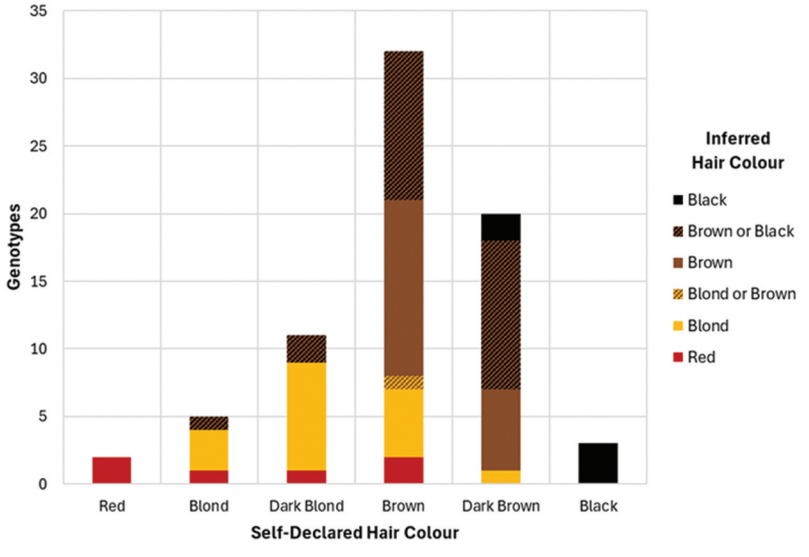

Figure 5. Hair colours inferred using HIrisPlex by the self-declared hair colour of the volunteers.

3.4. Biogeographical ancestry

The inferences generated using the PCoA pipeline had the highest consistency (86%) with the self-declared data, followed by the UAS pipeline (84%; Figure 6). The inferences that were partially consistent with the self-declared data (6% for PCoA, 8% for UAS) were due to only one of the admixed populations being detected; for the UAS pipeline, this was due to the population not being represented in the reference data (Figure 6, Table 8). Both MDS methods yielded inconclusive results for six samples each (8%), with four samples in

Table 7. Sensitivity and specificity for universal analysis software (UAS) and HIrisPlex methods by hair colour category.

Hair Colour	Method	Sensitivity (TPR)	Specificity (TNR)	Positive Predictive Value (PPV)	Negative Predictive Value (NPV)	Balanced Accuracy
Red	UAS	0.50	0.94	0.20	0.98	0.59
	HIrisPlex	1.00	0.94	0.33	1.00	0.67
Blond	UAS	0.67	0.80	0.48	0.90	0.69
	HIrisPlex	0.71	0.89	0.67	0.91	0.79
Brown	UAS	0.69	0.85	0.92	0.52	0.72
	HIrisPlex	0.81	0.86	0.93	0.64	0.79
Black	UAS	1.00	0.96	0.50	1.00	0.75
	HIrisPlex	1.00	0.96	0.93	1.00	0.96

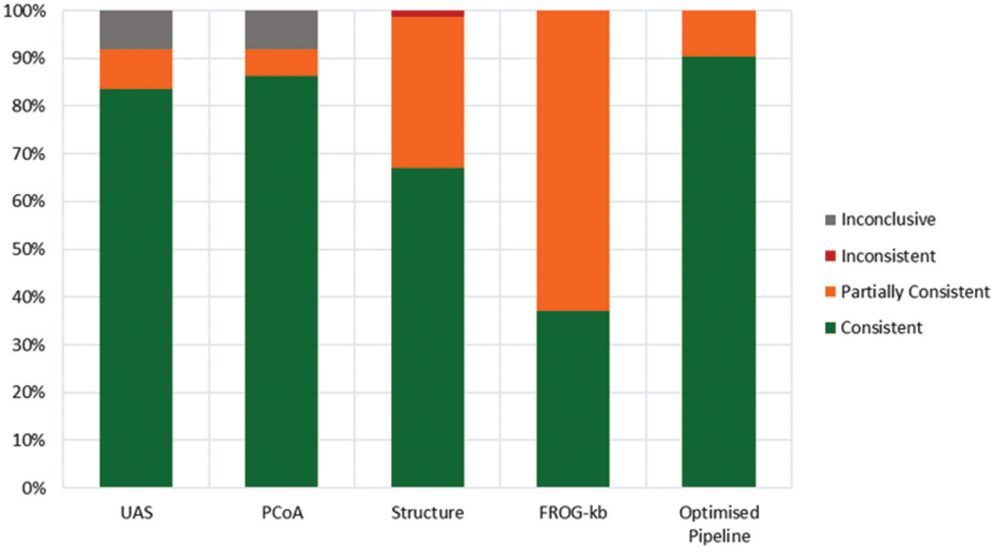

Figure 6. Biogeographical ancestry (BGA) inference consistency with self-declared ancestry of the volunteers by method: universal analysis software (UAS), principle coordinate analysis (PCoA), structure, forensic resource reference on genetics – knowledge base (FROG-kb) and the optimised pipeline.

Table 8. Categorization of biogeographical ancestry (BGA) results that were partially consistent with the self-declared ancestry of the volunteers by method: universal analysis software (UAS), principle coordinate analysis (PCoA), structure, forensic resource reference on genetics – knowledge base (FROG-kb) and the optimized pipeline.

Partially Consistent Inference	UAS	PCoA	Structure	FROG-kb	Optimised Pipeline
Some populations from self-declared admixed individuals detected and additional population(s) detected	0	0	1 (1.37%)	5 (6.85%)	1 (1.37%)
All populations from self-declared admixed individuals detected and additional population(s) detected	0	0	2 (2.74%)	3 (4.11%)	2 (2.74%)
Population from self-declared non-admixed individual detected and additional population(s) detected	0	0	18 (24.66%)	37 (50.68%)	0
Some populations from self-declared admixed individual detected	6 (8.22%)	4 (5.48%)	2 (2.74%)	1 (1.37%)	4 (5.48%)

common between the pipelines. All inconclusive results were derived from volunteers with self-declared admixture.

The Structure pipeline produced consistent BGA inferences for 67% of samples, with 32% being partially consistent and one sample exhibiting an inconsistent inference (Figure 6). In one instance, structure inferred the BGA as likely Middle Eastern ancestry with over 90% contribution, which was inconsistent with the self-declared European ancestry. FROG-kb was the least consistent, with only 37% of inferences being consistent with the self-declared ancestry and the remaining 62% being partially consistent. The majority of partially consistent inferences were samples with self-declared non-admixed ancestry, with the results inferring the self-declared population as well as additional populations (Table 8).

The optimized pipeline involved application of PCoA followed by Structure analysis for only those genotypes with inconclusive PCoA results, which improved the overall proportion of consistent results to 90%. The remaining samples were partially consistent with the self-declared ancestry and all inconclusive results were eliminated. Three samples that were inconclusive with PCoA produced consistent inferences for the self-declared admixed ancestry when analysed with Structure, while the other three samples were partially consistent (Table 8). Overall, the inferences were consistent across the five pipelines for 32% of the samples, with an additional 26% producing consistent inferences with UAS, PCoA, Structure and optimized pipelines.

3.5. Discussion

In agreement with previous studies of non-Australian populations, the HIrisPlex online tool and associated published interpretation guidelines provide flexibility for analysing partial profiles[25–27]. While the UAS uses the HIrisPlex model in an offline format and generates similar p-values to the online tool, its functionality for inferring EVCs is limited to samples with full piSNP profiles. Furthermore, additional information is provided by the online tool to assist with interpretation, such as AUC values for each category and p-values for hair shade[25–27].

Relatively high error rates for eye colour have been observed in other studies that rely on the maximum p-value approach[5,27,45,46]. These error rates were reduced following recommendations to only make inferences when the maximum p-value exceeds 0.7; however, this approach increases the likelihood of obtaining an inconclusive result[27,45]. In this study, uncertainty was incorporated by inferring both the highest p-value eye colour and intermediate eye colour if the highest p-value did not exceed 0.90, increasing the reliability of the inference pipeline. For reporting purposes, the conclusion that 'the eye colour of the individual is likely to be brown or intermediate' could alternatively be written as 'the eye colour of the individual is not likely to be blue'.

In previous studies, intermediate eye colours were the most difficult to infer as p-values tend to favour blue or brown eye colours[27,46]. In this study, 17 individuals self-declared green or hazel eye colours. The p-values generated by the HIrisPlex online tool for intermediate eye colour ranged from 0.05 to 0.34 and were never the maximum p-value produced. The sensitivity and specificity for inferring intermediate eye colour using the HIrisPlex online tool with the maximum p-value approach at the time of writing was 0.001 and 0.999, respectively[37]. However, after applying the additional thresholds in this study

to incorporate the possibility of intermediate eye colour in the inference, both sensitivity and specificity metrics were 0.97.

Inferring hair colour of an individual was more difficult than for eye colour and the consistency with self-declared data was lower. Prediction guidelines published by Walsh et al. incorporate a spectrum of hair colours[27]. Unlike the UAS, the HIrisPlex online tool generates p-values for inferring light and dark hair shades which can be used to further refine the hair colour inference with the published decision tree (i.e. dark blonde, light brown and dark brown hair colour inferences)[27]. These additional categories increased the consistency of the hair colour inferences generated by the HIrisPlex online tool compared to the in-built UAS tool.

The majority of volunteers in this study (93%) self-declared having blond or brown hair colour at 20 years old, reducing the impact of environmental and age-related hair colour changes. Both pipelines generated similar results, with the majority of inferences indicating blond or brown hair (UAS pipeline: 83%; HIrisPlex: 85%). Red hair colour had the lowest PPV for both pipelines (UAS: 0.20; HIrisPlex: 0.33), as four individuals with self-declared blond or brown hair colour were incorrectly inferred as likely having red hair. The proportion of individuals with self-declared blonde or brown hair colour who were inferred as having red hair was higher in this study than observed in previous studies[27,45,47,48].

Other studies have previously shown a correlation between pigmentation for hair, eye and skin colour and BGA[49,50]. The majority (81%) of volunteers in this study had self-declared European ancestry with no recent admixture. Other self-declared populations included East Asian (2.7%), South Asian (2.7%) and individuals with self-declared recent admixture (13%). The most important component of BGA inference is the suitability of the reference database for the jurisdiction. The UAS was designed primarily for American populations, with a 2D PCA plot of three super populations and one superimposed population representing admixed American individuals. This latter population was not deemed relevant for an Australian application of this pipeline.

The MDS approaches (UAS and PCoA) produced BGA inferences most consistent with the self-declared ancestries of the volunteers but were unable to infer admixture. The bespoke reference database used for PCoA represented a greater number of population groups and allowed for an additional PC, which was important for distinguishing between clusters[51]. BGA inferences for individuals with self-declared recent admixture, when analysed with PCoA, were either inconclusive (60%) or partially consistent (40%). The partially consistent inferences included the major ancestral contributor, where three grandparents were from the same population group and the fourth was from a different population group. The Structure and FROG-kb pipelines were more likely to infer multiple population contributions. When applying the same bespoke reference database as for PCoA, Structure helped to interpret the inconclusive results using PCoA in the optimized pipeline proposed in this study.

One limitation of BGA inferences is that they are currently restricted to continental population groups. Inference of subcontinental BGA will require the selection of DNA markers designed to reflect fine scale genetic distances within these populations[52]. Furthermore, Structure and FROG-kb often inferred both European and Middle Eastern ancestry for individuals with self-declared European ancestry only, likely due to the close geographic proximity of these regions. Use of these inferences in forensic casework will

require careful consideration of the risks associated with potential misinterpretation and misuse of the intelligence.

4. Conclusions

Where an identification cannot be achieved through direct comparison or database searching with STR profiles, inferences of biological sex, EVCs and BGA from SNP genotypes can assist law enforcement in reducing a pool of potential candidates for either a coronial or criminal investigation. The in-built UAS algorithms for hair and eye colour inference did not perform as well as the HIrisPlex online tool due to the limitation that all piSNPs must be typed in order to generate p-values. For BGA inferences, the PCoA method using our curated reference database was the highest performing pipeline but was unable to infer admixture. However, Structure was effective in refining BGA inferences for individuals with self-declared admixture that produced inconclusive results using PCoA, when integrated into an optimized analysis pipeline. By combining these inference methods, the interpretation and reporting of DNA intelligence can be improved for Australian jurisdictions.

Acknowledgments

The authors would like to thank the volunteers for providing their DNA and associated meta data for this study.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This research was funded by the Australian Federal Police (AFP) Innovation Fund and the AFP National DNA Program for Unidentified and Missing Persons. Jessica L Watson is supported by an Australian Government Research Training Program Scholarship.Institutional Review Board Statement.This study was conducted according to the guidelines of the Declaration of Helsinki and approved by the Ethics Committee of University of Technology Sydney (UTS) Human Research Ethics Committee (HREC) (UTS HREC NO. ETH21-5821).

ORCID

Jessica L Watson http://orcid.org/0000-0002-4649-3395
Kelly Grisedale http://orcid.org/0009-0000-3869-2725
Jodie Ward http://orcid.org/0000-0001-9359-9806
Dennis McNevin http://orcid.org/0000-0003-1665-3367

Author contributions

Conceptualization: J.L.W, D.M., K.G. and J.W.; Methodology: J.L.W., D.M., K.G. and J.W.; Formal Analysis: J.L.W., D.M. and K.G.; Writing – Original Draft Presentation: J.L.W.; Writing – Review & Editing: J.L.W., D.M., K.G. and J.W.; Supervision: D.M. and J.W.; Funding Acquisition: J.W.

Informed consent statement

Written informed consent was obtained from all volunteer sample donors involved in this study.

Data availability statement

Data are stored at the Australian Federal Police and may be made available to approved entities upon written request and subject to consent provisions.

References

1. Kayser M, Branicki W, Parson W, Phillips C. Recent advances in forensic DNA phenotyping of appearance, ancestry and age. Forensic Sci Int Genet. 2023;65:102870. doi:10.1016/j.fsigen.2023.102870
2. Kayser M, De Knijff P. Improving human forensics through advances in genetics, genomics and molecular biology. Nat Rev Genet. 2011;12(3):179–192. doi:10.1038/nrg2952
3. Novroski NMM. The use of massively parallel sequencing in forensic DNA typing: what will the future hold? Forensic Sci Virtual Event Ser 2020. 2020. https://www.labroots.com/webinar/massively-parallel-sequencing-forensic-dna-typing-future-hold?srsltid=AfmBOopADFEeqo4piWqzUQ5ulPoEtq7Qbft9aav8ZpBoFEW_SJ6D4Fse
4. Chaitanya L, Iz P, Walsh S, J B, Zupanc T, Kayser M. Bringing colour back after 70 years: predicting eye and hair colour from skeletal remains of world war ii victims using the hirisplex system. Forensic Sci Int Genet. 2017;26:48–57. doi:10.1016/j.fsigen.2016.10.004
5. Draus-Barini J, Walsh S, Pośpiech E, Kupiec T, Głąb H, Branicki W, Kayser M. Bona fide colour: DNA prediction of human eye and hair colour from ancient and contemporary skeletal remains. Investig Genet. 2013;4(1):3. doi:10.1186/2041-2223-4-3
6. Watherston J, McNevin D, Gahan ME, Bruce D, Ward J. Current and emerging tools for the recovery of genetic information from post mortem samples: new directions for disaster victim identification. Forensic Sci Int Genet. 2018;37:270–282. doi:10.1016/j.fsigen.2018.08.016
7. Ion ampliseq ngs panels for targeted sequencing. 2024. [accessed 2024 14 March]. https://www.thermofisher.com/au/en/home/life-science/sequencing/next-generation-sequencing/ion-torrent-next-generation-sequencing-workflow/ion-torrent-next-generation-sequencing-select-targets/ampliseq-target-selection.html
8. Tillmar A, Sturk-Andreaggi K, Daniels-Higginbotham J, Thomas JT, Marshall C. The force panel: an all-in-one snp marker set for confirming investigating genetic genealogy leads and for general forensic applications. Genes. 2021;12(12):1968. doi:10.1101/2021.11.30.470354
9. Antunes J, Walichiewicz P, Forouzmand E, Barta R, Didier M, Han Y, Perez JC, Snedecor J, Zlatkov C, Padmabandu G. Developmental validation of the forenseq® kintelligence kit, miseq fgx® sequencing system and forenseq universal analysis software. Forensic Sci Int Genet. 2024;71:103055. doi:10.1016/j.fsigen.2024.103055
10. Verogen. Forenseq DNA signature prep reference guide. ForenSeq DNA signature prep kit. 2020.
11. Verogen. Forenseq imagen kit. 2022.
12. Thermo Fisher Scientific. Quantifiler™ hp and trio DNA quantification kits user guide. Life Technologies. 2017.
13. Thermo Fisher Scientific. Globalfiler pcr amplification kits. Life Technologies. 2019.
14. Kidd JR, Friedlaender FR, Speed WC, Pakstis AJ, La Vega Fm D, Kidd KK. Analyses of a set of 128 ancestry informative single-nucleotide polymorphisms in a global set of 119 population samples. Investig Genet. 2011;2(1):1. doi:10.1186/2041-2223-2-1
15. Sanchez JJ, Phillips C, Børsting C, Balogh K, Bogus M, Fondevila M, Harrison CD, Musgrave-Brown E, Salas A, Syndercombe-Court D, et al.A multiplex assay with 52 single nucleotide polymorphisms for human identification. Electrophoresis. 2006;9(27):1713–1724. doi:10.1002/elps.200500671

16. Kidd KK, Speed WC, Pakstis AJ, Furtado MR, Fang R, Madbouly A, Maiers M, Middha M, Friedlaender FR, Kidd JR. Progress toward an efficient panel of SNPs for ancestry inference. Forensic Sci Int Genet. 2014;10:23–32. doi:10.1016/j.fsigen.2014.01.002
17. McNevin D. Forensic inference of biogeographical ancestry from genotype: the genetic ancestry lab. Wiley Interdiscip Rev. 2020;2(2):e1356. doi:10.1002/wfs2.1356
18. Verogen. 2021. Universal analysis software v2.0 and above, reference guide. Universal analysis software.
19. Gaspar HA, Breen G. Probabilistic ancestry maps: a method to assess and visualize population substructures in genetics. BMC Bioinf. 2019;20(1):1–11. doi:10.1186/s12859-019-2680-1
20. Paradis E, Claude J, Strimmer K. Ape: analyses of phylogenetics and evolution in r language. Bioinformatics. 2004;20(2):289–290. doi:10.1093/bioinformatics/btg412
21. Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. Genetics. 2000;155(2):945–959. doi:10.1093/genetics/155.2.945
22. Porras-Hurtado L, Ruiz Y, Santos C, Phillips C, Carracedo A, Lareu MAn overview of structure: applications, parameter settings, and supporting softwareFron Genet. 2013;4(98): 1–13. doi:10.3389/fgene.2013.00098
23. Kidd KK, Soundararajan U, Rajeevan H, Pakstis AJ, Moore KN, Ropero-Miller JD. The redesigned forensic research/reference on genetics-knowledge base, frog-kb. Forensic Sci Int Genet. 2018;33:33–37. doi:10.1016/j.fsigen.2017.11.009
24. Rajeevan H, Soundararajan U, Pakstis AJ, Kidd KK. Introducing the forensic research/reference on genetics knowledge base, frog-kb. Investig Genet. 2012;3(1):18. doi:10.1186/2041-2223-3-18
25. Chaitanya L, Breslin K, Zuñiga S, Wirken L, Pośpiech E, Kukla-Bartoszek M, Sijen T, Knijff PD, Liu F, Branicki W, et al. The hirisplex-s system for eye, hair and skin colour prediction from DNA: introduction and forensic developmental validation. Forensic Sci Int Genet. 2018;35:123–135. doi: 10.1016/j.fsigen.2018.04.004
26. Walsh S, Chaitanya L, Breslin K, Muralidharan C, Bronikowska A, Pospiech E, Koller J, Kovatsi L, Wollstein A, Branicki W. Global skin colour prediction from DNA. Hum Genet. 2017;136 (7):847–863. doi:10.1007/s00439-017-1808-5
27. Walsh S, Chaitanya L, Clarisse L, Wirken L, Draus-Barini J, Kovatsi L, Maeda H, Ishikawa T, Sijen T, De Knijff P, et al. Developmental validation of the hirisplex system: DNA-based eye and hair colour prediction for forensic and anthropological usage. Forensic Sci Int Genet. 2014;9:150–161. doi:10.1016/j.fsigen.2013.12.006
28. QIAGEN. 2022. Ez1&2 DNA investigator kit.
29. Thermo Fisher Scientific. Quantstudio™ 5 real-time pcr instrument (for human identification) user guide. Life Technologies. 2017.
30. Thermo Fisher Scientific. Applied biosystems 3500/3500xl genetic analyzer user guide. Life Technologies. 2010.
31. Genemapper™ id-x software v1.6. 2024. [accessed 2024]. https://www.thermofisher.com/order/catalog/product/A39975.
32. Verogen miseq fgx sequencing reagents. 2024. [accessed 2024 15 August]. https://www.qiagen.com/us/products/human-id-and-forensics/nextgeneration-sequencing/verogen-miseq-fgx-sequencing-reagents
33. Verogen. 2021. Miseq fgx sequencing system reference guide. MiSeq FGx sequencing system.
34. Verogen. 2021. Universal analysis software v1.3, reference guide. Universal analysis software.
35. Verogen forenseq kintelligence kit. 2024. [accessed 2024 24 May]. https://www.qiagen.com/us/products/human-id-and-forensics/nextgeneration-sequencing/verogen-forenseq-kintelligence-kit.
36. Watson JL, McNevin D, Grisedale K, Spiden M, Seddon S, Ward JOperationalisation of the forenseq® kintelligence kit for Australian unidentified and missing persons casework. Forensic Sci Int Genet. 2023;68: 102972. doi: 10.1016/j.fsigen.2023.102972
37. Hirisplex-s eye, hair and skin colour DNA phenotyping webtool. 2022. [accessed 2023 19 April]. https://hirisplex.erasmuscmc.nl.

38. Sudmant PH, Rausch T, Gardner EJ, Handsaker RE, Abyzov A, Huddleston J, Zhang Y, Ye K, Jun G, Fritz M H-Y, et al. An integrated map of structural variation in 2,504 human genomes. Nature. 2015;526(7571):75–81: doi: 10.1038/nature15394
39. Auton A, Abecasis GR, Altshuler DM, Durbin RM, Abecasis GR, Bentley DR, Chakravarti A, Clark AG, Donnelly P, Eichler EE. Genomes Project Consortium. A global reference for human genetic variation. Nature. 2015;526(7571):68–74. doi:10.1038/nature15393
40. Cann HM, De Toma C, Cazes L, Legrand MF, Morel V, Piouffre L, Bodmer J, Bodmer WF, Bonne-Tamir B, Cambon-Thomsen A. A human genome diversity cell line panel. Science. 2002;296 (5566):261–262. doi:10.1126/science.296.5566.261b
41. Mallick S, Li H, Lipson M, Mathieson I, Gymrek M, Racimo F, Zhao M, Chennagiri N, Nordenfelt S, Tandon A, et al. The Simons genome diversity project: 300 genomes from 142 diverse populations. Nature. 2016;538(7624):201–206. doi:10.1038/nature18964
42. R Core Team2020R: a language and environment for statistical computing, Vienna, Austria.
43. Paradis E, Bolker B, Claude J, Cuong HS, Desper R, Durand B, Dutheil J, Gascuel O, Heibl C, Lawson D, et al. Ape: analyses of phylogenetics and evolution. R package version 5.5. 2021.
44. Pritchard JK. Structure software 2.3.4. California, United: Pritchard Lab, Stanford University; 2012. https://web.stanford.edu/group/pritchardlab/structuresoftware/releaseversions/v2.3.4/html/structure.html
45. Salvo NM, Janssen K, Kirsebom MK, Meyer OS, Berg T, Olsen G-H. Predicting eye and hair colour in a Norwegian population using Verogen's ForenSeq™ DNA signature prep kit. Forensic Sci Int Genet. 2021;56:102620. doi:10.1016/j.fsigen.2021.102620
46. 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(3):170–180. doi:10.1016/j.fsigen.2010.02.004
47. Walsh S, Liu F, Wollstein A, Kovatsi L, Ralf A, Kosiniak-Kamysz A, Branicki W, Kayser M. The hirisplex system for simultaneous prediction of hair and eye colour from DNA. Forensic Sci Int Genet. 2013;7(1):98–115. doi:10.1016/j.fsigen.2012.07.005
48. Carratto TMT, Marcorin L, Do Valle-Silva G, de Oliveira MLG, Donadi EA, Simões AL, Castelli EC, Mendes-Junior CT. Prediction of eye and hair pigmentation phenotypes using the hirisplex system in a Brazilian admixed population sample. Int J Legal Med. 2021;135(4):1329–1339. doi:10.1007/s00414-021-02554-7
49. Palmal S, Adhikari K, Mendoza-Revilla J, Fuentes-Guajardo M, de Cerqueira Ccs, Bonfante B, Chacón-Duque JC, Sohail A, Hurtado M, Villegas V. Prediction of eye, hair and skin colour in Latin americans. Forensic Sci Int Genet. 2021;53:102517. doi:10.1016/j.fsigen.2021.102517
50. CCSd C, Hünemeier T, Gomez-Valdés J, Ramallo V, Volasko-Krause CD, Barbosa AAL, Vargas-Pinilla P, Dornelles RC, Longo D, Rothhammer F, et al. Implications of the admixture process in skin color molecular assessment. PLoS One. 2014;9(5):e96886. doi:10.1371/journal.pone.0096886
51. Cheung EY, Gahan ME, McNevin D. Predictive DNA analysis for biogeographical ancestry. Aust J Forensic Sci. 2018;50(6):1–8. doi:10.1080/00450618.2017.1422021
52. Schneider PM, Prainsack B, Kayser M. The use of forensic DNA phenotyping in predicting appearance and biogeographic ancestry. Deutsches Ärzteblatt Int. 2019;116(51):873. doi:10.3238/arztebl.2019.0873