Forensic Science International: Genetics 86 (2027) 103576

Contents lists available at ScienceDirect

Forensic Science International: Genetics

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

Quo vadis, BGA? A collaborative EDNAP exercise on the challenges and progress in forensic biogeographical ancestry inference

Marta Diepenbroek [a,*], António Amorim [b,c,1], Katja Anslinger [a], Natasha Arora [d], Christina Amory [e], David Ballard [f], Bram Bekaert [g], Caroline Bouakaze [h], Heloisa Costa [b,i], Ronny Decorte [g], Lena Ewers [e], Siri Aili Fagerholm [j], Oscar Garcia [k], Kris van der Gaag [l], Mario Gysi [d], Cordula Haas [d], Clemence Hollard [h], Daniel Kling [m], Leire Palencia-Madrid [k], Vania Pereira [n], María de la Puente [o], Jorge Ruiz-Ramírez [o], Maarja Sadam [p], Maja Sidstedt [j], Helle Smidt Mogensen [n], Adam Staadig [m], Monika Stoljarova-Bibb [p], Denise Syndercombe Court [f], Andreas Tillmar [m], Torben Tvedebrink [q], Catarina Xavier [e,r], Natalie Weiler [l], Walther Parson [e,s,**], Christopher Phillips [f,o]

[a] Institute of Legal Medicine, LMU Munich, Germany
[b] Instituto Nacional de Medicina Legal e Ciências Forenses, Serviço de Genética e Biologia Forenses, Portugal
[c] Faculdade de Ciências da Universidade de Lisboa, Portugal
[d] Zurich Institute of Forensic Medicine, University of Zurich, Switzerland
[e] Institute of Legal Medicine, Medical University of Innsbruck, Austria
[f] King's Forensics, Department of Analytical, Environmental and Forensic Sciences, Faculty of Life Sciences and Medicine, King's College London, London, United Kingdom
[g] Forensic Biomedical Sciences, Department of Imaging & Pathology, Leuven, KU, Belgium
[h] Institut National de Police Scientifique, Laboratoire de Police Scientifique de Lyon, France
[i] Faculdade de Ciências e Tecnologia da Universidade Nova de Lisboa, Portugal
[j] National Forensic Centre, Swedish Police Authority, Linköping, Sweden
[k] Forensic Science Unit, Forensic Genetics Lab, Basque Country Police-Ertzaintza, Spain
[l] Netherlands Forensic Institute, the Hague, Netherlands
[m] Department of Forensic Genetics and Forensic Toxicology, National Board of Forensic Medicine, Linköping, Sweden
[n] Section of Forensic Genetics, Department of Forensic Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Denmark
[o] Forensic Genetics Unit, Institute of Forensic Sciences, University of Santiago de Compostela, Spain
[p] Estonian Forensic Science Institute, Tallinn, Estonia
[q] Department of Mathematical Sciences, Aalborg University, Denmark
[r] Population Genetics and Evolution, i3S, Instituto de Investigação e Inovação em Saúde, Universidade do Porto, Portugal
[s] Forensic Science Program, The Pennsylvania State University, University Park, PA, USA

ARTICLE INFO

Keywords:
Externally visible characteristics (EVC)
Biogeographical ancestry (BGA)
Admixture analysis
Ancestry-informative markers (AIMs)
Population genetics

ABSTRACT

There is a broad consensus that forensic tests for the prediction of externally visible characteristics (EVC) and analysis of biogeographic ancestry (BGA) of an individual are technically reliable. However, interpretation of the results and population-specific genotype distribution patterns remains challenging. EVC and BGA analyses provide valuable information for population genetics studies and as investigative leads for criminal cases, as well as for historical and contemporary identification tests. However, inaccurate or incorrect predictions, for example, from subjective bias in the interpretations made, have the potential to misdirect police investigations. The legal situation regarding EVC and BGA testing varies by country: ranging from countries where it is explicitly prohibited, to those without specific regulations on biogeographic ancestry prediction, and others that have already enacted laws governing its use. The reluctance to utilize these analyses is not only due to legal restrictions and

* Correspondence to: Institute of Legal Medicine, LMU Munich, Nußbaumstraße 26, Munich 80336, Germany.
** Correspondence to: Institute of Legal Medicine, the Medical University of Innsbruck, Müllerstraße 44, Innsbruck 6020, Austria.
E-mail addresses: marta.diepenbroek@med.uni-muenchen.de (M. Diepenbroek), walther.parson@i-med.ac.at (W. Parson).

[1] Dedication: We dedicate this paper to Professor António Amorim, a long-standing friend and supportive colleague in EDNAP to many of its other co-authors. Professor Amorim sadly died in March 2024, before the paper's completion and submission.

https://doi.org/10.1016/j.fsigen.2026.103576
Received 17 December 2025; Received in revised form 15 June 2026; Accepted 22 June 2026
Available online 2 July 2026

data protection concerns, but also to initial limited sets of sufficiently comprehensive forensic DNA assays. Forensic BGA marker panels typically contain up to ~300 SNPs. This relatively small number of genetic markers, along with limited reference population data, complicates the interpretation of results from donors of unknown origin. This paper presents the results of a collaborative EDNAP study, which, for the first time, evaluated the approach to reporting EVC and BGA data between international laboratories. For the study, DNA from nine individuals with self-reported ancestry was collected and analysed using various forensic panels differing in the number and composition of ancestry-informative markers genotyped, comprising: the Precision ID mtDNA Whole Genome Panel, the VISAGE Basic Tool and the VISAGE Enhanced Tool for Appearance and Ancestry Prediction, and the Ion AmpliSeq™ PhenoTrivium Panel. To ensure full data protection, all SNP genotypes and uniparental marker haplotypes obtained were not shared with third parties. Instead, the genetic data were analysed using a range of commonly used population analysis software packages. These analysis outcomes were then distributed to twelve European forensic laboratories (both academic and law enforcement institutions), who were asked to prepare reports based on their interpretation of the phenotypes and ancestry they inferred from the analysis data. A questionnaire sent alongside the genetic information, aimed to evaluate which difficulties were encountered by the participants in processing the BGA analysis data they were given.

1. Introduction

In the last thirty-five years, the European DNA Profiling working group (EDNAP) has been at the forefront of collaborative initiatives to evaluate, optimise, and promote the introduction of a broad range of new DNA analysis techniques. Massively parallel sequencing (MPS) is being increasingly used in forensic laboratories for a range of applications, including analysis of externally visible characteristics (EVC) and biogeographic ancestry (BGA) [1]. EDNAP continues to aid in the development of optimal approaches for using MPS in forensic practice [2–4]. Recently, three large-scale MPS single nucleotide polymorphism (SNP) panels were developed for BGA and EVC. These combined-marker panels genotype SNPs informative for a DNA donor's ancestry and appearance. Such tests can enhance routine STR profiling by obtaining additional genetic information about the DNA donors. The three MPS assays comprise the Ion AmpliSeq™ PhenoTrivium Panel [5], the VISAGE Basic Tool for appearance and ancestry prediction [6], and the VISAGE Enhanced Tool for appearance and ancestry prediction [7]. All panels include the HIrisPlex-S EVC SNPs for predicting eye, hair and skin colour [8], with the Enhanced Tool including SNPs to predict the additional traits of early onset male hair loss; eyebrow colour; freckling; and hair shape [7]. In addition to the large set of autosomal ancestry informative SNPs in each panel, PhenoTrivium and the Enhanced Tool include Y-SNPs chosen for globally distributed Y haplotype analysis, and the latter includes 16 X-SNPs informative for discerning the maternal co-ancestry in males with admixed backgrounds. In 2015, EDNAP conducted a collaborative exercise for BGA where participants analysed 6 test samples of undisclosed ancestry using capillary electrophoresis (CE) to genotype a 34-plex SNP panel and a 46-plex Indel panel [9]. In a separate exercise, prediction of eye colour was evaluated by analysing 15 test samples of undisclosed eye colour by CE genotyping of 6 IrisPlex SNPs [10]. Both exercises were informative for evaluating SNP genotyping accuracy amongst participant laboratories, some using SNaPshot technology for the first time. However, the BGA exercise highlighted the need to begin building data interpretation guidelines for laboratories considering extension of routine forensic DNA analysis to specialised SNP tests to predict ancestry and appearance, where the genetic data and inferences made are more complex.

With the increased multiplexing capacity of MPS-based SNP genotyping assays, forensic BGA and EVC panels have been extended to combine many more SNPs which can compare a wider range of worldwide populations or externally visible characteristics. While EVC methods continue to centre on predicting the presence of a trait using multinomial logistic regression (MLR), ancestry tests have applied an expanding range of population analysis frameworks. These approaches aim to improve ancestry inference for individuals from less divergent populations, such as those from Eurasian sub-regions, and to detect the complex patterns of co-ancestry that can occur in admixed individuals. Differentiation of the Eurasian sub-regions of Europe, Middle East and South Asia has required adding SNPs with less contrasted allele frequencies into MPS BGA assays, bringing the need for more detailed population analyses of the genotype data. These developments prompted EDNAP to propose a "paper" exercise that only shared the analysis outcomes of a limited number of donor samples, not their SNP genotypes or mitochondrial DNA (mtDNA) sequences. Participants would be asked to select the marker set data they wanted to use, to make their own interpretations from the statistical output or haplotype assignments sent, and to prepare reports to investigators (as made now, or as they would envisage them in the future). The data analyses of nine DNA donors with self-declared ancestry and anonymous sample IDs, were sent to participants. Data included mitogenome haplotype assignments; Y-chromosomal haplogroup assignments; phenotype predictions based on HIrisPlex-S; then for autosomal SNPs, admixture analysis (STRUCTURE and Converge software), Principal Component Analysis (PCA) and GenoGeographer (GG) profile-to-reference data similarity tests and admixture analyses. Aiming to harmonize the autosomal SNP analyses, STRUCTURE, PCA, and GG all used a standardised reference population dataset of samples from seven population groups (African, European, East Asian, South Asian, Middle East, Native American, Oceanian), with the first five listed above having near identical numbers (~100 individuals) to minimize sample size bias. To compile the views and experiences of participants in handling complex and multi-component genetic data from DNA samples of unknown origin, a questionnaire was sent out to explore marker set choices, data interpretation problems, participant's own selection of reference databases for the haplotype data, and overall challenges when comparing output from different statistical analyses of population data. This feedback will create the basis for the formulation of interpretative guidelines and help to define the limitations of current forensic ancestry tests, particularly when the individuals analysed have admixed backgrounds and therefore may show complex and variable patterns of co-ancestry in their genetic data.

2. Material and methods

2.1. Sample collection and study design

Nine samples originally collected as part of the forensic validation of the Ion AmpliSeq™ PhenoTrivium Panel [5] were selected (data unpublished). Buccal swab samples were obtained from volunteers with the approval of the Bioethical Commission of Ludwig Maximilian University, Munich (.

reference number 18–870). Volunteers filled a questionnaire for self-declared ancestry based on their family tree up to grandparents, with an additional column for those with knowledge of more distant ancestors (Supplementary File S2). Self-described physical appearance was recorded and pictures of the iris, the back of the head/roots, and the forearm were taken for comparison to these descriptions (Supplementary File S17). All samples were anonymized by assigning random

numbers after collection.

Genetic data, comprising SNP genotypes and mtDNA sequences (i.e., tests supplementary to the standard forensic identification markers used in profiling) was compiled per individual. To safeguard donor privacy, the genotypes and sequence data was not shared with third parties, so data passed to participants consisted of analysis outcomes from the most commonly used population analysis software packages, plus uniparental marker haplotypes indicated by the genetic data.

Twelve forensic laboratories, both academic and police institutions, participated. Participants were free to select the data they considered most useful or relevant from the range of analysis results provided. To better understand the usefulness of the shared data and the challenges in analysing and interpreting it, a questionnaire accompanied the analysis data sent (see Supplementary File S3). Participants were asked to prepare a report for each sample treating it as a casework sample which they anticipate they will analyse (e.g., in response to future changes in the legal framework for BGA in their country) or already analyse in their laboratory.

2.2. Nuclear and mitochondrial DNA sequencing

For all samples used in the study, genomic DNA (gDNA) was extracted on the Maxwell® RSC 48 Instrument using the Maxwell® FSC DNA IQ™ Casework Kit as recommended by the manufacturer (Promega). The extracts were quantified using a Quantifiler™ Trio DNA Quantification Kit (Thermo Fisher Scientific, TFS) as recommended by the manufacturer and using the custom assay SDquants [11]. Samples were diluted to the recommended DNA input of 1 ng of nuclear DNA for the autosomal markers and 3000 copies for mitochondrial DNA.

Nuclear DNA was analysed using three assays: the Ion AmpliSeq™ PhenoTrivium Panel (herein PT), the VISAGE Basic Tool (herein BT) and the VISAGE Enhanced Tool (herein ET). The library preparation, MPS, primary and secondary analyses were carried out as described in [5–7]. Mitochondrial DNA was amplified and sequenced using the Precision ID mtDNA Whole Genome Panel following manufacturer's recommendations (TFS). The primary sequence analysis used the Torrent Server (TSS 5.10.1) with TMAP alignment to map the sample reads against the rCRS and secondary analysis performed in Converge v2.2 using the HIDGenotyper v2.2 plugin.

2.3. Population reference data

Ancestry analyses using different software were harmonized by applying a unified set of population reference data to each of the three SNP assay genotype sets and for the following analyses: i. genetic cluster analysis using STRUCTURE; ii. admixture analysis using Converge; iii. Principal Component Analysis (PCA) using Snipper, and iv. GenoGeographer, an online software used for population assignment of an individual in a forensic genetic context (the profile is compared to several reference populations present in a database to assess its most likely biogeographical origin based on the z-score statistical similarity metric) [16].

The reference population sets comprised 635 samples from seven populations: Africa, Europe, South Asia, East Asia, Oceania, America and Middle East (alternatively termed Southwest Asia). To balance population sample sizes as much as possible, single populations were used from 1000 Genomes [12], comprising: 108 African Yoruba from Nigeria (YRI); 99 CEPH Europeans from Utah (CEU); 103 South Asian Gujarati Indians from Houston (GIH) and 103 East Asian Chinese Han from Beijing (CHB). Oceanian data comprised 28 samples from combined Papua New Guinea and Bougainvillea populations from the CEPH human genome diversity panel [13]. Additionally, data included 61 Native Americans from combined CEPH Maya, Colombia, Pima, Surui, Karitiana samples plus a subset of eighteen 1000 Genomes (1KG) Peruvians from Lima (PEL) previously identified to have 98%–100% Native American ancestry proportions. Lastly, Middle East variation was represented by 106 samples from seven combined populations of Saudi, Yemeni and Emirati regions [7,14]. Therefore, the reference sample set compiled four populations representative of continental SNP variation patterns ranging from 99 to 108 samples. The combined-population Middle East sample showed a small degree of ancestry heterogeneity, with detected co-ancestry from South Asia in Emirati samples. The Oceanian and Native American reference population data were taken from the largest unadmixed sample sets it was possible to compile for these regions.

In parallel to STRUCTURE, for PhenoTrivium, admixture analysis was also conducted using Converge v2.2 software (TFS) and a built-in reference database dedicated to the Precision ID Ancestry Panel SNP set, which is a part of PT. This used a reference frequency dataset with allele frequencies for 146 SNPs of the SNPs in the Precision ID Ancestry panel from seven root populations created by hierarchal clustering of 66 populations in the ALFRED database (http://alfred.med.yale.edu) [15]. The seven ALFRED populations comprised: 1040 Africans; 628 Americans; 1110 East Asians; 1122 Europeans; 90 Oceanians; 420 South Asians; 318 Southwest Asians (corresponding to Middle East population distributions). Note this dataset could not be used for BT and ET, as only a proportion of their composite SNPs are present in ALFRED.

2.4. Data analysis of the autosomal markers

2.4.1. Ancestry markers

Ancestry prediction was based on the analysis of 163 autosomal SNPs in PT, 115 autosomal SNPs in BT, and 104 autosomal SNPs in ET (Supplementary Files S4-S6). PT corresponds with the Precision ID Ancestry Panel, except for SNPs rs6464211 and rs12439433 omitted in PT, while BT and ET consist of unique sets of SNPs developed by the VISAGE consortium with some overlap of the most informative autosomal SNPs with PT.

Autosomal SNP panels were analysed with STRUCTURE, PCA, Converge, and GenoGeographer. For the inference of ancestry of forensic samples, STRUCTURE genetic cluster analysis provides the most informative system for analysing potential co-ancestry patterns in individuals with admixed backgrounds. Using the Snipper online classifier to run Bayes likelihood ratio (LR) analysis and generate PCA plots adds further interpretative aids for inference of ancestry, but Snipper's reporting of the highest LR value to assign ancestry can disguise a minor co-ancestry component in favour of the major component, so Snipper LR values were not provided to participants. PCA plots are limited to showing samples with parental admixture from well differentiated contributor populations as points in positions between reference clusters, and only a portion of the autosomal data contributes to the principal component (PC) values of the most informative PC1 vs PC2 plots. Converge and GenoGeographer analyses gave the additional tools for participant's interpretation of autosomal SNP data. Note also that PCA was used to infer the ancestry of the X-chromosome(s) of the test samples on behalf of the participants (see Section 2.5).

2.4.1.1. STRUCTURE analysis. STRUCTURE genetic cluster analysis recognised K:7 clusters, in line with all publications for PT, BT and ET indicating this is the optimum differentiation of global variability represented by the population groups used for the exercise analyses. STRUCTURE analyses applied standard parameter settings of 100,000 burnin steps and 100,000 MCMC steps, using correlated allele frequencies under the Admixture model and updating allele frequencies using reference samples pre-designated with POPFLAG = 1, and population labels 1–7. Cluster membership proportion results from five runs were merged into average values per sample using the Structure Selector web tool (https://lmme.ac.cn/StructureSelector/). For each sample, average estimated cluster membership proportions were provided as a table and a single cluster column placed next to the narrowed reference sample cluster plot columns (Supplementary File S7).

2.4.1.2. Converge analysis. Admixture analysis with Converge was performed using the bootstrapping admixture analysis feature of the HIDGenotyper v2.2 plugin applying the built-in Converge reference dataset as well as the unified dataset compiled for the other analyses. Converge bootstrapping admixture analysis produces admixture predictions using a maximum likelihood approach to infer the most likely co-ancestry proportions from seven root populations (as described above). The predictions are bootstrapped across a random subset of SNP genotypes, user-specified percentage values, with each bootstrapping replication run through the core admixture algorithm N times using a different subset of SNPs for each replication to capture prediction uncertainty. The results are displayed as an average of the bootstrapping replications for each population group and a 95% confidence interval reflecting the probable range of variability of the estimated co-ancestry percentages. The predicted ancestry is presented as a percentage of each population with the corresponding likelihood. Sample admixture was run using the 75% resampling size and 1000 replications. The participants were given the original Converge output in PDF format (Supplementary File S8). Output included plots generated with built-in Converge reference data set and the unified population reference set.

2.4.1.3. Principal Component Analysis. Three Snipper-based PCA plots (comparing principal components, PCs: PC1 vs PC2; PC1 vs PC3; PC2 vs PC3) were provided for each assay. Samples were analysed separately to maintain clarity, i.e., PCA plots had a single black point for the unknown sample superimposed on coloured clouds of reference population points (Supplementary File S9).

2.4.1.4. GenoGeographer. GenoGeographer is an online forensic software suite used to assign an individual to a population by running two statistical calculations [16,17]. First, an outlier (z-score) test assesses whether the SNP profile fits within any of the reference populations (null hypothesis: "the profile originates from the reference population"). If the z-score is ≤ 1.64, the null hypothesis is accepted, and the profile is considered to originate from the population; if > 1.64, the profile is considered to be an outlier. Second, it calculates likelihood ratios to weigh evidence under two user-defined hypotheses: H1 (the individual belongs to population A) vs. H2 (the individual belongs to population B). LR calculations are only valid for interpretation if at least one reference population is a likely match ($z \leq 1.64$), with population A being the most likely. For samples with admixture patterns in STRUCTURE and Converge, GenoGeographer analyses applied the 1st order admixture option. Using a 95% confidence level, results were grouped by populations with the minimum sample size set to 25 (to include Oceanian population data). The participants were provided with analysis results as PDFs generated from the original HTML reports from GenoGeographer (Supplementary File S10).

2.4.2. Phenotypic markers

Phenotype predictions used the 41 HIrisPlex-S (HPS) SNPs in PT, BT and ET assays. Full and concordant HPS SNP profiles were used as input for the HPS Webtool and predictions made with the MPS SNP analysis R-script (https://walshlab.sitehost.iu.edu/pages/tools.html). Participants were provided with the *p-values* obtained for eye colour (3 phenotypes), hair shade (2) and hair colour (4), and skin colour (5), as shown in Supplementary File S17.

2.5. Data analysis of X-SNPs

The inclusion of 16 ancestry-informative X-SNPs for the first time in the ET panel [7,14] represents a novel approach to the forensic analysis of maternal co-ancestry in admixed individuals. For this reason, the X-chromosome ancestry of each sample was analysed and data provided as a summary report of the inferred X ancestry of each sample. Supplementary File S11 shows the guidance document given to participants outlining the interpretation of X-SNP data.

The 16 ET X-SNPs comprise five loci widely distributed on the p-arm, six widely distributed on the q-arm, and five more closely sited at the centromere. The five centromeric X-SNPs have little or no detected recombination between them and are treated as a single haplotype of phased genotypes in male samples. The very low recombination at the X centromere (~7% overall recombination rate across these five centromeric SNPs) means the 5-SNP motif is very likely to be preserved as an undisrupted set of SNP alleles in phase for many generations after population admixture has occurred. The 5-allele centromeric haplotypes are highly specific for African ancestry (90.3% of observed haplotypes in 1000 Genomes' Africans being specific); for East Asian ancestry (88.9% of haplotypes specific); and to a lesser extent for European ancestry (75.7%) and Native American ancestry (60%). A dual approach was taken to analysing the X-SNP patterns in the nine samples: i. comparison of the 5-SNP haplotype (analysed in female haplotype pairs when only 1/5 SNPs was heterozygous, so the phase was known) to a database of 1000 Genomes male sample 5-SNP haplotypes from four population groups; and ii. A Bayes likelihood ratio and PCA test in Snipper of the full 16 X-SNP profiles using three reference population datasets of male 1000 Genomes' X-SNP genotypes. The three reference populations used are subset combinations of the four most differentiated population groups, i.e. either African-European-East Asian, or African-European-Native American, the latter when there are indications of co-ancestry from amongst these three populations. This reduced reference data approach creates triangular PCA plots with expanded separation of reference clouds on which the unknown X-SNP PCA position can be compared. The 5-SNP centromeric haplotype, when specific, and the likelihood ratio value plus PCA plots were used to make an X ancestry inference reported directly to the participants with the interpretation guidance document.

2.6. Data analysis of the uniparental markers

Full mitogenomes obtained with the Precision ID mtDNA Whole Genome Panel were used for haplogroup classification using EMPOP v4/R13. Participants were provided with the assigned mitochondrial lineages (see Figs. 1B, 2B, 3B).

The 85 Y-SNPs in ET and 116 Y-SNPs in PT allow the inference of 60 lineages that overlap between both panels. The participants were provided with the most derived lineage within both panels which was obtained using Yleaf for ET data and the HIDGenotyper v2.2 plugin in Converge v2.2 for PT data (see Figs. 1B, 2B, 3B).

3. Results and Discussion

All genetic analysis data was shared with the participants as folders labelled KS9, KS71, KS96, KS102, KS106, KS110, KS126, KS153 and KS180. To present the exercise results in the most comprehensive manner here, we have divided samples into three groups: "Level 1" samples KS106, KS110 and KS153, which showed simple, unequivocal ancestries; "Level 2" samples KS102, KS126 and KS180, which showed very similar admixture patterns and "Level 3" samples KS71, KS96 and KS9, which had the most complex ancestry patterns.

As the primary purpose of the exercise was to evaluate the challenges faced when reporting forensic ancestry tests, based on the provided data rather than to compare the relative performance of the assays or software used; for clarity, only the key findings from the analyses of STRUCTURE, Snipper-based PCA, non-autosomal ancestry markers and externally visible characteristics are discussed below. Nevertheless, a detailed presentation of the analyses performed, and their outcomes is outlined in Supplementary Files S1, S7–S11, and S17, allowing independent scrutiny and interpretation of the genetic data originally provided to exercise participants.

3.1. Genotyping results

3.1.1. Analysis of autosomal ancestry markers

Fig. 1 shows STRUCTURE analysis of "Level 1" KS106, KS110 and KS153 (leftmost three column sets, three rows for each autosomal SNP panel) indicated clear, near-complete East Asian, European and European cluster membership proportions respectively, in each panel. Fig. 2 (PC1-PC2 plots for each panel) shows KS106, KS110 and KS153 all had PCA positions in the non-overlapping parts of each of the relevant reference clusters. All STRUCTURE cluster membership proportions and PCA positions matched the self-declared Chinese, Russian and German ancestries of these donors, respectively.

The middle three columns of Fig. 1 show STRUCTURE cluster plots of "Level 2" samples KS102, KS126, KS180, with self-declared ancestries which can be all described as geographic outliers of Europe, or peripheral to the western European reference data used. The main interpretative challenge for exercise participants from these STRUCTURE patterns arose from the visible disparity between cluster membership proportions in PT and BT vs ET, due, in large part to the focus during ET development, on expanding the numbers of Middle East-informative component SNPs. In the absence of ancestry ‚ground truth' for donor samples from e.g., large-scale SNP sets or whole-genome sequence based analysis, it is not possible to know whether ET gives a more accurate estimate of the mainly European cluster proportion in KS102, self-declared as Albanian, compared to the significant proportion of Middle East joint cluster membership detected by PT and BT. For KS126, self-declared to be from Türkiye, the disparity between panels is even more marked as almost no European cluster membership is detected by ET, contrasting with very low Middle East cluster membership from PT and BT SNPs. KS180 was an individual with German-Turkish parental co-ancestry, with STRUCTURE cluster plots slightly more consistent between panels. Turkish individuals can show combined ancestry from Europe and the Middle East, or show wholly European or wholly Middle East ancestry. Therefore, autosomal SNP analysis alone is unable to differentiate origins from different regions in Türkiye, strongly influenced by Europe to the west, and the Middle East to the south and east. PCA positions for all three "Level 2" samples show varied degrees of displacement away from European and towards Middle East clusters, but this was not a consistent pattern across SNP panels.

The rightmost columns show the most complex patterns obtained from the "Level 3" samples KS71, KS96, KS9. Sample KS71 consistently showed reasonably balanced European and East Asian cluster membership proportions, matching the self-declared German-Japanese parental co-ancestry of this donor. KS96 showed the most varied cluster patterns between panels, with ET possibly the best match to the self-declared German-Brazilian parental co-ancestry. Lastly, sample KS9 self-declared as South African, and had a degree of consistency amongst the panels in showing African and Middle East joint cluster memberships. The PCA positions of each sample aligned between the contributing ancestries indicated by STRUCTURE, but were also close to some reference clusters in similar intermediate positions, highlighting the problem of PCA tests where multiple clusters position closely to the African-European-East Asian triangle vertices.

Two factors influencing data interpretation with STRUCTURE and PCA should be highlighted. First, indications of joint cluster membership in a sample should always prompt cross-checks with the uniparental data obtained for such samples. It was expected that participants would follow the same process of interpretation by comparing autosomal data to complementary marker sets with lineage-informative characteristics. We also emphasised in the data sent that minor cluster membership proportions below 15% were unlikely to accurately reflect true low-level co-ancestry (i.e., beyond grandparental admixture contributions) when using relatively small numbers of SNPs typical of forensic panels. Second, PCA positions that show some ambiguity or are positioned close to overlapping, or very closely aligned reference clusters, can be simplified and clarified by performing a reduced population PCA in Snipper, concentrating on just the reference clusters close to the queried sample position. Reduced reference data PCA has the effect of expanding the space between clusters closely sited in ‚broad-brush' plots that include all population data. It was not possible to provide participants with a ‚pro-active' nested analysis in this way, so PCA data was limited in its informativeness for samples not positioned unambiguously within the reference clusters.

3.1.2. Analysis of non-autosomal ancestry markers

For the uniparental markers, participants received lineage assignments and for X-SNPs, they were given the interpretation and guidance notes on how these were made (Supplementary File S11). For consistency and clarity, the maternal and paternal lineage data shared with the study participants are shown in Fig. 1 with the same colour labels used for the STRUCTURE and PCA plots (note these correspond to the populations for which the lineages are specific, but shading is used when multiple populations are associated). Blank cells Y-SNP cells denote female samples and colour labels used for X-SNP ancestry inferences when these were interpreted to be population-specific, or in gray when not specific.

Among "Level 1" samples, KS106 showed clear East Asian ancestry across all markers: mtDNA F1e3, Y haplogroup O-M119, and East Asian-specific X-SNPs. KS110 showed mainly European/West Eurasian ancestry, with mtDNA H2a1c and X-SNPs strongly supporting European origin, although the centromeric X haplotype could not be phased unambiguously. KS153 indicated West Eurasian/European ancestry, with mtDNA U4c1 and Y haplogroup R-U152; the X-chromosome data also suggested European origin, but with weaker support than for KS110 (Fig. 1B).

Among the "Level 2" samples, KS102 showed consistent West Eurasian/European ancestry across all markers, with mtDNA W1, Y haplogroup R-CTS1078, and European-specific X-SNPs. In KS126, the uniparental markers pointed to different geographic affinities: mtDNA U7a is associated with the Middle East/Central–South Asia, whereas the Y haplogroup was R-CTS1078, and X-SNPs were interpreted as European with weak support. KS180 indicated predominantly West Eurasian/European ancestry, with European mtDNA U2e2a1a, West Eurasian Y haplogroup J-M67, and X-SNPs indicating a European X-chromosome (Fig. 1B).

Among the "Level 3" samples, KS71 had the East Asian mtDNA haplogroup M7b1a1a1, and X-SNPs indicated European and East Asian X-chromosome ancestries, although the 5-SNP haplotypes were not fully unambiguous. KS96 showed African-associated mtDNA and X-chromosome profiles, while the Y haplogroup R-M343 could not be assigned to a specific geographic region in the reference dataset. KS9 carried the Southeast Asian mtDNA haplogroup E1a1a1, and the X-chromosome data indicated at least one African X-chromosome, in agreement with both PCA and Bayes analysis (Fig. 1B).

3.1.3. Analysis of externally visible characteristics

The *p-values* obtained for pigmentation traits are summarised in Supplementary Fig. S17 and the donor's phenotype reference pictures in Supplementary File S18. Among the "Level 1" samples, KS106 had highest observed values for predictions of brown eyes, black hair colour and dark shade, plus intermediate skin colour. Sample KS110 had highest observed values for predictions of brown eyes, brown hair colour and light shade, plus intermediate skin colour. Sample KS153 had highest observed values for predictions of blue eyes, brown hair colour and light shade, plus pale skin colour.

Among the "Level 2" samples, KS102 and KS126 had highest observed values for brown eyes, for brown hair colour and dark shade, plus intermediate skin colour. Sample KS180 had highest observed values for brown eyes, for brown hair colour and light shade, plus pale skin colour.

Among the "Level 3" samples, KS71 had highest observed values for brown eyes, for black hair colour and dark shade, plus intermediate skin

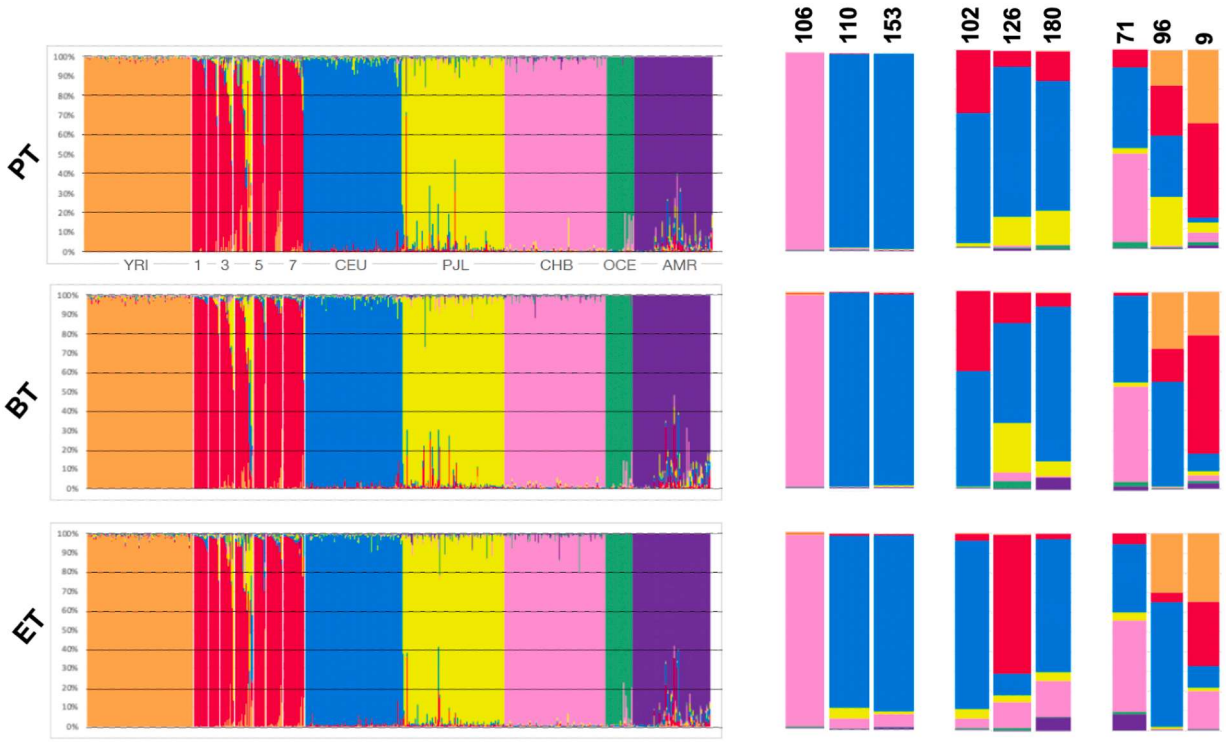
A)

	106	110	153
Origin of the individuals	China	Russia	Germany

102	126	180
Albania	Türkiye	Germany (M) Türkiye (P)

71	96	9
Japan (M) Germany (P)	Brazil (M) Germany (P)	South Africa Coloured

B)



C)

	106	110	153
mtDNA	F1e3	H2a1	U4c
87 ET Y-SNPs	O-M119	♀	R-P312
116 PT Y-SNPs	O-M119	♀	R-U152
16 ET X-SNPs	EA specific	EU specific	EU*

102	126	180
W1	U7a*	U2e2a1a
R-CTS1078	R-CTS1078	J-M172
R-M269	R-M269	J-M67
EU specific	EU**	EU specific

71	96	9
M7b1a1a1	L3e2a1a*	E1a1a1
♀	R-M343	♀
♀	R-M343	♀
EU + EA specific	AFR specific	AFR** + AFR specific

Fig. 1. Summary of the key findings from the full data analyses of exercise samples. Supplementary Files S7-S11 present more detailed results. A) The self-declared, grandparental ancestry (M – maternal; P – paternal)B) Genetic cluster analysis of autosomal SNPs using STRUCTURE with K= 7, 100,000 burnin steps and 100,000 MCMC steps, with correlated allele frequencies under the Admixture model. PT – PhenoTrivium; BT – VISAGE Basic Tool; ET – VISAGE Enhanced Tool. YRI – African; 1–7 – Middle East; CEU – European; PJL – South Asian; CHB – East Asian; OCE – Oceanian; AMR – Native American. C) Maternal and paternal lineage assignments provided to participants, plus interpretation of X-SNP data. Note the colours used here to visualise uniparental marker distributions only apply to this figure and were not given to participants. *The gray colour denotes a haplotype which is not population-specific, but was interpreted as most likely European.

colour. Sample KS96 had highest observed values for brown eyes, for brown hair colour and dark shade, plus dark skin colour. Sample KS9 had highest observed values for brown eyes, for black hair colour and dark shade, plus dark to black skin colour.

3.2. Reports

The laboratories that participated in this exercise have experience with forensic DNA phenotyping analyses, with the majority preparing reports as part of their normal casework (only Labs 2 and 4 are not). Therefore, participants focused on the analysis outcomes most closely aligned to their own routine workflow. Of the 12 laboratories involved in the study, six used all the ancestry marker set data provided to produce their reports and three laboratories used only the autosomal SNP data. The three remaining laboratories used different combinations of markers. Information about these marker combinations, including details of the SNP assay and software to analyse the autosomal markers, is provided in Supplementary File S12.

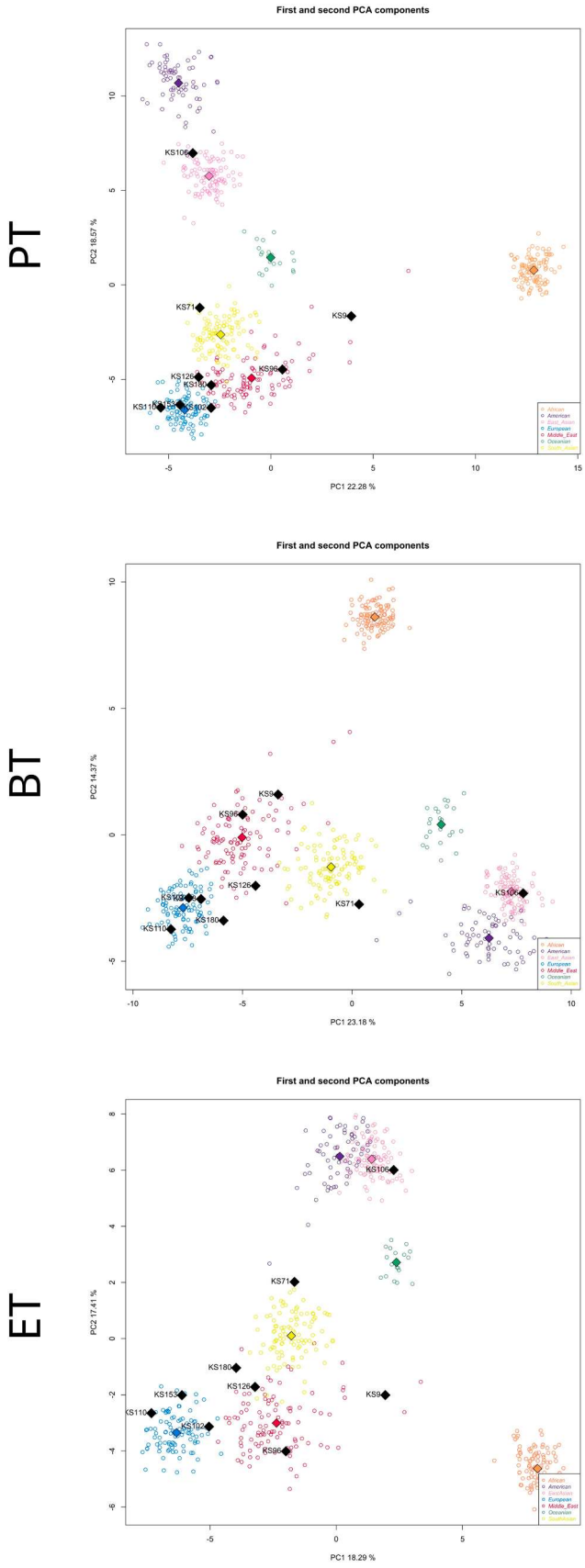


Fig. 2. PCA plots generated by Snipper for principal components PC1 vs PC2 for all samples and each SNP panel. Note original PCA colours were adjusted to match STRUCTURE cluster plots (See Fig. 1).

Because participant's reports were very complex and differed substantially between laboratories, only the summaries of the predicted ancestries are presented in the text below for all groups and additionally in Table 1 and Fig. 3. The full version of the final conclusions regarding the ancestry of the tested individuals can be found in Supplementary Files S13-S15. We provide the original report texts submitted by participants to demonstrate the complexity and range of styles shown in the reports. In Fig. 3, ancestry reports have been simplified for clarity and highlighted with the same colour codes used for the STRUCTURE and PCA plots. If an individual was reported to have admixed ancestry, shading is used for contributing populations inferred in the co-ancestry components. Reports mentioning possible recent admixture (first-generation admixture, parental) are highlighted in black. Samples reported as inconclusive are highlighted in white.

For the externally visible characteristics, out of 12 laboratories, nine performed the phenotype predictions - seven of them used the HIrisPlex-S guidelines provided by this panel's authors and two used updated / in house guidelines (Supplementary Files S16-S17). Four of the laboratories did not consider the impact of the EVC predictions on the inference of BGA, and five stated that the predictions were used to confirm BGA inferences.

3.2.1. *"Level 1" samples*

3.2.1.1. Ancestry. The simple unambiguous ancestry observed for the "Level 1" samples was reflected in the reports, as all 12 laboratories assigned the individuals to a single population: KS106 to East Asia, KS110 and KS153 to Europe (Fig. 3A). In the final reports for KS106, only nuances in the data summaries were observed, from laboratories reporting in a more "direct" way, e.g., "Individual of EAS origin" or "Belongs genetically to East Asian population", or were less direct, e.g., "Most likely has an East Asian biogeographic origin". The same applied to samples KS110 and KS153, e.g., "European ancestry" or "Most likely European ancestry" (Supplementary File S13).

3.2.1.2. Appearance. The *p-values* obtained for KS106 were interpreted in the same way by all nine laboratories reporting phenotype predictions, with the individual predicted to have: brown eyes, black hair and intermediate skin colour. Only one (Lab 2) reported intermediate to dark, which cannot be explained by the data (Supplementary File S17–18).

For KS110, all nine laboratories reported brown eyes, and most laboratories reported intermediate skin colour (some pale to intermediate). However, hair colour predictions showed more discrepant predictions, with HPS guidelines stating the obtained *p-values* should be interpreted as dark brown/black hair (the self-reported, natural colour of the individual was brown). This interpretation was given by four laboratories, of which three used the HPS guidelines and one used their own updated version. Other laboratories reported KS110 as: dark brown, brown, brown or dark blonde, brown to blond, using HPS guidelines, and blond or brown, using their own guidelines (Supplementary File S17–18).

Similar observations were made for KS153, for which all laboratories predicted blue eyes and pale to intermediate skin colour (one reported only intermediate). For hair colour, the observed *p-values* interpreted according to HPS guidelines suggested brown/dark brown hair (with the actual colour dark brown). This prediction was reported by five laboratories, of which four used HPS guidelines and one used updated guidelines. Other laboratories reported: brown, brown to blonde, blond or brown (Supplementary File S17–18).

For all three samples, the observed *p-values* for eye colour were higher than the 0.7 threshold (KS110 close to it, with 0.767). This was reflected by the confidence with which eye colour predictions were reported by all laboratories. For the skin colour of KS110 and KS153, the highest *p-values* were below 0.7, with $p = 0.681$ and $p = 0.563$,

Table 1
A short summary of the self-reported, genotyped and reported ancestries for all samples. For the "Level 1" samples with no observed admixture, the reported ancestries were consistent across all participating laboratories. As shown for "Level 2" and "Level 3" samples, the participants considered different possible ancestries, which are presented in descending order of frequency of reporting. The most popular answers are marked in **bold**. Note that, as shown in Fig. 3, many laboratories considered more than one possible ancestry per sample, hence the total number of reports is higher than the number of participants. More detailed genotyping and reporting results are presented in Figs. 1–3.

	Sample name	Self-reported ancestry	Genotyped ancestry	Reported ancestry
„Level 1" samples	KS106	China	East Asia	**12 labs: East Asia**
	KS110	Russia	Europe	**12 labs: Europe**
	KS153	Germany	Europe	**12 labs: Europe**
„Level 2" samples	KS102	Albania	Admixed	**9 labs: admixed (Europe & Middle East)** 7 labs: Europe 4 labs: Middle East 1 lab: admixed (Europe & South Asia) 1 lab: recent admixture
	KS126	Türkiye	Admixed	**5 labs: Middle East** **5 labs: admixed (Europe & Middle East & South Asia)** 4 labs: admixed (Europe & Middle East) 3 labs: recent admixture 2 labs: Europe 2 labs: admixed (Europe & South Asia) 1 lab: admixed (Middle East & South Asia) 1 lab: South Asia 1 lab: inconclusive
	KS180	Germany (M) / Türkiye (P)	Admixed	**7 labs: admixed (Europe & Middle East & South Asia)** 3 labs: admixed (Europe & South Asia) 3 labs: recent admixture 2 labs: Europe 2 labs: Middle East 2 labs: admixed (Europe & Middle East) 1 lab: South Asia 1 lab: admixed (Middle East & South Asia)
„Level 3" samples	KS71	Japan (M) / Germany (P)	Admixed	**8 labs: admixed (Europe & East Asia)** 3 labs: recent admixture 1 lab: admixed (Middle East & East Asia) 1 lab: admixed (South Asia & East Asia) 1 lab: South Asia 1 lab: inconclusive
	KS96	Brazil (M) / Germany (P)	Admixed	**3 labs: Middle East** **3 labs: recent admixture** 3 labs: admixture (Europe & Middle East & Africa) 2 labs: admixture (Europe & Africa) 2 labs: inconclusive
	KS9	South Africa (coloured)	Admixed	**5 labs: admixture (Africa & Middle East & Southeast Asia)** 4 labs: admixture (Africa & Middle East) 3 labs: Middle East 3 labs: inconclusive 1 lab: Africa 1 lab: admixed (Africa & Middle East & East Asia)

respectively, for intermediate and pale skin colour. Laboratories followed the guidelines, according to which only KS153 the second highest predicted category would have affected the final prediction (pale/intermediate) and this was reflected by predictions reported. KS110 was reported as intermediate by five of seven laboratories using the HPS guidelines.

For hair colour, in KS110 and KS153, the highest p-value (for brown) was below the 0.7 threshold. The principal discrepancy was evident in the *p-values* observed for hair shade, with the first sample exhibiting a p-value of 0.742 and the second displaying a p-value of 0.812. According to the HPS guidelines, this discrepancy should result in different predictions, namely dark brown/black and brown/dark brown, respectively. Labs 2 and 12 reported a lighter colour than that suggested by the HPS guidelines. The Lab 12 report mentioned that their predictions referred also to the second highest predicted p-value, namely blond (for both samples, around 30%). Lab 7 employed the same methodology, utilising their own updated guidelines. It is important to note that both individuals had natural hair colour lighter than predicted by the HPS.

3.2.2. *"Level 2" samples*

3.2.2.1. Ancestry. The complex ancestry patterns observed amongst the "Level 2" samples were reflected in the ancestry reports from all laboratories. As shown in Fig. 3B, some reports considered more than one prediction to indicate the ancestry for most samples.

KS102 was assigned to a single population, Europe, by Labs 9 and 11 using all markers, although Lab 9 suggested "a slight chance the individual is admixed and has a European male lineage and female lineage from further east (e.g., the Middle East)". Lab 5 (using all markers) considered two scenarios: European ancestry, or recent parental admixture between Europe and the Middle East. Labs 6, 7 and 8 (using only autosomal SNPs or autosomal and Y-SNPs) reported several possible scenarios: Europe, the Middle East, a Europe-Middle East admixture, or a Europe-South Asia admixture. The remaining laboratories reported the sample as being admixed between Europe and the Middle East. Among them, Lab 10 suggested "it is likely that the maternal linage comes from Europe.".

Reports for KS126 were even more complex, as only Labs 2, 3, 4 and 10 considered just the one possible scenario (Labs 4 and 10 using all markers) comprising admixture of Europe, Middle East and South Asia; while Lab 11 (all markers) considered recent admixture of Europe and Middle East. Their report stated that the individual „probably has a mixed biogeographical background of European (probably paternal) - Middle Eastern (probably maternal) origin". The other laboratories reported several possible scenarios: Europe or Middle East or South Asia or admixture between these contributing populations. Lab 12 (all markers) reported that the "predominant biogeographical origin is inconclusive" however they mentioned that recent admixture is possible".

Reports for sample KS180 were very similar, with most laboratories reporting the same possible ancestries for both samples. Lab 5 (all markers) considered only parental admixture for KS180, whereas for KS126, other possibilities included European or Middle Eastern ancestry. Similarly, Lab 9 (all markers) considered historical or recent admixture, which they had not reported for KS126.

A comparison of the reports for "Level 2" samples revealed a significant difference: only KS102 was considered to be European or European/Middle Eastern in origin, or an admixture of these two

populations. Two laboratories considered recent admixture as a second scenario to a European origin. KS102 was the only non-admixed European sample in this group with a self-declared, Albanian origin. As previously reported by other studies, a Middle East (West Asian) component is often observed among southern European populations [5,19,20].

Overall, KS126 and KS180 were reported to be admixed between Europe, the Middle East, and South Asia. Some laboratories considered parental admixture for both samples. For KS180, two laboratories reported only recent admixture as the sample's ancestry, correctly reflecting the individual's maternal German and paternal Turkish origins. In contrast, KS126 was of Turkish origin, with five laboratories considering Middle Eastern ancestry for this sample compared with two for KS180.

As can be seen from the reports, laboratories that used all markers usually considered fewer scenarios for each sample than those using only autosomal markers or a combination of autosomal and selected markers. Recent admixture was reported almost exclusively by laboratories using all data (although one laboratory reported it using autosomal markers only). Six laboratories analysed data from only one autosomal SNP assay (Supplementary File S12) but they reported the same or very similar ancestries to laboratories using data from all three panels. Lab 2 used only PT data analysed with Converge plus uniparental markers, with their reports being more general compared to the others. Lab 6 used only PT SNPs analysed with GenoGeographer and no additional markers or software, reporting three possible scenarios for each sample. Labs 7 and 8, using BT and PT data analysed with GenoGeographer, STRUCTURE and PCA (Supplementary File S12).

3.2.2.2. Appearance. The reports for KS102 were fully consistent regarding eye colour (brown) and almost identical for hair colour (dark brown to black) and skin colour (pale to intermediate or intermediate). The same predictions concerning eye and hair colour were reported for sample KS126, with the skin colour being reported as intermediate. Lab 2 using HPS guidelines predicted it to be intermediate to dark, however, the dark category had p-value of 0.136 which should not affect the prediction (0.15 is recommended). Predictions for KS180 were consistent for eye colour (brown) and skin colour (pale to intermediate), but not hair colour. In the group of seven laboratories that employed HPS guidelines for interpretation, there was a divergence in the reported results, which ranged from brown or blond / brown to dark brown. The same colours were predicted by laboratories utilising their own guidelines (Supplementary File S17–18).

The highest p-value for hair colour (brown) was below the 0.7 threshold for all three samples, but the reports are only consistent for KS102 and KS126 which were predicted adhering to HPS guidelines. For these samples, the hair shade was predicted as dark, with *p-values* of 0.940 and 0.754 respectively. The light shade predicted for KS180, together with blond colour predicted with ~25% probability, resulting again in some laboratories (using HPS or their own guidelines) reporting lighter hair coloured than suggested. Again, the natural hair colour of this individual was lighter than predicted by HPS (Supplementary File S17–18).

3.2.3. "Level 3" samples

3.2.3.1. Ancestry. The reports for "Level 3" samples reflected why this sample group was universally considered the most challenging to interpret. It is notable that in this group alone, a large proportion of laboratories decided to report ancestries as inconclusive (Fig. 3C).

For KS71, admixture of Europe and East Asia (or parental admixture of these two populations) was reported by all laboratories, except Lab 1 reporting it as inconclusive (using autosomal SNPs only). The Lab 1 report stated that the individual "doesn't present an African ancestry. The biogeographical origin of this individual could not be predicted more precisely." Lab 1 used only ET data analysed with STRUCTURE, PCA and GenoGeographer. The STRUCTURE results showed East Asian and European ancestry (46.3% and 34.4%) with minor South Asian and American components. In the PCA plot, KS71 was positioned on the edge of the South Asian reference cluster. GenoGeographer rejected all reference populations for this sample. Among the remaining laboratories using only autosomal SNP data, either two or three possible scenarios were reported, including European and East Asian admixture. The labs that used all markers or a combination of autosomal and maternal/paternal markers, were more direct in the reporting of the recent (parental) admixture of Europe and East Asia, although with some variation (Lab 5 and Lab 11 were less direct or more nuanced,). Lab 5 reported "one parent having East Asian ancestry. Under the hypothesis of admixture, it is possible the other parent has European ancestry – although the ME and South Asia cannot be excluded as origins for the second parent" and Lab 11 reported "likely to have a recent mixed biogeographical background of East Asian (probably maternal) – European (probably paternal) origin." These reports corresponded well with the donor's self-reported German paternal and Japanese maternal parentage.

For KS96, Labs 8 and 12 opted not to report ancestry (using all markers, and autosomal and Y-SNPs, respectively). Lab 8 reported that "prediction of biogeographic ancestry provides support against the unidentified individual's biogeographical ancestry being East Asia, America or Oceania. It cannot be clarified whether the ancestry is Africa, Europe, Middle East, South Asia or an admixture of these meta-populations." Lab 12 reported the individual is "most likely of admixed origin although it cannot be excluded that the ancestry is from a population not represented by the references database. East Asian or Oceania ancestry is not likely, but any other ancestry component cannot be excluded." Of the remaining laboratories that used all marker data, Labs 4 and 10 described the ancestry as admixed Europe, Africa and Middle East, and Labs 5, 9 and 11reported parental admixture from these same populations. Of the three laboratories using only autosomal SNPs, Lab 6 assigned the individual to one likely population, stating that it is "more likely if the investigated DNA comes from an individual from the Middle Eastern reference population than if the investigated DNA comes from an individual from the European, East Asian, South Asian, Oceanian, Native American, or African reference populations." This lab analysed only PT data with GenoGeographer. Labs 1 and 7 considered recent or historical admixture from Europe, the Middle East and Africa. The origin of the individual was mixed ancestry from paternal German and maternal Brazilian parentage.

Three of 12 laboratories did not report the ancestry of sample KS9 and all used only autosomal marker data (Labs 1, 6 and 7). Lab 1 reported that "the origin of this individual could not be predicted" and Lab 7 that "ancestry only from Europe, Africa or East Asia, respectively, is unlikely". The remaining laboratories considered four possible ancestries: Africa, Middle East, admixture of these two ancestries or admixture of Africa, Middle East and Southeast Asia. The latter was reported by all except one laboratory (all markers) and Lab 2 (autosomal markers, and mtDNA). Lab 9 wrote that the individual has "either ancestry from the North-East African/Middle East area or admixed ancestry with male ancestry from Africa and ancestry on the female side from somewhere in the Middle East/South Asia/South-East Asia region. Given the dark-to-black skin tone, the admixed ancestry option is the most likely", thus directly including EVC prediction results into the BGA report. The donor's self- described ancestry was South African "Coloured".

Both KS96 and KS9 presented challenges in interpreting historical admixture. The KS96 donor described themselves as half-German, half-Brazilian, but did not elaborate on their African maternal heritage. The admixture patterns, supported by the mitochondrial lineage L3e2a1a and the African X-chromosome, were interpreted by most laboratories as indicating European and African admixture. However, it is important to note that this maternal lineage, which is African, occurs at high frequencies in populations descended from Africans in the Americas, as

most of the captives in the Atlantic slave trade came from West and West-Central Africa [21]. The same applies to the mitochondrial lineage E1a1a1 observed for the South African individual KS9. The population of South Africa termed "Coloured" is admixed, with maternal lineages showing contributions from: Khoisan (indigenous southern African), Bantu-speaking Africans, Europeans (mainly Dutch, German, Portuguese) and South and Southeast Asians (brought via the Indian Ocean slave trade). Importantly, the Dutch East India Company (VOC) established the Cape Colony (1652), importing tens of thousands of enslaved people from Indonesia, Malaysia, Madagascar, India, and East Africa [22–24].

These two examples demonstrate that the analysis of uniparental lineages can reflect local and global distribution patterns created by historical events.

3.2.3.2. Appearance. For all "Level 3" samples the predicted eye colour, based on HPS guidelines, should be brown - successfully reported by all laboratories with confident and consistent predictions.

For KS71 and KS9, most laboratories reported black hair colour, with only a few reporting brown, black, or dark brown hair. For KS96, the reports were less consistent. According to the HPS guidelines, the predicted colour should be brown to dark brown. Only four laboratories using these guidelines reported this, with the others reporting brown to black or dark brown. Laboratories using their own guidelines reported the hair colour as brown or dark brown (Supplementary File S17–18).

The predicted skin colour was intermediate to dark for KS71 (one laboratory reported only intermediate), and dark to black for KS9, except for one laboratory which reported intermediate to black. This cannot be explained by the data (HPS guidelines were used). For KS96, the expected prediction should be dark skin, which was reported by six out of nine laboratories; the rest reported dark to intermediate skin colour (both HPS and their own guidelines were used). For this sample, the p-value for dark skin was 0.685 which from HPS guidelines should not be affected by the second highest category if pale or intermediate (intermediate with $p = 0.262$, Supplementary File S17–18).

As mentioned, for KS9, one laboratory reported that the predicted skin colour influenced the interpretation of ancestry, favouring admixed ancestry over a Middle Eastern ancestry.

3.3. Findings from participant feedback

The short questionnaire (Supplementary File S3) provided valuable insights into the challenges faced by participating laboratories, as well as important feedback on the study itself. Less complex questions for which identical answers were received are summarised in Fig. 4. These and other more complex questions and replies are discussed in more detail below.

Three of the 12 laboratories did not provide ancestry reports for KS9, and Fig. 4 shows this sample was identified as the most challenging to interpret. This was followed by all other samples from "Level 3", as well as those of "Level 2", where admixture patterns were very similar. When asked if any of the provided results were confusing and/or difficult to interpret, the most common answer was the interpretation of admixed ancestry, with some laboratories mentioning that the provided examples were more complex than the samples they are used to working with. Participants said that PCA patterns, particularly for admixed samples, were difficult to understand and that there are no recommendations on how to interpret such patterns. At the same time, without guidelines for the interpretation of STRUCTURE results, genetic cluster data was hard to analyse, and the limitations of this widely used population analysis algorithm were not outlined to users beforehand. Comments were also made about GenoGeographer analysis results, highlighting that LR and *p-values* were difficult to interpret without experience, especially when contradictory z-scores and likelihoods were observed for many samples. As Fig. 4 shows, at the same time, when asked which method of biparental SNP analysis they found most informative, participants listed all the methods discussed above, along with Converge, as the most useful tools. The feedback also mentioned that it would be interesting to see how using alternative reference datasets would affect the autosomal analyses outcome and subsequent data interpretation, especially for "Level 2" and "Level 3" samples.

This study offered participants the opportunity to combine phenotype and ancestry markers to make final ancestry predictions. One of the questions explored whether appearance influenced the conclusions about the biogeographical origin of the individuals studied. While most laboratories stated that it did not, some commented that it supported the interpretation of ancestry data. As previously highlighted, at least one laboratory mentioned that the phenotype prediction played a role in reporting the ancestry of sample KS9. Interestingly, it was observed that the interpretation of the *p-values* calculated by the HPS Webtool varied between laboratories, even when the guidelines recommended by the HPS authors were used. Eye colour was the only trait for which all the reports were fully concordant, but it should be noted that for this phenotype prediction all the observed *p-values* were higher than 0.7. The biggest discrepancies were observed for hair colour, particularly brown hair. Skin colour predictions were less varied, but still not fully concordant. Interestingly, the known negative effect of non-European co-ancestry on the predictive performance of HPS SNPs [25] was not a factor that was raised in discussions, nor did it prevent predictions being made by some participants.

Only six of the 12 laboratories decided to use the full range of genetic data provided to them, comprising autosomal, X- and Y-SNPs, and mitogenomes. Participants that analysed only the autosomal data stated that this aligns with the approach adopted in their own laboratory, and/or that they lack experience with non-autosomal markers or the capacity to incorporate them into their workflow. When asked which markers, besides autosomal SNPs, were the most useful for interpreting the data, maternal lineage mtDNA data was the most common answer, followed by paternal lineage Y-SNP data and X-chromosome data. It is worth noting that although laboratories lacked prior experience with X-SNP data, it was found to be very useful, particularly for admixed individuals and males in general.

For the uniparental markers, participants were provided only with observed maternal and parental lineages and were free to choose the tools with which to interpret these inferred lineages. The most chosen databases were EMPOP (for maternal lineages only), YHRD (for paternal lineages only), followed by literature searches and other online resources for both markers. However, it was mentioned that there are currently no guidelines available for using uniparental markers as tools to support ancestry inference. This is evident from the fact that YHRD was the second most popular choice, despite this database not being designed for drawing ancestry conclusions, unlike EMPOP. This is especially pertinent given exercise data was based on Y-SNPs rather than Y-STRs, which form the basis of the YHRD database. Although Y-SNP data are presented in YHRD, it is limited and largely uninformative. The Haplogroup Browser tool in EMPOP also has limitations, as the estimated lineage frequencies are based on the data available in the database. However, EMPOP is a considerably more extensive database for mtDNA data than the information on Y-SNPs in YHRD. Participants also mentioned that a literature search is both time-consuming and difficult for those with no experience of interpreting haploid markers other than haplotypes and their frequencies, and that this requires a population genetics perspective.

4. Concluding remarks

It is evident from all the participant's ancestry analyses reported back to the exercise organisers, and summarised here, that the unadmixed "Level 1" samples (KS106; KS110; KS153), presented genetic marker patterns allowing unequivocal ancestry inferences to be made. This contrasts with the "Level 2" and "Level 3" sample groups,

A)

Origin of the individuals			
	Sample 106	Sample 110	Sample 153
	China	Russia	Germany
Reports from the participants			
Lab 1	EA	EU	EU
Lab 2	EA	EU	EU
Lab 3	EA	EU	EU
Lab 4	***EA***	***EU***	***EU***
Lab 5	***EA***	***EU***	***EU***
Lab 6	EA	EU	EU
Lab 7	EA	EU	EU
Lab 8	EA	EU	EU
Lab 9	***EU***	***EU***	***EU***
Lab 10	***EA***	***EU***	***EU***
Lab 11	***EU***	***EU***	***EU***
Lab 12	***EA***	***EU***	***EU***

B)

Origin of the individuals				
	Sample 102	Sample 126	Sample 180	
	Albania	Türkiye	Germany (M)	Türkiye (P)
Reports from the participants				
Lab 1	EU / admixed (EU / ME)	ME / admixed (EU / ME / SA)	ME / admixed (EU / ME / SA)	
Lab 2	admixed (EU / ME)	admixed (EU / ME)	admixed (EU / SA)	
Lab 3	admixed (EU / ME)	admixed (EU / ME / SA)	admixed (EU / ME / SA)	
Lab 4	***admixed (EU / ME)***	***admixed (EU / ME / SA)***	***admixed (EU / ME / SA)***	
Lab 5	***EU / recent admixture****	***EU / ME / recent admixture****	***recent admixture****	
Lab 6	EU / ME / admixed (EU / ME) / admixed (EU / SA)	admixed (EU / ME) / admixed (EU / SA) / admixed (ME / SA)	admixed (EU / ME) / admixed (EU / SA) / admixed (ME / SA)	
Lab 7	EU / ME / admixed (EU / ME)	admixed (EU / ME) / ME / admixed (EU / SA)	admixed (EU / ME) / EU / admixed (EU / SA)	
Lab 8	EU / ME / admixed (EU / ME)	EU / ME / SA / admixed (EU / ME / SA)	EU / ME / SA / admixed (EU / ME / SA)	
Lab 9	***EU***	***ME / admixed (EU / ME)***	***admixed (EU / ME / SA) / recent admixture****	
Lab 10	***admixed (EU / ME)***	***admixed (EU / ME / SA)***	***admixed (EU / ME / SA)***	
Lab 11	***EU***	***recent admixture****	***recent admixture****	
Lab 12	***EU and / or ME***	***inconclusive / recent admixture****	***admixed (EU / ME / SA)***	

C)

Origin of the individuals				
	Sample 71		Sample 96		Sample 9
	Germany (P)	Japan (M)	Germany (P)	Brazil (M)	South Africa (coloured)
Reports from the participants					
Lab 1	inconclusive		ME / admixed (EU / ME)		inconclusive
Lab 2	admixed (EU / EA)		admixed (EU / AFR)		admixed (AFR / ME / SEA)
Lab 3	admixed (EU / EA)		admixed (EU / AFR)		admixed (AFR / ME)
Lab 4	***admixed (EU / EA)***		***admixed (EU / ME / AFR)***		***admixed (AFR / ME / SEA)***
Lab 5	***recent admixture****		***recent admixture****		***ME / admixed (AFR/ ME) / admixed(AFR/ME/EA)***
Lab 6	admixed (EU / EA) / admixed (ME / EA) / admixed (SA / EA)		ME		inconclusive
Lab 7	admixed (EU / EA)	SA	ME / admixed (EU / ME / AFR)		inconclusive
Lab 8	admixed (EU / EA)		inconclusive		ME / AFR / admixed (AFR / ME)
Lab 9	***recent admixture****		***recent admixture****		***ME / admixed (AFR / ME / SEA)***
Lab 10	***admixed (EU / EA)***		***admixed (EU / ME / AFR)***		***admixed (AFR / ME / SEA)***
Lab 11	***recent admixture****		***recent admixture****		***admixed (AFR / ME)***
Lab 12	***admixed (EU / EA)***		***inconclusive***		***admixed (AFR / ME / SEA)***

Fig. 3. Summary of each laboratory's reported ancestry all samples (A for "Level 1", B for "Level 2" and C for "Level 3"); with the same colour labels used by STRUCTURE. More detailed results are presented in Supplementary File S14. Laboratories using all markers for their interpretation are marked in ***bold italics***; those using only autosomal markers are underlined; and the rest used combinations (detailed in Supplementary File S12). The self-declared, grandparental ancestry is provided at the top of the table (M – maternal; P – paternal). Samples associated with a single population are given one colour, while admixed donors (historical) have colour shading representing the populations involved. *The term "recent admixture" refers to reporting of parental co-ancestry; more detailed reporting is provided in Supplementary File S14.

Fig. 4. Selected questions and answers from the questionnaire sent to participants. Bar plots show all the answer alongside the number of times it was given.

characterised by a degree of inconsistency in the population analysis data provided, notably the STRUCTURE and GenoGeographer outputs. These algorithms tend to reveal their limitations when an individual's co-ancestry components originate from less differentiated populations or have admixed ancestries themselves. If the exercise had just used samples such as those of "Level 1", there would not be enough data interpretation challenges for the participants. More importantly, the simple self-reported ancestries of KS106; KS110; KS153 would not be reflective of real-world forensic samples in many countries, where modern urban demographic profiles often show a wide range of population admixture contributors and ratios. Overall, the reporting from "Level 1" was the most consistent, however it is important to mention that even within this group, the reports slightly differed in terms of how the ancestry was verbally reported and how the nomenclature used could affect law enforcement's understanding of the report's statements. For example, police could recognise a different reporting "value" if a phrase like "most likely East Asian" was used instead of "an individual of East Asian origin". One of the key messages from participants' feedback was the need for a more thorough exploration of the limitations, data-interpretation thresholds and expected precision of the four main autosomal SNP data analysis systems used in the exercise. This was considered particularly important for STRUCTURE analysis, especially when interpreting samples with admixed backgrounds. One outcome from this feedback has been the introduction of reference population cloud centroids in the Snipper PCA plots (the mid-cloud solid colour diamonds in Fig. 2), with accompanying distance-to-centroid estimates. This allows users to more confidently assess whether the sample's position is close to, or distant from the nearest reference population clouds in the plot.

Although half of the twelve participants used all the marker data provided, three of the other six relied solely on autosomal markers, and this choice to limit the data accessed may have been made with a view to the marker systems planned to be used, or in use in their laboratories. It should also be noted that implementing many extra DNA analyses requires budgetary support, and that the amount of DNA extract available is often limited, particularly following prior STR analysis. Nevertheless, reliance on autosomal SNPs alone can reduce the level of detail that can be obtained for the inference of admixed ancestry or the identification of populations slightly divergent from those used for the reference data (e.g., Albania vs. the mainly Western Europeans represented by 1000 Genomes sampling). We note that in the case of KS96 the three autosomal SNPs-only laboratories did not infer European ancestry, in contrast to the others. For the most challenging sample of KS9, these laboratories were the only ones reporting inconclusive data analyses. One final point about using the full breadth of marker data provided was the obvious advantage we observed of being able to estimate the "ancestry" of the X-chromosome(s) independently of the other marker sets. Although an entirely novel marker type for all participants, use of this data not only provided an informative point of comparison with the STRUCTURE patterns in each sample but potentially aided the interpretation of the mtDNA maternal lineage (e.g., the KS9 E1a1a1 lineage having a Southeast Asian geographic distribution that contrasted with the other

ancestry findings for this sample). Further integration of X-SNPs into forensic panels beyond ET will clearly enhance ancestry analysis in the future and small sets of these SNPs should amplify well when included in large multiplexes of autosomal and Y-SNPs.

While it is tempting to compare the three different SNP assays' STRUCTURE cluster plots and membership coefficients for the "Level 2 and 3" samples with complex ancestries, our experience has been that genetic cluster analysis can produce inconsistent and irregular patterns when analysing populations with complex historical admixture that has created imbalanced distribution of variation across a geographic region. This effect is accentuated when using SNPs in the numbers limited by forensic multiplex size constraints. Two notable examples being Brazil, with three main contributing ancestries and Türkiye, with predominantly European variation in the west and predominantly Middle East variation in the east. Therefore, it is not surprising that KS126, KS180 and KS96 had the most inconsistent genetic cluster patterns from independent STRUCTURE analysis of PT, BT and ET SNP genotypes. Furthermore, we feel it is not instructive at this stage to focus on differences in the inferred admixture components between these panels unless the donors' DNA can be analysed with a much more extensive population marker panel such as the Affymetrix Origins SNP array [18].

The main intended outcome of this exercise was to move towards consensus within EDNAP on the frameworks for reporting ancestry and phenotype inferences from forensic samples. However, this is likely to be challenging given the differences in phraseology and levels of caution used amongst participants. An approach which is often discussed is to consider not to report an ancestry when the population analysis does not support a strong inference for a forensic sample [26]. Furthermore, the use of donors in this exercise with complex (but real-world) backgrounds has highlighted the shortcomings of admixture analysis using the current SNP-based forensic ancestry panels, constrained in their scale by multiplex size considerations. As SNP panels get larger (and we note the increased ancestry informativeness of FORCE, Kintelligence and MPSplex SNP sets in a recent study [18]), the ability to detect and characterise admixture is likely to improve, particularly when combined with mtDNA, Y-SNPs and X-SNPs. However, even large-scale assays will not solve all potential problems, as demonstrated by the KS96 sample. Our heritage and how we self-report it does not have to be reflected in our biogeographical ancestry; therefore, it does not have to correspond with the data provided by any SNP assay. It is highly likely that whole genome sequencing of the KS96 sample would reveal more complex admixture patterns, possibly including some Native American component. Nevertheless, the maternal admixture would most likely remain predominantly African, reflecting the individual's African-Brazilian ancestry. The individual identified as simply Brazilian and did not mention in the questionnaire that their maternal ancestors were necessarily descendants from African immigrants.

Participant feedback has also emphasised the need for more carefully constructed interpretative guidelines for the statistical and graphical output of all the population analysis systems we provided to them. These guidelines will need to be in place before EDNAP as a collaborative group starts the process of standardising forensic population analysis and the reporting of ancestry inferences to investigators. We also note that, despite set guidelines for the interpretation of HPS *p-values*, there was still divergent approaches to their inference, and even use of in-house guidelines by some participants.

Finally, we acknowledge that the outcome of the exercise helped to answer some questions, but also raised many new ones. We believe it is too early to make official recommendations for forensic practitioners reporting the ancestry inferences they make, and therefore we plan to organise a follow-up exercise in which the lessons learned from this study can be implemented.

Editorial declaration statement

Given his role as Editor-in-Chief, Walther Parson had no involvement in the peer review of this article and had no access to information regarding its peer review. Full responsibility for the editorial process for this article was delegated to another journal editor.

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

CRediT authorship contribution statement

Torben Tvedebrink: Methodology. **Christopher Phillips:** Writing – review & editing, Writing – original draft, Supervision, Project administration, Data curation, Conceptualization. **Siri Aili Fagerholm:** Writing – review & editing, Methodology, Formal analysis, Data curation. **Walther Parson:** Writing – review & editing, Visualization, Supervision, Project administration, Conceptualization. **Lena Ewers:** Writing – review & editing, Methodology, Formal analysis, Data curation. **Natalie Weiler:** Methodology, Formal analysis, Data curation. **Ronny Decorte:** Methodology, Formal analysis, Data curation. **Catarina Xavier:** Methodology, Formal analysis, Data curation. **Cordula Haas:** Writing – review & editing, Methodology, Formal analysis, Data curation. **Mario Gysi:** Writing – review & editing, Methodology, Formal analysis, Data curation. **Kris van der Gaag:** Methodology, Formal analysis, Data curation. **Oscar Garcia:** Methodology, Formal analysis, Data curation. **Leire Palencia-Madrid:** Writing – review & editing, Methodology, Formal analysis, Data curation. **Daniel Kling:** Writing – review & editing, Methodology, Formal analysis, Data curation. **Clemence Hollard:** Writing – review & editing, Methodology, Formal analysis, Data curation. **Maarja Sadam:** Writing – review & editing, Methodology, Formal analysis, Data curation. **Marta Diepenbroek:** Writing – original draft, Methodology, Formal analysis, Data curation, Conceptualization. **Jorge Ruiz-Ramírez:** Methodology, Formal analysis, Data curation, Conceptualization. **María de la Puente:** Methodology, Formal analysis, Data curation, Conceptualization. **Vania Pereira:** Writing – review & editing, Methodology, Formal analysis, Data curation. **Monika Stoljarova-Bibb:** Writing – review & editing, Methodology, Formal analysis, Data curation. **Christina Amory:** Methodology, Formal analysis, Data curation. **Adam Staadig:** Writing – review & editing, Methodology, Formal analysis, Data curation. **Natasha Arora:** Writing – review & editing, Methodology, Formal analysis, Data curation. **Helle Smidt Mogensen:** Writing – review & editing, Methodology, Formal analysis, Data curation. **Katja Anslinger:** Writing – review & editing, Supervision, Formal analysis. **Maja Sidstedt:** Writing – review & editing, Methodology, Formal analysis, Data curation. **António Amorim:** Formal analysis. **Heloisa Costa:** Methodology, Formal analysis, Data curation. **Caroline Bouakaze:** Writing – review & editing, Methodology, Formal analysis, Data curation. **Andreas Tillmar:** Writing – review & editing, Methodology, Formal analysis, Data curation. **Bram Bekaert:** Methodology, Formal analysis, Data curation. **Denise Syndercombe Court:** Methodology, Formal analysis, Data curation. **David Ballard:** Methodology, Formal analysis, Data curation.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. The author is an Editorial Board Member/Editor-in-Chief/Associate Editor/Guest Editor for this journal and was not involved in the editorial review or the decision to publish this article.

Acknowledgments

The authors would like to thank Ricky Ansell, Barbara Dell'Amico and Mathieu Gabut for their valuable contributions and insightful discussions that supported this work.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.fsigen.2026.103576.

References

[1] M. Kayser, W. Branicki, W. Parson, C. Phillips, Recent advances in Forensic DNA phenotyping of appearance, ancestry and age, Forensic Science International Genetics 65 (2023) 102870, https://doi.org/10.1016/j.fsigen.2023.102870.

[2] S. Ingold, G. Dørum, E. Hanson, A. Berti, W. Branicki, P. Brito, P. Elsmore, et al., Body fluid identification using a targeted mRNA massively parallel sequencing approach – Results of a EUROFORGEN/EDNAP collaborative exercise, Forensic Science International Genetics 34 (2018) 105–115, https://doi.org/10.1016/j.fsigen.2018.01.002.

[3] S. Ingold, G. Dørum, E. Hanson, D. Ballard, A. Berti, K.B. Gettings, F. Giangasparo, et al., Body fluid identification and assignment to donors using a targeted mRNA massively parallel sequencing approach – Results of a second EUROFORGEN / EDNAP collaborative exercise, Forensic Science International Genetics 45 (2020) 102208, https://doi.org/10.1016/j.fsigen.2019.102208.

[4] N.V. Hänggi, A. Amorim, H.A. Costa, J.D. Andersen, M.L. Kampmann, C. Courts, M. Neis, et al., mRNA profiling and donor association of mock casework samples: Results of a 3rd and 4th EDNAP collaborative exercise, Forensic Science International Genetics 79 (2025) 103303, https://doi.org/10.1016/j.fsigen.2025.103303.

[5] M. Diepenbroek, B. Bayer, K. Schwender, R. Schiller, J. Lim, R. Lagacé, K. Anslinger, Evaluation of the Ion AmpliSeq™ PhenoTrivium panel: MPS-based assay for ancestry and phenotype predictions challenged by casework samples, Genes 11 (12) (2020) 1398, https://doi.org/10.3390/genes11121398.

[6] Xavier, C. de La Puente, M. Mosquera-Miguel, A. Freire-Aradas, A. Kalamara, V. Vidaki, A. Gross, et al., Development and validation of the VISAGE AmpliSeq basic tool to predict appearance and ancestry from DNA (Article), Forensic Science International Genetics 48 (2020) 102336, https://doi.org/10.1016/j.fsigen.2020.102336.

[7] C. Xavier, M. de la Puente, A. Mosquera-Miguel, A. Freire-Aradas, V. Kalamara, A. Ralf, A. Revoir, et al., Development and inter-laboratory evaluation of the VISAGE Enhanced Tool for Appearance and Ancestry inference from DNA (Article), Forensic Science International Genetics 61 (2022) 102779, https://doi.org/10.1016/j.fsigen.2022.102779.

[8] L. Chaitanya, K. Breslin, S. Zuñiga, L. Wirken, E. Pośpiech, M. Kukla-Bartoszek, T. Sijen, et al., The HIrisPlex-S system for eye, hair and skin colour prediction from DNA: introduction and forensic developmental validation, Forensic Science International Genetics 35 (2018) 123–135, https://doi.org/10.1016/j.fsigen.2018.04.004.

[9] C. Santos, M. Fondevila, D. Ballard, R. Banemann, A.M. Bento, C. Børsting, W. Branicki, et al., Forensic ancestry analysis with two capillary electrophoresis ancestry informative marker (AIM) panels: Results of a collaborative EDNAP exercise, Forensic Science International Genetics 19 (2015) 56–67, https://doi.org/10.1016/j.fsigen.2015.06.004.

[10] L. Chaitanya, S. Walsh, J.D. Andersen, R. Ansell, K. Ballantyne, D. Ballard, R. Banemann, et al., Collaborative EDNAP exercise on the IrisPlex system for DNA-based prediction of human eye colour, Forensic Science International Genetics 11 (1) (2014) 241–251, https://doi.org/10.1016/j.fsigen.2014.04.006.

[11] C. Xavier, M. Eduardoff, C. Strobl, W. Parson, SD quants — Sensitive detection tetraplex-system for nuclear and mitochondrial DNA quantification and degradation inference, Forensic Science International Genetics 42 (2019) 39–44, https://doi.org/10.1016/j.fsigen.2019.06.004.

[12] P.H. Sudmant, T. Rausch, E.J. Gardner, R.E. Handsaker, A. Abyzov, J. Huddleston, Y. Zhang, et al., An integrated map of structural variation in 2,504 human genomes, Nature 526 (7571) (2015) 75–81, https://doi.org/10.1038/nature15394.

[13] A. Bergström, S.A. McCarthy, R. Hui, M.A. Almarri, Q. Ayub, P. Danecek, Y. Chen, et al., Insights into human genetic variation and population history from 929 diverse genomes, Article aay5012, Science 367 (6484) (2020), https://doi.org/10.1126/science.aay5012.

[14] J. Ruiz-Ramírez, M. de la Puente, C. Xavier, A. Ambroa-Conde, J. Álvarez-Dios, A. Freire-Aradas, A. Mosquera-Miguel, et al., Development and evaluations of the ancestry informative markers of the VISAGE Enhanced Tool for Appearance and Ancestry (Article), Forensic Science International Genetics 64 (2023) 102853, https://doi.org/10.1016/j.fsigen.2023.102853.

[15] K.H. Cheung, M.V. Osier, J.R. Kidd, A.J. Pakstis, P.L. Miller, K.K. Kidd, ALFRED: An allele frequency database for diverse populations and DNA polymorphisms, Nucleic Acids Res. 28 (1) (2000) 361–363, https://doi.org/10.1093/nar/28.1.361.

[16] T. Tvedebrink, P.S. Eriksen, H.S. Mogensen, N. Morling, GenoGeographer – A tool for genogeographic inference, Forensic Science International Genetics. Supplement Series 6 (2017) e463–e465, https://doi.org/10.1016/j.fsigss.2017.09.196.

[17] H.S. Mogensen, T. Tvedebrink, C. Børsting, V. Pereira, N. Morling, Ancestry prediction efficiency of the software GenoGeographer using a z-score method and the ancestry informative markers in the Precision ID Ancestry Panel, Forensic Science International Genetics 44 (2020) 102154, https://doi.org/10.1016/j.fsigen.2019.102154.

[18] M. de la Puente, L. Casanova-Adán, J. González-Bao, J. Pardo-Seco, A. Mosquera-Miguel, A. Ambroa-Conde, J. Ruiz-Ramírez, et al., Evaluating the effect of marker panel sizes on estimation of bio-geographical co-ancestry proportions, Forensic Science International Genetics 78 (2025) 103275, https://doi.org/10.1016/j.fsigen.2025.103275.

[19] Lazaridis, I. Alpaslan-Roodenberg, S, A. Acar, A. Açıkkol, A. Agelarakis, L. Aghikyan, U. Akyüz, et al., The genetic history of the Southern Arc: A bridge between West Asia and Europe, Science 377 (6609) (2022) eabm4247, https://doi.org/10.1126/science.abm4247.

[20] S. Sarno, A. Boattini, L. Pagani, M. Sazzini, S. De Fanti, A. Quagliariello, G. A. Gnecchi Ruscone, et al., Ancient and recent admixture layers in Sicily and Southern Italy trace multiple migration routes along the Mediterranean, Sci. Rep. 7 (2017) 1984, https://doi.org/10.1038/s41598-017-01802-4.

[21] M.H. Gouveia, V. Borda, T.P. Leal, R.G. Moreira, A.W. Bergen, F.S.G. Kehdy, I. Alvim, et al., Origins, admixture dynamics, and homogenization of the African gene pool in the Americas, Mol. Biol. Evol. 37 (6) (2020) 1647–1656, https://doi.org/10.1093/molbev/msaa033.

[22] L. Quintana-Murci, C. Harmant, H. Quach, O. Balanovsky, V. Zaporozhchenko, C. Bormans, et al., Strong maternal Khoisan contribution to the South African coloured population: A case of gender-biased admixture, Am. J. Human. Genet. 86 (4) (2010) 611–620, https://doi.org/10.1016/j.ajhg.2010.02.014.

[23] I. Lankheet, K.J. van der Gaag, H. Meijers-Heijboer, K.J. Duran, M. Kayser, Wide-scale geographical analysis of genetic ancestry in the South African Coloured population, BMC Biol. 23 (2025) 260, https://doi.org/10.1186/s12915-025-02317-5.

[24] South African History Online. (n.d.). History of slavery and early colonisation in South Africa. ⟨https://www.sahistory.org.za/article/history-slavery-and-early-colonisation-south-africa⟩ (accessed 16 Oct 2025).

[25] A. Freire-Aradas, Y. Ruiz, C. Phillips, O. Maroñas, J. Söchtig, A.G. Tato, et al., Exploring iris colour prediction and ancestry inference in admixed populations of South America, Forensic Science International Genetics 13 (2014) 3, https://doi.org/10.1016/j.fsigen.2014.06.007.

[26] C. Phillips, L. Prieto, M. Fondevila, A. Salas, A. Gómez-Tato, J. Álvarez-Dios, A. Alonso, et al., Ancestry analysis in the 11-M Madrid bomb attack investigation, PLOS ONE 4 (8) (2009) e6583, https://doi.org/10.1371/journal.pone.0006583.