

Research paper



## Predicting eye and hair colour in a Norwegian population using Verogen's ForenSeq™ DNA signature prep kit

Nina Mjølunes Salvo<sup>a,\*</sup>, Kirstin Janssen<sup>a</sup>, Maria Kristine Kirsebom<sup>a</sup>, Olivia Strunge Meyer<sup>b</sup>, Thomas Berg<sup>a</sup>, Gunn-Hege Olsen<sup>a</sup>

<sup>a</sup> Centre for Forensic Genetics, Department of Medical Biology, Faculty of Health Sciences, UiT - The Arctic University of Norway, Norway

<sup>b</sup> Section of Forensic Genetics, Department of Forensic Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Denmark

### ARTICLE INFO

#### Keywords:

Genotyping performance  
Massively parallel sequencing  
Genetic prediction  
Forensic DNA phenotyping  
Human pigmentation  
HirisPlex

### ABSTRACT

Prediction of eye and hair colour from DNA can be an important investigative tool in forensic cases if conventional DNA profiling fails to match DNA from any known suspects or cannot obtain a hit in a DNA database. The HirisPlex model for simultaneous eye and hair colour predictions was developed for forensic usage. To genotype a DNA sample, massively parallel sequencing (MPS) has brought new possibilities to the analysis of forensic DNA samples. As part of an in-house validation, this study presents the genotyping and predictive performance of the HirisPlex SNPs in a Norwegian study population, using Verogen's ForenSeq™ DNA Signature Prep Kit on the MiSeq FGx system and the HirisPlex webtool. DNA-profiles were successfully typed with DNA input down to 125 pg. In samples with DNA input < 125 pg, false homozygotes were observed with as many as 92 reads. Prediction accuracies in terms of AUC were high for red (0.97) and black (0.93) hair colours, as well as blue (0.85) and brown (0.94) eye colours. The AUCs for blond (0.72) and brown (0.70) hair colour were considerably lower. None of the individuals was predicted to have intermediate eye colour. Therefore, the error rates of the overall eye colour predictions were 37% with no predictive probability threshold (pmax) and 26% with a probability threshold of 0.7. We also observed that more than half of the incorrect predictions were for individuals carrying the rs12913832 GG genotype. For hair colour, 65% of the individuals were correctly predicted when using the highest probability category approach. The main error was observed for individuals with brown hair colour that were predicted to have blond hair. Utilising the prediction guide approach increased the correct predictions to 75%. Assessment of phenotype-genotype associations of eye colours using a quantitative eye colour score (PIE-score), revealed that rs12913832 AA individuals of Norwegian descent had statistically significantly higher PIE-score (less brown eye colour) than individuals of non-northern European descent. To our knowledge, this has not been reported in other studies. Our study suggests that careful assessment of the target population prior to the implementation of forensic DNA phenotyping to case work is beneficial.

### 1. Introduction

Forensic DNA phenotyping (FDP) is an emerging DNA intelligence area for use in forensic genetics [1]. When conventional DNA profiling fails to match DNA from any known suspects or cannot obtain a hit in a DNA database, FDP can aid police investigations by predicting a person's externally visible characteristics from DNA alone [2,3]. Eye and hair colour are prominent visible traits that are often used to describe an individual. These traits are also highly heritable and are associated with relatively few genes, which makes them suitable for FDP [2].

DNA variants found in the regions of pigmentation genes are widely

used to predict eye and hair colour [4–8], including FDP tests in forensic casework [9,10]. One of the most popular models for these tests, namely HirisPlex, is a simultaneous eye and hair colour prediction system [11–13]. The HirisPlex DNA marker panel consists of 23 SNPs and one insertion/deletion (INDEL), of which six (referred to as IrisPlex SNPs) are used for eye colour prediction and 22 are used for hair colour prediction [8]. The model predicts colour into categories and reports prediction probability values (p-values) of blue, intermediate and brown eye colour, as well as blond, brown, black and red hair colour. For hair colour, light and dark shade is also reported. IrisPlex is reported to have very good accuracies for blue and brown eye colours, whereas the

\* Correspondence to: UiT – The Arctic University of Norway, Post Box 6050, 9037 Tromsø, Norway.

E-mail address: [nina.mjolsnes@uit.no](mailto:nina.mjolsnes@uit.no) (N.M. Salvo).

<https://doi.org/10.1016/j.fsigen.2021.102620>

Received 7 May 2021; Received in revised form 24 September 2021; Accepted 21 October 2021

Available online 24 October 2021

1872-4973/© 2021 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

accuracy for intermediate eye colour is considerably lower [13]. This is because the model is mainly dependent on the SNP rs12913832 in *HERC2*, which is known as the most important SNP for prediction of blue and brown coloured eyes [4,5,14–17]. The SNP rs12913832 is located upstream of the pigmentation gene *OCA2* and functions as an enhancer. The ancestral rs12913832 A-allele enhances transcription of *OCA2*, and thereby increases melanin production. In contrast, the derived rs12913832 G-allele has the opposite effect on melanin production [18]. Individuals with rs12913832 AA and AG are expected to have brown eyes, and individuals with rs12913832 GG are expected to have blue eyes [19,20]. The remaining five IrisPlex SNPs have only minor additive effects on the eye colour predictions [17]. The genetic basis of hair colour is more complex than iris colour. Although rs12913832 is also associated with hair colour, this association has been shown to be much weaker than for eye colour [8]. Red hair has been extensively studied and is strongly associated with mutations in *MC1R*, often in a homozygous or a compound-heterozygous state [7,21,22]. For the other hair colours ranging from blond via brown to black, predictions are made based on combinations of the 22 DNA variants with various degrees of association [8]. Hence, the hair colour prediction model is reported to have highest accuracy for red hair colour, followed by slightly lower accuracies for black, blond and brown hair [8].

In addition to an accurate prediction model, FDP also depends on a reliable typing method. Because forensic trace samples often contain small amounts of DNA and/or degraded DNA, the typing method needs to be suitable for short DNA fragments and low DNA input amounts. In this study, the HIRisPlex SNPs are typed by massively parallel sequencing (MPS, also known as next generation sequencing, NGS), using the commercially available ForenSeq™ DNA Signature Prep Kit on the MiSeq FGx system (Verogen, San Diego, CA). Primer mix B supplied with the kit contains primers to multiplex 231 forensically relevant genetic markers, including the 24 HIRisPlex SNPs. This kit has been demonstrated to be robust and to produce reliable results with 250 pg DNA [23, 24]. Compared to traditional PCR-CE methods, the application of MPS has brought new possibilities to the analysis of forensic DNA samples by enabling simultaneous typing of hundreds of markers in one analysis [25]. Thus, markers for FDP can be genotyped together with ancestry informative marker (AIM) SNPs, human identification (HID) SNPs and short tandem repeats (STR) used for standard DNA profiling. Thereby, several major issues in relation to a forensic DNA sample can be addressed in a single analysis.

As part of an in-house validation of the FDP analysis of eye and hair colour with the ForenSeq™ DNA Signature Prep Kit and the HIRisPlex webtool, we provide data on both technical- and predictive performance of the HIRisPlex SNPs in a Norwegian study population. In this population, individuals with intermediate or blue eyes as well as blond or brown hair are common. Although HIRisPlex is regarded to be an informative and robust model for predicting eye and hair colour, the precise genotype-phenotype associations for these traits are yet to be identified [8,26–29]. Supplementing information on inaccurate predictions of genotype-phenotype combinations might be educational for the interpretation of prediction results in a specific population, as well as helpful when searching for new candidate markers. Other forensic laboratories who are considering implementing FDP in their analysis repertoire might also benefit from the evaluation presented here.

## 2. Materials and methods

### 2.1. Samples

Blood samples were collected from 540 unrelated volunteers residing in Tromsø and Bodø, northern Norway from 2015 to 2017. All samples were collected with fully informed consent and subsequently anonymised. The project is approved by the Faculty of Health Sciences, UiT – The Arctic University of Norway (reference number 2021/2034). Digital photographs of the participants' eyes were taken (see below).

Photographs from 519 individuals had sufficient quality to be used for further evaluation of eye colour predictions. The participants also self-reported their natural hair colour at the age around 20, birthplace and grandparents' ancestry. All 540 individuals were used for evaluation of hair colour predictions. Of the 519 individuals included for eye colour predictions, 480 self-reported to be of European descent (all four grandparents were Europeans), of which 441 were Scandinavian (424 Norwegian). The remaining 39 individuals were of non-European, admixed or unknown ancestry. Of the 21 individuals that were only included for hair colour predictions, 20 were European, of which 18 were Scandinavian (17 Norwegian) and one individual was non-European.

### 2.2. Eye and hair colour phenotyping

High resolution photographs of eyes were taken with a Canon EOS 5D Mark III camera using a Canon EF 100 mm f/2.8 L Macro IS USM lens and a MT-24EX Macro Twin Lite Flash. The photographs were taken at approximately 5–10 cm in "Raw" format with ISO 100, shutter 1/100, AV 22, exposure compensation + 1 and manual focus. According to Andersen et al. [30], the white balance of "Raw" format photographs was changed to "Flash" using Digital Photo Professional 4 version 4.10.40.0 (Canon INC).

For qualitative determination of eye colour, nine untrained individuals were asked to intuitively assign each photograph to one of four categories (blue, intermediate-blue, intermediate-brown and brown eyes), similar to Meyer et al. [31]. To decide the final eye colour for each photograph, a majority-vote was employed. The IrisPlex model reports eye colour according to three categories (blue, intermediate and brown). Therefore, intermediate-blue and intermediate-brown, which includes all eye colours neither perceived as blue nor brown, were afterwards grouped into one category (intermediate). A total of 52% (n = 269) of individuals were categorised as blue-eyed, 34% (n = 179) as intermediate eye coloured (25% intermediate blue and 9% intermediate brown) and 14% (n = 71) as brown-eyed (see Fig. 2 in Section 3.2). This eye colour categorisation was further used as reference for assessment of eye colour predictions. For 31% (n = 163) of the photographs, all the observers agreed completely on the eye colour. For the remaining 69% (n = 356), on average two observers disagreed with the concluded category, demonstrating that perception of eye colours varies. The vast majority of the disagreements were between the categories blue and intermediate-blue, as well as brown and intermediate-brown.

For an objective evaluation of eye colour, a quantitative eye colour score (Pixel Index of the Eye (PIE)- score) was calculated for each individual eye photograph using the custom-made Digital Analysis of Iris Tool software (DIAT) v1\3 [30]. The software labels the pixels in the photograph as either blue or brown and calculates a score on a continuous scale from – 1–1. The value 1 equals to blue pixels only and – 1 equals to brown pixels only. Each photograph was manually corrected for eye-lid boundaries to ensure reliable scores.

Hair colour was self-reported as the participant's natural hair colour at the age around 20 and categorised as either blond, brown, red or black, together with shade (light or dark). In some cases, photographs of the hair were used as a control of the self-reported hair colour. Based on the self-reported hair colour, a vast majority of individuals in the Norwegian study population had blond (53.0%, n = 287), or brown (40.6%, n = 219) hair. Only 3.5% (n = 19) of the individuals had red hair and 2.8% (n = 15) had black hair (see Fig. 5 in Section 3.4).

### 2.3. Library preparation and sequencing

DNA was purified using either the QIA Symphony DNA Midi Kit (Qiagen) or the PrepFiler™ Express DNA Extraction kit (Thermo Fisher Scientific). The human male reference DNA 2800 M (Verogen) and 007 (Thermo Fisher Scientific) were used for performance testing of the ForenSeq™ DNA Signature Prep kit (Verogen). DNA was quantified

using the Quantifiler™ Trio DNA Quantification kit on the 7500 Real-time PCR system (Thermo Fisher Scientific).

One ng DNA was used to construct sequencing libraries with two different lots of the ForenSeq™ DNA Signature Prep kit (Verogen), following the manufacturer's instructions (VD 2018005 Rev. A). For the technical sensitivity study, serial dilutions of human male reference DNA 2800 M and 007 were prepared for DNA inputs of 500, 250, 125, 62.5 and 31.3 pg. 2800 M was also analysed at 1000 pg and 007 at 15.6 pg. All dilutions were analysed in triplicates.

Briefly, libraries were amplified with Primer mix B. Amplification, tagging and enrichment was carried out on a Veriti Thermal Cycler (Thermo Fisher Scientific). The libraries were purified and normalised using magnetic beads. Batches of 32 libraries, including a positive and negative control, were pooled and denatured. Twelve µl of this pool were loaded on the MiSeq® FGx Reagent Cartridge. Sequencing was otherwise performed on the MiSeq® FGx instrument according to the manufacturer's instructions.

## 2.4. Data analysis

### 2.4.1. Analysis of sequence data

Run metrics and sequence data were processed using the ForenSeq™ Universal Analysis Software v1.2 (UAS, Verogen). For interpretation, the default analytical threshold of 1.5% (minimum 10 reads) and interpretation threshold of 4.5% (minimum 30 reads) were applied. If both alleles had between 10 and 30 reads, the genotype was considered heterozygous. Genotypes for which only one allele was detected, and the read count was between 10 and 30 reads were considered inconclusive due to potential allele dropout. Furthermore, a noise limit of 10% was applied for calling homozygous genotypes.

Six representative sequencing runs with cluster densities of 1200 – 1600 K/mm<sup>2</sup> were chosen to study the typing performance of the DNA ForenSeq™ DNA Signature Prep kit. Of 180 samples (30 samples per run), 177 samples fulfilled the recommended total read count requirement of 85,000 and were included for further analyses. Libraries were prepared with two different reagent lots, but no lot-specific effect was observed (data not shown). For performance evaluation of the multiplex, profile completeness, depth of coverage (DoC) and allele balance were calculated. DoC per locus was defined by all reads per locus. Allele balances, also referred to as allele coverage ratios, were calculated for all heterozygous genotypes by dividing the number of reads of the allele with the lowest read depth, with number of reads of the allele with the highest read depth.

### 2.4.2. HirisPlex model-based prediction of eye and hair colour

Predictions of eye and hair colours were obtained by the HirisPlex multinomial logistic regression model using the current version of the webtool (<https://hirisplex.erasmusmc.nl/>). The model is based on 9466 individuals for eye colour [4,12,17] and 1878 individuals for hair colour [8,12]. Notably, the UAS (Verogen) also generates eye and hair colour predictions based on the first HirisPlex model [13]. However, this software was not used because the HirisPlex webtool generates additional predictions of hair colour shades (light/dark). Prediction results were reported as predictive probability values (p-values) for each category. For eye colour, the colour with the highest p-value (pmax) was considered as the predicted eye colour. Following the recommendation by Walsh et al., probability thresholds of 0.5 and 0.7 were also evaluated [17,32]. For hair colour prediction, both the highest probability category approach (HPCA) and the recommended prediction guide approach (PGA), which also takes shade into account, were evaluated [13].

The accuracy of prediction performance by the HirisPlex webtool was evaluated by calculating AUC (area under the receiver operating characteristic curve), sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) for each colour category. ROC curves and their corresponding AUC values were computed in R software v3.6.1 [33] using the R packages ROCR [34] and klaR [35].

All other statistical calculations and plot generations were performed using the Real Statistics Resource Pack software (release 6.3), copyright (2013 – 2020) Charles Zaiontz, [www.real-statistics.com](http://www.real-statistics.com). Mann-Whitney tests were employed for between group comparisons, and multiple Bonferroni correction was applied. The Spearman's correlation test was employed to test if PIE-scores and eye colour categories obtained by subjective categorisation were correlated.

## 3. Results

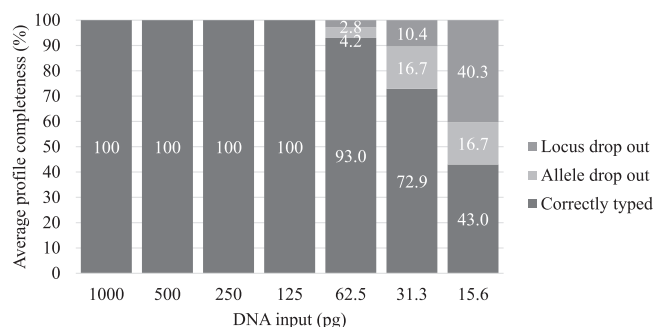
### 3.1. Genotyping performance of the HirisPlex SNPs using the ForenSeq™ DNA Signature prep kit

In total, 177 samples with 1 ng DNA input from six representative runs were evaluated for genotyping performance of the 24 HirisPlex SNPs. These samples had a median read count of 371,573, ranging from 100,008 to 715,552 reads. Complete HirisPlex profiles were obtained for all samples. The median DoC was 939 reads across the 24 SNPs, ranging from 61 reads in rs12896399 to 9780 reads in rs201326893\_Y152OCH (Table S1, Fig. S1). In addition to rs12896399, five other loci, rs2378249, rs12821256, rs12203592, rs28777 and rs1393350, had minimum read numbers below 100 (Table S1). All loci except rs201326893\_Y152OCH were heterozygous in one or more samples (see allele frequencies of all SNPs in Table S2). The average allele balance for these loci was  $0.82 \pm 0.12$  (mean  $\pm$  SD). SNP rs1800407 had the lowest average allele balance with  $0.73 \pm 0.17$  (mean  $\pm$  SD) (Fig. S1). There were very few observations of allele balances below 0.4, and these occurred in rs28777, rs683, rs2402130 and rs1800407 (Fig. S1).

The concordance of the six eye colour SNPs (IrisPlex) was assessed by comparing ForenSeq genotyping data with data obtained using a PCR-SBE assay [36]. SNP profiles for these markers were 100% concordant between methods.

#### 3.1.1. Technical sensitivity

Dilution series of control DNA were sequenced to assess the kit's ability to produce reliable HirisPlex profiles at low DNA input amounts. DoC decreased nearly linearly with decreasing amount of input DNA, and allele balances became more variable with decreasing amounts of input DNA (ranging from 0.71 to 0.99 and 0.24–1 with 1000 pg and 31.3 pg DNA input, respectively. Fig. S2). However, complete Hirisplex profiles were obtained in all samples with 125 pg DNA input or higher (Fig. 1). With 62.5 pg DNA input, allele dropout was observed in the most important SNP for eye colour prediction, rs12913832. Furthermore, locus dropout was observed in rs12821256, rs16891982, rs12203592 and rs12821256. One of the SNPs with low DoC, rs12203592 (Fig. S1), had the highest number of locus dropout (19% of the samples with 62.5 pg DNA). Applying the default analytical and interpretation thresholds, false homozygotes (drop out of one allele) and thus incorrect genotypes were observed in 56% of the samples with DNA



**Fig. 1.** Sensitivity study of the HirisPlex SNPs genotyped with the ForenSeq™ DNA Signature Prep Kit, showing profile average completeness in relation to DNA input.

**Table 1**

: Parameters describing the prediction accuracy of the HirisPlex webtool using the highest probability category approach for (A) eye colour, categorised by observation from photographs, and (B) hair colour, self-reported by the participants.

A) Eye colour		Blue	Intermediate	Brown	
Threshold					
pmax	Sensitivity	0.96	0.00	1.00	
	Specificity	0.56	1.00	0.82	
	PPV	0.70	0.00	0.47	
	NPV	0.93	0.66	1.00	
	AUC	0.85	0.69	0.94	
0.7 *	Sensitivity	0.98	0.00	1.00	
	Specificity	0.44	1.00	0.90	
	PPV	0.72	0.00	0.55	
	NPV	0.94	0.70	1.00	
B) Hair colour		Blond	Brown	Red	Black
Threshold					
pmax	Sensitivity	0.76	0.50	0.68	0.73
	Specificity	0.60	0.79	0.98	0.99
	PPV	0.68	0.62	0.50	0.61
	NPV	0.69	0.70	0.99	0.99
	AUC	0.72	0.70	0.97	0.93

\* Inconclusive samples (samples below the threshold) were excluded from accuracy calculations

input lower than 62.5 pg. This was observed in rs12913832, rs1042602, rs12821256, rs1800407 and rs12896399. Although the number of reads was above the interpretation threshold of 30 reads (ranging from 32 to 92 reads), the other allele had dropped out. The SNP rs1042602 had the overall highest number of allele dropout (in four of 15 samples with DNA input < 125 pg).

### 3.2. Eye colour prediction performance of IrisPlex

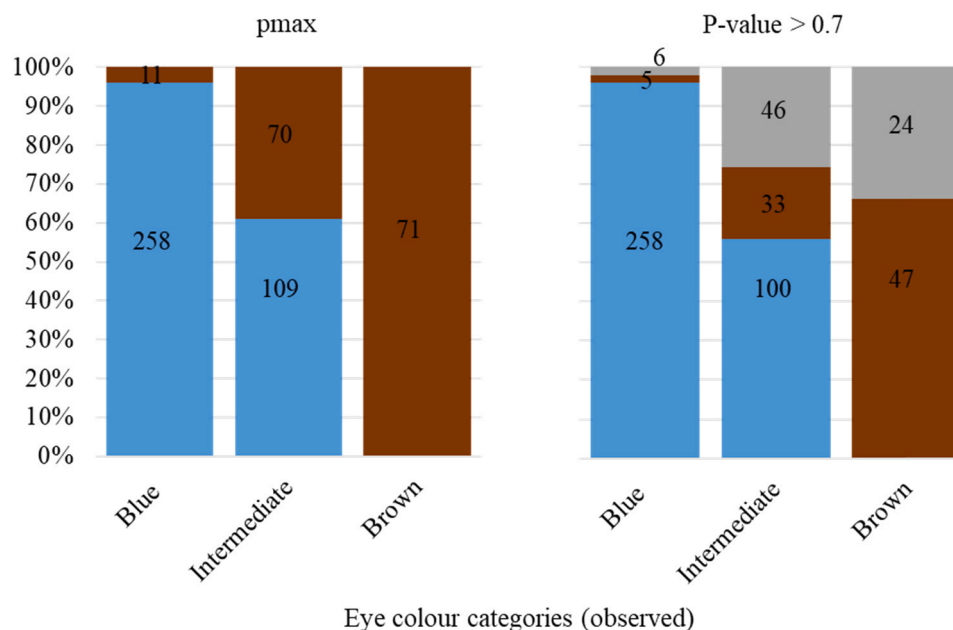
Using the IrisPlex webtool without threshold (pmax), 71% of the individuals (n = 367) were predicted as blue-eyed, 29% as brown-eyed (n = 152) and none as having eyes with intermediate colour. Thus, all individuals with intermediate eyes were either classified as blue (61%) or brown (39%) (Fig. 2).

The AUC reached acceptable values for blue (0.85) and brown (0.94)

eye colours but was rather low for intermediate-coloured eyes (0.69) (Table 1A). With pmax, 96% of the blue-eyed individuals and 100% of the brown-eyed individuals were classified correctly (Table 1A). A slight increase of correctly predicted blue eyes was observed when applying a probability threshold of 0.7, e.g. the sensitivity increased from 0.96 to 0.98 (Fig. 2 and Table 1A). In contrast to high sensitivities, the specificities for blue and brown eye colours were comparably low, with pmax of 0.56 and 0.82, respectively (Table 1A). This was mainly due to incorrect predictions of intermediate eye colour that resulted in high numbers of falsely positive predicted blue and brown eyes. With the probability threshold of 0.7, the proportion of false positive prediction of blue eyes increased, reducing the specificity to 0.44, whereas the proportion of false positive predictions of brown eyes decreased, increasing the specificity to 0.90 (Table 1A). With pmax, the number of false positive prediction of brown eyes was slightly higher than the true positives, resulting in a positive predictive value (PPV) of 0.47, whereas the PPV for blue eye colour was 0.70 (Table 1A). The high proportion of correctly classified blue and brown eye colours resulted in high negative predictive values (NPVs) of 0.93 and 1, respectively. The PPVs for blue and brown eye colours, as well as the NPV for blue eye colour increased slightly when applying the probability threshold of 0.7.

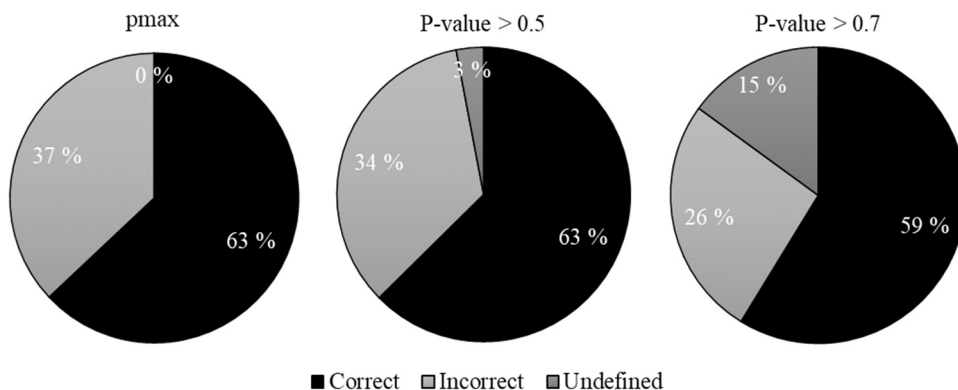
The overall correct predictions with pmax were 63% (Fig. 3). When applying thresholds, the incorrect predictions across the eye colour categories decreased from 37% to 34% with p-values > 0.5% and 26% with p-values > 0.7 (Fig. 3). Additionally, the proportion of inconclusive predictions were 3% and 15% with thresholds of 0.5 and 0.7, respectively. Most of the inconclusive predictions with a threshold of 0.7 (52/76) were incorrect predictions with pmax (Fig. 2). Thus, application of thresholds affected the overall error rate. However, with pmax, 94% of individuals with incorrect predictions had intermediate eyes, of which 74% were also incorrect with a threshold of 0.7. Additionally, 34% of the individuals correctly predicted to have brown eye colour with pmax were inconclusive with a threshold of 0.7 (n = 24) (Fig. 2). It is noteworthy to mention that assignment of intermediate eye colour when p-values for blue and brown are similar (as proposed in [13,17]) did not improve the prediction accuracy in the Norwegian study population considerably (data not shown).

Because eye colour predictions are mainly dependent on HERC2



**Fig. 2.** IrisPlex eye colour predictions obtained from 519 individuals from Norway with no threshold (pmax) and a probability threshold of 0.7. Colours within bars represent percentage of blue (blue), brown (brown) and inconclusive (grey) eye colour predicted within the respected observed eye colour category. Numbers in the bars represent numbers of individuals predicted within the category.





**Fig. 3.** The overall frequency of correct, incorrect and inconclusive (undefined) eye colour predictions using the IrisPlex webtool, with and without thresholds (n = 519).

rs12913832, eye colours of the Norwegian study group were also analysed based on this SNP (Table S3). Only 70.9% of rs12913832 GG individuals had the expected blue eye colour, whereas 79% and 39% of rs12913832 AA and AG individuals, respectively, had the expected brown eye colour. Intermediate eye colour was observed with all genotypes, but the majority was observed in rs12913832 GG individuals.

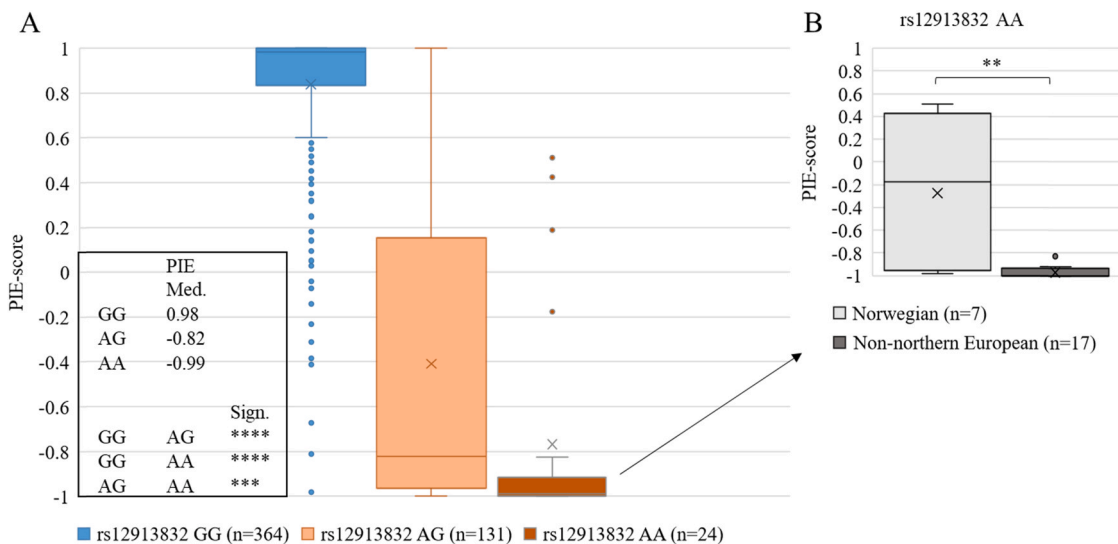
**3.3. Eye colour phenotype-genotype associations using quantitative PIE-score**

For the evaluation of phenotype-genotype associations, the eye colour phenotype was assessed using the quantitative PIE-score [30], an objective method that was found to be highly correlated with the qualitative categorisation by observation from photographs (Spearman’s correlation coefficient  $-0.88$ ;  $p < 0.001$ ) (Fig. S3). Among the six IrisPlex SNPs, rs12913832 showed the strongest association with PIE-score. The PIE-scores were significantly different for all pairwise comparisons between the three genotypes GG, AG and AA (Fig. 4A,  $p < 0.001$ ). The remaining five IrisPlex SNPs were also significantly associated with PIE-score, but not as strongly as rs12913832 (Fig. S4).

The PIE-score in individuals with the genotype rs12913832 AG varied considerably, ranging from 1 (blue eyes) to  $-1$  (brown eyes) (PIE median  $-0.82$ ) (Fig. 4A). Although there was generally little spread in the PIE-score among individuals with genotype GG (PIE median 0.98)

and genotype AA (PIE median  $-0.99$ ), we observed some individuals with genotype GG having negative PIE values (perceived as intermediate to brown) and individuals with genotype AA having positive PIE-values (perceived as intermediate). Notably, 94% of individuals carrying the AG genotype were reported to have Norwegian ancestry (n = 123). Norwegian ancestry was also reported in 91% individuals with genotypes GG and low PIE-score (PIE-score  $< 0$ , n = 11), as well as the four outliers with genotype AA and high PIE-score (Fig. 4A).

Although all individuals with genotype rs12913832 AA were predicted as brown-eyed by the IrisPlex model with high p-values ( $> 0.96$ ), we observe that individuals of Norwegian ancestry had statistically significantly higher PIE-scores (perceived as intermediate towards blue eyes) than non-northern Europeans ( $p < 0.01$ ; Fig. 4B). rs12913832 AA individuals in our study population were either of Norwegian or non-northern European ancestry. Additionally, we observed a statistically significant effect of rs16891982 on the PIE-score of rs12913832 AA individuals ( $p < 10^{-6}$ ; Fig. S5). All the individuals with Norwegian ancestry and genotype rs12913832 AA also carried the genotype rs16891982 GG, whereas all except two of the non-northern Europeans carried the ancestral C allele (13 individuals with rs16891982 CC and two individuals with rs16891982 CG). A small but statistically significant effect on the PIE-score was also observed when combining rs12913832 AA with rs1393350 or rs1800407, regardless of ancestry ( $p < 0.05$ ; Fig. S5). It is noteworthy to mention that only few



**Fig. 4.** : A) Boxplot showing PIE-scores of the rs12913832 genotypes (n = 519) and B) subgroups of rs12913832 AA individuals based on ancestry (n = 24), with pairwise comparisons between the genotypes (Mann-Whitney tests). Statistical significance (Sign.) after multiple Bonferroni correction:  $p < 0.02$ ,  $*p < 10^{-2}$ ,  $**p < 10^{-3}$ ,  $***p < 10^{-6}$ . Med.:Median.

rs12913832 AA individuals were observed with rs1393350 AG (n = 4) and rs1800407 AG (n = 2) in our data set.

Among individuals with the genotype rs12913832 GG, we observed a small but statistically significant effect on the PIE-score with the genotypes of rs12203592, rs1393350, rs12896399 and rs168991982 ( $p < 0.02$ ; Fig. S6). Notably, one individual of Asian ancestry with brown eyes had the rare combination of rs12913832 GG and rs168991982 CC (PIE-score:  $-0.98$ ). The phenotype was correctly predicted as brown by IrisPlex ( $p$ -value: 0.55). However, the prediction was inconclusive when applying the probability threshold of 0.7.

Among individuals with the genotype rs12913832 AG, a statistically significant effect on the PIE-score was only observed for rs168991982 ( $p < 0.01$ ; Fig. S7).

### 3.4. Hair colour prediction performance of the HirisPlex webtool

Using the HirisPlex webtool with the HPCA, the hair colour of 59% of the individuals (n = 321) were predicted as blond-haired, 32% as brown-haired (n = 175), 5% as red-haired (n = 26) and 3% as black-haired (n = 18) (Fig. 5).

We observed high AUC values for both red and black hair (0.97 and 0.93, respectively) (Table 1B). The AUC values for blond and brown hair were lower (0.72 and 0.70, respectively). The high AUC values for red and black hair reflected the high proportion of true negatives relative to false positives, resulting in high specificity values of 0.98 for red and 0.99 for black hair. For blond and brown hair, false positives were more frequent. Thus, specificities were rather low, with 0.60 for blond and 0.79 for brown hair. In terms of sensitivity, individuals with blond hair had the highest proportion of correct predictions, followed by black and red hair (0.76, 0.73 and 0.68, respectively), whereas only 50% of individuals with brown hair were correctly predicted. PPVs were relatively low for all hair colours (0.50–0.68), whereas the NPVs were higher (0.99 for red and black hair, 0.70 for brown and 0.69 for blond hair).

With the HPCA, the hair colour was predicted correctly in 65% of the individuals (Table 2). The most common prediction error was found in individuals with brown hair who were predicted to have blond hair (44%), or vice versa (21%). Furthermore, neither blond nor red-haired individuals were predicted to have black hair, and none of the black-

haired individuals were predicted to have red hair (Fig. 5, Table 2). Notably, half of the red predictions were false positives (n = 13), and most of them were for individuals with either blond or brown hair (Fig. 5). Two of the false positive predictions were also heterozygous in rs312262906\_N29insA. Individuals with this genotype are expected to have red hair and were predicted so with a  $p$ -value of 0.93 and 0.61, respectively. When inspecting the photographs that were taken during sample collection, hints of red could be observed in almost half of the individuals with false positive red hair prediction, including one of the individuals with the N29insA variant.

The overall correct predictions increased to 75% when utilising the PGA (Table 2). The PGA combines light/dark hair colour shade probabilities with the categorical hair colour probabilities. Because the shades of hair colour from blond via brown to black are overlapping, adding a  $p$ -value for light/dark shade to the categorical predictions provides an additional level of information. Thus, it is possible to differentiate light blond from dark blond and light brown from dark brown/black. Predictions made using HPCA are still considered as correct (e.g., predicted Blond, Blond/D-blond and D-blond/Brown is considered as correct for an individual with self-reported light blond or dark blond hair colour, Table 2). In combination with the HPCA, PGA allows for additional predictions to be considered as correct as there are overlaps between initial categories (e.g., predicted brown and self-reported as dark blond, and predicted D-brown/black and self-reported as black, Table 2). Additionally, dark blond and light brown hair colour were set into one category. The most pronounced increase in correct predictions when applying this approach was for dark blond/light brown (61–76%) and black (73–93%) hair colour categories (Table 2). When applying the rule that  $p$ -values above 0.25 for red hair are predictors for red, two additional individuals with red hair were predicted correctly, thereby increasing the correct prediction for red hair from 68% to 79% (Table 2).

## 4. Discussion

In this study, we assessed both the typing performance of the 24 HirisPlex SNPs using the ForenSeq™ DNA Signature prep kit (Verogen) and the eye and hair colour prediction by the HirisPlex webtool in a Norwegian study population. We have shown the performance metrics for each of the HirisPlex SNPs (typing success, read depth and allele balance) together with parameters describing prediction performance (AUC, sensitivity, specificity, PPV and NPV). Additionally, we assessed genotype-phenotype associations of eye colours using the quantitative PIE-score. Thereby, we show that population specific testing on the intended target population disclose rare and possibly population specific variation in association patterns, which might be useful for future improvements of prediction models.

### 4.1. MPS genotyping performance

We evaluated only the performance of the HirisPlex SNPs within the kit and found the typing performance of the multiplex to be satisfactory as complete profiles were obtained for all samples with optimal DNA input. Although the profiles were complete, in agreement with other studies we observed a locus-to-locus variation in terms of DoC and allele balances [24,29,37]. Despite these variations, the sensitivity study demonstrated that the system was also able to type complete and correct HirisPlex profiles with DNA input as low as 125 pg. This is slightly better than the developmental validation data showing complete and correct profiles with DNA input  $\geq 250$  pg [23]. However, locus dropout has been reported for samples with optimal DNA input in other studies [24, 29,37,38]. Notably, Sharma et al. [29] demonstrate that sequencing runs with cluster densities within the recommended range by Illumina (1200–1400 K/mm<sup>2</sup>) perform better in terms of dropouts than runs with lower cluster densities. Our runs were within this recommended cluster density range or slightly above.

As expected, read counts and allele balances decreased with

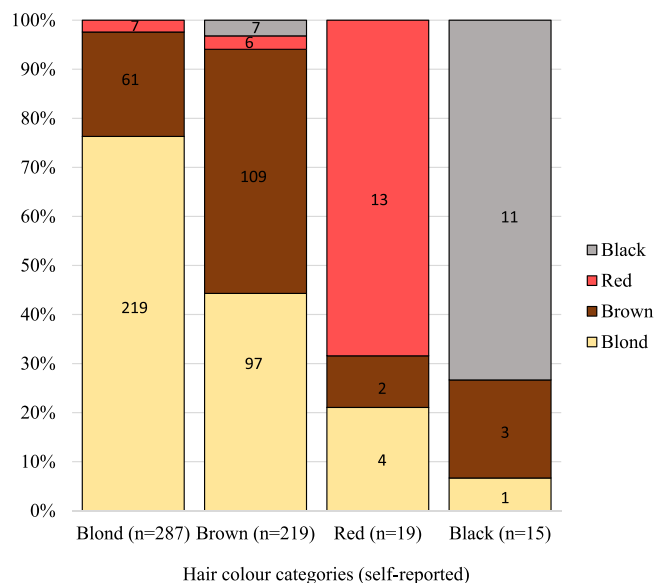


Fig. 5. HirisPlex hair colour predictions obtained from 540 individuals from Norway. Bars represent percentage of eye colour predicted within the respected self-reported hair colour category. Numbers in the bars represent numbers of individuals predicted within the category.

**Table 2**

HirisPlex hair colour predictions obtained from 540 individuals from a Norwegian study population using the highest probability category approach (HPCA) and the prediction guide approach (PGA). Predictions that were considered as correct using the HPCA are highlighted in dark grey, and additional predictions that were considered as correct using the PGA are highlighted in light grey. For the HPCA, the highest p-value for one of the four hair colours (blond, brown, black or red) was considered the predicted phenotype. The PGA combines the highest p-value approach with a p-value for shade (light or dark) in a step-wise model (for the prediction guide see [13]). According to the PGA, hint of red will be seen if a p-value for the red category is more than 0.25.

		Self-reported hair colour						Total (n=540)
		Dark blond/light brown (n=331)				Black (n=15)	Red (n=19)	
		Light blond (n=68)	Dark blond (n=219)	Light brown (n=112)	Dark brown (n=107)			
Predicted	Blond	Blond	20	52	15	7	1	0
		Blond/D-blond	37	99	41	21	0	4(-2)*
		D-blond/Brown	4	7	6	7	0	0
	Brown	Brown	0	1	0	1	0	0
		Brown/D-brown	1	22	26	23	0	1
		D-brown/Black	5	32	19	40	3	1
	Black	D-brown/Black	0	0	0	0	0	0
		Black	0	0	1	6	11	0
	Red	Red	1	6	4	2	0	13(+2)*
	HPCA		90 %	61 %	60 %	73 %	68 %	<b>65 %</b>
PGA		90 %	76 %	60 %	93 %	79 %	<b>75 %</b>	

\*Two individuals with highest p-value for blond hair but also a p-value > 0.25 for red hair.

\*Two individuals with highest p-value for blond hair but also a p-value > 0.25 for red hair.

decreasing amounts of DNA. This caused the poorer performing SNPs to drop out. One of the affected SNPs was rs12913832, which is the most important SNP for eye colour prediction. This SNP also performed poorly in other studies [29,37]. Without this marker, it is not possible to predict eye colour with the IrisPlex webtool [32]. In addition, three other highly ranked SNPs in the IrisPlex model, namely rs16891982, rs12203592 (for eye and hair colour predictions), and rs28777 (for hair colour prediction), were among the lower performing SNPs (also observed by Churchill et al. [37]). These four loci accounted for 36% of all dropouts with 62.5 and 31.3 pg DNA input in the sensitivity study. The consequence of failing to type rs12913832, rs16891982 and rs12203592 combined, is that eye and hair colour cannot be predicted [13]. Importantly, undiscovered allele dropout (false homozygotes) in highly ranked SNPs, especially rs12913832, may lead to false predictions in samples with little DNA input. We observed false homozygotes in these loci with as many as 92 reads, which is far above the default interpretation threshold of 30 reads. This supports the need for more stringent thresholds for genotype interpretation as previously discussed by others [39]. Therefore, to avoid misleading results, we

suggest increasing the interpretation threshold to 100 reads for low input samples or not analysing such samples at all. We emphasise that we did not observe false homozygotes with DNA input of 125 pg or more despite low read counts.

#### 4.2. Eye colour predictions in a Norwegian study population

When testing the IrisPlex model on the Norwegian study population, we obtained high AUCs for blue and brown eye colours, as well as high sensitivities. However, the AUC value for blue eye colour (0.85) was lower than the reported value of 0.94–0.97 in the IrisPlex-S DNA Phenotyping Webtool User Manual Version 2.0 (2018) and other studies [26–28]. This may be because our study resulted in a higher false positive rate for blue eye colour.

In concordance with other studies, no individuals in our study were predicted to have intermediate eyes [26–28]. This immensely affected the error rate, that was found to be 37%. Although the overall prediction error decreased when applying thresholds, most of the intermediate-eyed individuals were still incorrect with a threshold of

0.7. Consequently, the application of probability thresholds only slightly increased the prediction accuracy in the Norwegian population. The overall error rate of 26% with a threshold of 0.7 was still considerably higher than the reported 6% by Walsh et al., [17] who included seven populations across Europe in their dataset. The reason for this substantial difference may be that our Norwegian study population had a higher frequency of intermediate eyes than the other European populations. In addition, the discrepancies might also be partly due to the challenge in categorising eye colours. Both the IrisPlex model and our study were mainly based on subjective categorisation of eye colour by observation. It has been shown that general perception of eye colour varies between observers [31], suggesting that comparisons between studies may be difficult. This challenge of subjective perception was also demonstrated in the Norwegian population by the overlap in PIE-score especially between perceived blue and perceived intermediate eyes. Perception of eye colour might also be influenced by an observer's reference background. For example, a person residing in Norway, where the frequency of light blue eyes is high, might more easily categorise slightly darker eyes as intermediate or brown than a person who is more exposed to darker and brown coloured eyes. Furthermore, categorisation of a continuous trait such as eye colour, will inevitably contribute to some errors.

To be accustomed for forensic applications, a method must be highly reliable. Although the lack of intermediate eye colour predictions leads to a high error rate, none of the brown eyes were classified as blue and only 4% of the blue eyes were classified as brown. Therefore, a blue eye colour prediction indicates that the individual has most likely either blue or intermediate eyes, and a brown eye colour prediction indicates that the individual has most likely brown or intermediate eyes. This observation was also supported when the intermediate-coloured eyes were sub-categorised into intermediate-blue and intermediate-brown (Fig. S8). Most intermediate-blue eyes were predicted as blue (77%), and most intermediate-brown eyes were predicted as brown (83%), a trend also reported by Salvo et al. [28]. Thus, the obstacles of the IrisPlex model with a three-category system might at least in part be improved by using a two-category system of eye colour (blue and brown), as already suggested by others [15,31,40].

We also observed that more than half of the individuals with incorrect predictions had the *HERC2* genotype rs12913832 GG. This was in contrast to other studies of populations with high frequencies of intermediate eye colours who report highest numbers of incorrect predictions in individuals with rs12913832 AG [26,28,40]. The derived G allele in rs12913832 is known to reduce the production of eumelanin by negatively affecting *OCA2* expression. Hence, homozygous individuals of the derived G allele are expected to have blue eyes [18,19,41]. As only 71% of our individuals with rs12912823 GG genotype had blue eyes, some of the incorrect predictions might be due to the overlap between perceived blue and intermediate-blue, as discussed above. However, when analysing the eye colour based on the objective PIE-score, we also observed rs12913832 GG individuals with intermediate-brown and brown eye colour (low PIE-scores). Brown eye colour in rs12913832 GG individuals has also been observed in other Scandinavian populations (Swedes and Danes), suggesting that other variants might upregulate the expression of *OCA2* and overrule the G-allele down regulation of *OCA2* in some individuals [30]. A recent study showed that other SNPs than the IrisPlex SNPs, such as rs1126809 in *TYR*, rs62538956 and rs35866166 in *TYRP1* and rs1289469 in *SLC24A4*, were associated with brown eye colour in rs12913832 GG individuals [42]. Inclusion of these SNPs in the model for eye colour prediction, might improve the accuracy in the Norwegian population.

In addition to the rs12913832 GG individuals with non-blue eye colour, we observed some individuals with the AA and AG genotype having the unexpected intermediate and intermediate towards blue eye colours (based on PIE-score). Although the sample size was limited ( $n = 24$ ), rs12913832 AA individuals of Norwegian ancestry had less brown eye colours (based on PIE-score) than individuals of non-northern

European ancestry. The rs12913832 A allele is positively affecting transcription of *OCA2* and has been shown to be strongly associated with dark iris colour [18,43]. However, a nonsynonymous mutation (A-allele) in rs1800407 is suggested to be a penetrance modifier of rs12913832, associated with lighter eye colour [19,30,44]. Only two of the Norwegians carried the A-allele in rs1800407 with rs12913832 AA, suggesting that other SNPs might explain the lighter eye colour. All the individuals with Norwegian ancestry had the rs16891982 GG genotype, in contrast to non-northern Europeans who nearly all carried the ancestral C-allele. SNP rs16891982 in *SLC45A2* is extensively used as a predictor for European ancestry as the G-allele has a high frequency in Europeans (frequency 97%) and is nearly absent in Africans, South-East Asians and South Americans (frequency of  $\sim 1$ –2%) [45,46]. A recent study in mouse models showed that *SLC45A2* was involved in melanosome maturation, similar to *OCA2* [47]. They also demonstrated that the G variant in rs16891982 lead to an instable protein that negatively affected the melanin synthesis. This might explain the lighter eye colour that we observed in rs12913832 AA individuals. However, some of these individuals were not correctly predicted by IrisPlex. Other nonsynonymous mutations in *OCA2*, such as the T-alleles in s7465330 and rs121918166, have previously been shown to be associated with blue eye colour in heterozygous rs12913832 individuals [44]. These variants might also be associated with lighter eye colour in the Norwegian population. It is also important to note that other factors such as gender, age and iris patterns can affect eye colour [40,48,49]. Reported gender effects on eye colour seems to be population specific [40,49,50], and was not observed in the two Scandinavian populations analysed so far [40]. The number of rs12913832 AA individuals in our sample population was too small to fully explore this effect. However, the individuals that were observed to have lighter eye colour were both males and females of age 30–55 with perceived intermediate-brown and intermediate-blue eye colour.

#### 4.3. Hair colour predictions in a Norwegian study population

When testing the HirisPlex model on a Norwegian study population, we obtained similar AUC values for hair colour as reported in the HirisPlex-S DNA Phenotyping Webtool User Manual Version 2.0 (2018). However, the AUC value for blond hair (0.72) was considerably lower than reported by HirisPlex (0.81). This may be because the most common incorrect prediction across all hair colours were for individuals with brown hair who were predicted to have blond hair, which is also in line with data from a Swedish population [51]. Kukla-Bartoszek et al. suggested that incorrect predictions may be the result of age-dependent hair colour darkening [52]. They regularly observed children with blond hair developing brown hair with increasing age. Most of these individuals were predicted to have blond hair. Although we did not have information about our participants' hair colour as children, we suggest that age-dependent hair colour darkening increased the incorrect predictions of individuals with brown hair. We reduced some of these errors when applying the PGA (prediction guide approach) for interpretation, as recommended by Walsh et al. [8,13]. This approach increased the correct predictions in the dark blond/light brown category considerably, and the overall correct predictions across the hair colours from 65% to 75%. Our findings are in agreement with the reported 77% accuracy by HirisPlex and another study [8,53]. Therefore, the PGA should always be utilised for interpretation of a prediction, especially in populations in which blond and brown hair is common. However, blond-haired individuals (light and dark shades) were also predicted to have dark brown hair. This cannot be explained by age onset hair darkening. Even though some inaccuracies might arise due to people's subjective perception of hair colour, these findings suggest the need for further investigation of the genetics of hair colour. Moreover, in half of the individuals with false positive red hair prediction, hints of red were observed in photographs. This demonstrates the difficulties of self-categorisation. Additionally, "strawberry blond" and "auburn" might be difficult to distinguish from



blond and light brown. Thus, in a population with a high frequency of blond and brown-haired individuals, an individual that is predicted to have red hair might not have distinct red hair.

## 5. Conclusions

In this study, we demonstrated that genotyping the HirisPlex SNPs using the ForenSeq™ DNA Signature Prep kit was highly reliable with DNA input of 125 pg or higher. Using the HirisPlex webtool, eye and hair colour were predicted with high accuracies in terms of AUC for blue (0.85) and brown-eyed (0.94) individuals, as well as red (0.97) and black-haired (0.93) individuals. The Norwegian study population displayed high frequencies of blue (52%) and intermediate (34%) eyes, as well as blond (59%) and brown hair (32%). Because of the well-known limitation in predicting intermediate eye colours and the challenge with age-dependent hair colour darkening, a relatively high prediction error was observed (26% for eye colour with  $p > 0.7\%$  and 35% for hair colour using the HPCA). However, 75% of the hair colour predictions were correct when using the PGA as opposed to the HPCA approach. This study also disclosed rare phenotype-genotype associations, pointing to the importance of population specific testing before implementing FDP in casework.

## CRedit authorship contribution statement

**Nina Mjølunes Salvo:** Conceptualisation, Formal analysis, Visualisation, Writing – original draft. **Kirstin Janssen:** Conceptualisation, Supervision, Investigation, Writing – review & editing. **Maria Kristine Kirsebom:** Investigation. **Olivia Strunge Meyer:** Writing – review & editing. **Thomas Berg:** Conceptualisation, Supervision. **Gunn-Hege Olsen:** Conceptualisation, Supervision, Project administration, Writing – review & editing.

## 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.

## Acknowledgements

A special thanks to all participants. Thanks to Marita Olsen, Marthe Aune and Mari Zakariassen for sample collection and technical assistance. Thanks to Jeppe Dyrberg Andersen and Claus Børsting for comments on the manuscript and helpful discussions by personal communication. The project was funded by UiT – The Arctic University of Norway.

## Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at [doi:10.1016/j.fsigen.2021.102620](https://doi.org/10.1016/j.fsigen.2021.102620).

## References

- [1] P.M. Schneider, B. Prainsack, M. Kayser, The use of forensic DNA phenotyping in predicting appearance and biogeographic ancestry, *Dtsch. Arztebl.* Int 116 (2019) 873–880, <https://doi.org/10.3238/arztebl.2019.0873>.
- [2] M. Kayser, Forensic DNA phenotyping: predicting human appearance from crime scene material for investigative purposes, *Forensic Sci. Int. Genet.* 18 (2015) 33–48, <https://doi.org/10.1016/j.fsigen.2015.02.003>.
- [3] M. Kayser, P. de Knijff, Improving human forensics through advances in genetics, genomics and molecular biology, *Nat. Rev. Genet.* 12 (2011) 179–192, <https://doi.org/10.1038/nrg2952>.
- [4] F. Liu, K. van Duijn, J.R. Vingerling, A. Hofman, A.G. Uitterlinden, A.C.J. W. Janssens, M. Kayser, Eye color and the prediction of complex phenotypes from genotypes, *Curr. Biol.* 19 (2009) R192–R193, <https://doi.org/10.1016/j.CUB.2009.01.027>.
- [5] Y. Ruiz, C. Phillips, A. Gomez-Tato, J. Alvarez-Dios, M. Casares de Cal, R. Cruz, O. Maroñas, J. Söchtig, M. Fondevila, M.J. Rodriguez-Cid, Á. Carracedo, M. V. Lareu, Further development of forensic eye color predictive tests, *Forensic Sci. Int. Genet.* 7 (2013) 28–40, <https://doi.org/10.1016/J.FSIGEN.2012.05.009>.
- [6] J. Söchtig, C. Phillips, O. Maroñas, A. Gómez-Tato, R. Cruz, J. Alvarez-Dios, M.-Á. C. de Cal, Y. Ruiz, K. Reich, M. Fondevila, Á. Carracedo, M.V. Lareu, Exploration of SNP variants affecting hair colour prediction in Europeans, *Int. J. Leg. Med.* 129 (2015) 963–975, <https://doi.org/10.1007/s00414-015-1226-y>.
- [7] W. Branicki, F. Liu, K. van Duijn, J. Draus-Barini, E. Pośpiech, S. Walsh, T. Kupiec, A. Wojas-Pelc, M. Kayser, Model-based prediction of human hair color using DNA variants, *Hum. Genet.* 129 (2011) 443–454, <https://doi.org/10.1007/s00439-010-0939-8>.
- [8] S. Walsh, F. Liu, A. Wollstein, L. Kovatsi, A. Ralf, A. Kosiniak-Kamysz, W. Branicki, M. Kayser, The HirisPlex system for simultaneous prediction of hair and eye colour from DNA, *Forensic Sci. Int. Genet.* 7 (2013) 98–115, <https://doi.org/10.1016/J.FSIGEN.2012.07.005>.
- [9] G. Samuel, B. Prainsack, The regulatory landscape of forensic DNA phenotyping in Europe (2018). [http://www.visage-h2020.eu/Report\\_regulatory\\_landscape\\_FDP\\_in\\_Europe2.pdf](http://www.visage-h2020.eu/Report_regulatory_landscape_FDP_in_Europe2.pdf).
- [10] G. Samuel, B. Prainsack, Soc., Ethic-., Regul. Dimens. Forensic DNA phenotyping (2019). <http://www.visage-h2020.eu/#reports>.
- [11] L. Chaitanya, K. Breslin, S. Zuniga, L. Wirken, E. Pośpiech, M. Kukla-Bartoszek, T. Sijen, P. de Knijff, F. Liu, W. Branicki, M. Kayser, S. Walsh, The HirisPlex-S system for eye, hair and skin colour prediction from DNA: Introduction and forensic developmental validation, *Forensic Sci. Int. Genet.* 35 (2018) 123–135, <https://doi.org/10.1016/j.fsigen.2018.04.004>.
- [12] S. Walsh, L. Chaitanya, K. Breslin, C. Muralidharan, A. Bronikowska, E. Pospiech, J. Koller, L. Kovatsi, A. Wollstein, W. Branicki, F. Liu, M. Kayser, Global skin colour prediction from DNA, *Hum. Genet.* 136 (2017) 847–863, <https://doi.org/10.1007/s00439-017-1808-5>.
- [13] S. Walsh, L. Chaitanya, L. Clarisse, L. Wirken, J. Draus-Barini, L. Kovatsi, H. Maeda, T. Ishikawa, T. Sijen, P. de Knijff, W. Branicki, F. Liu, M. Kayser, Developmental validation of the HirisPlex system: DNA-based eye and hair colour prediction for forensic and anthropological usage, *Forensic Sci. Int. Genet.* 9 (2014) 150–161, <https://doi.org/10.1016/J.FSIGEN.2013.12.006>.
- [14] J.S. Allwood, S. Harbison, SNP model development for the prediction of eye colour in New Zealand, *Forensic Sci. Int. Genet.* 7 (2013) 444–452, <https://doi.org/10.1016/J.FSIGEN.2013.03.005>.
- [15] K.L. Hart, S.L. Kimura, V. Mushailov, Z.M. Budimlija, M. Prinz, E. Wurmbach, Improved eye- and skin-color prediction based on 8 SNPs, *Croat. Med. J.* 54 (2013) 248–256, <https://doi.org/10.3325/cmj.2013.54.248>.
- [16] J. Mengel-From, C. Børsting, J.J. Sanchez, H. Eiberg, N. Morling, Human eye colour and HERC2, OCA2 and MATP, *Forensic Sci. Int. Genet.* 4 (2010) 323–328, <https://doi.org/10.1016/J.FSIGEN.2009.12.004>.
- [17] S. Walsh, A. Wollstein, F. Liu, U. Chakravarthy, M. Rahu, J.H. Seland, G. Soubrane, L. Tomazozli, F. Topouzis, J.R. Vingerling, J. Vioque, A.E. Fletcher, K. N. Ballantyne, M. Kayser, DNA-based eye colour prediction across Europe with the IrisPlex system, *Forensic Sci. Int. Genet.* 6 (2012) 330–340, <https://doi.org/10.1016/J.FSIGEN.2011.07.009>.
- [18] M. Visser, M. Kayser, R.-J. Palstra, HERC2 rs12913832 modulates human pigmentation by attenuating chromatin-loop formation between a long-range enhancer and the OCA2 promoter, *Genome Res* 22 (2012) 446–455, <https://doi.org/10.1101/gr.128652.111>.
- [19] R.A. Sturm, D.L. Duffy, Z.Z. Zhao, F.P.N. Leite, M.S. Stark, N.K. Hayward, N. G. Martin, G.W. Montgomery, A Single SNP in an evolutionary conserved region within intron 86 of the herc2 gene determines human blue-brown eye color, *Am. J. Hum. Genet.* 82 (2008) 424, <https://doi.org/10.1016/J.AJHG.2007.11.005>.
- [20] S. Walsh, A. Lindenbergh, S.B. Zuniga, T. Sijen, P. de Knijff, M. Kayser, K. N. Ballantyne, Developmental validation of the IrisPlex system: determination of blue and brown iris colour for forensic intelligence, *Forensic Sci. Int. Genet.* 5 (2011) 464–471, <https://doi.org/10.1016/J.FSIGEN.2010.09.008>.
- [21] P. Valverde, E. Healy, I. Jackson, J.L. Rees, A.J. Thody, Variants of the melanocyte-stimulating hormone receptor gene are associated with red hair and fair skin in humans, *Nat. Genet.* 11 (1995) 328–330, <https://doi.org/10.1038/ng1195-328>.
- [22] N. Flanagan, E. Healy, A. Ray, S. Philips, C. Todd, I.J. Jackson, M.A. Birch-Machin, J.L. Rees, Pleiotropic effects of the melanocortin 1 receptor (MC1R) gene on human pigmentation, *Hum. Mol. Genet* 9 (2000) 2531–2537, <https://doi.org/10.1093/hmg/9.17.2531>.
- [23] A.C. Jäger, M.L. Alvarez, C.P. Davis, E. Guzmán, Y. Han, L. Way, P. Walichiewicz, D. Silva, N. Pham, G. Cavas, J. Bruand, F. Schlesinger, S.J.K. Pond, J. Varlaro, K. M. Stephens, C.L. Holt, Developmental validation of the MiSeq FGx forensic genomics system for targeted next generation sequencing in forensic DNA casework and database laboratories, *Forensic Sci. Int. Genet.* 28 (2017) 52–70, <https://doi.org/10.1016/J.FSIGEN.2017.01.011>.
- [24] F. Guo, J. Yu, L. Zhang, J. Li, Massively parallel sequencing of forensic STRs and SNPs using the Illumina® ForenSeq™ DNA Signature Prep Kit on the MiSeq FGx™ Forensic Genomics System, *Forensic Sci. Int. Genet.* 31 (2017) 135–148, <https://doi.org/10.1016/J.FSIGEN.2017.09.003>.
- [25] C. Børsting, N. Morling, Next generation sequencing and its applications in forensic genetics, *Forensic Sci. Int. Genet.* 18 (2015) 78–89, <https://doi.org/10.1016/J.FSIGEN.2015.02.002>.
- [26] G.M. Dembinski, C.J. Picard, Evaluation of the IrisPlex DNA-based eye color prediction assay in a United States population, *Forensic Sci. Int. Genet.* 9 (2014) 111–117, <https://doi.org/10.1016/J.FSIGEN.2013.12.003>.

- [27] V. Kastelic, E. Pošpiech, J. Draus-Barini, W. Branicki, K. Drobníč, Prediction of eye color in the Slovenian population using the IrisPlex SNPs. *Croat. Med. J.* 54 (2013) 381–386, <https://doi.org/10.3325/cmj.2013.54.381>.
- [28] C. Salvo, C. Faccinotto, L. Zucchelli, M. Porto, A. Marino, G. Occhi, G. de los Campos, G. Vazza, Performance of four models for eye color prediction in an Italian population sample, *Forensic Sci. Int. Genet.* 40 (2019) 192–200, <https://doi.org/10.1016/j.fsiggen.2019.03.008>.
- [29] V. Sharma, K. Jani, P. Khosla, E. Butler, D. Siegel, E. Wurmbach, Evaluation of ForenSeq™ signature prep Kit B on predicting eye and hair coloration as well as biogeographical ancestry by using universal analysis software (UAS) and available web-tools, *Electrophoresis* 40 (2019) 1353–1364, <https://doi.org/10.1002/elps.201800344>.
- [30] J.D. Andersen, P. Johansen, S. Harder, S.R. Christoffersen, M.C. Delgado, S. T. Henriksen, M.M. Nielsen, E. Sørensen, H. Ullum, T. Hansen, A.L. Dahl, R. R. Paulsen, C. Børsting, N. Morling, Genetic analyses of the human eye colours using a novel objective method for eye colour classification, *Forensic Sci. Int. Genet.* 7 (2013) 508–515, <https://doi.org/10.1016/j.fsiggen.2013.05.003>.
- [31] O.S. Meyer, C. Børsting, J.D. Andersen, Perception of blue and brown eye colours for forensic DNA phenotyping, *Forensic Sci. Int. Genet. Suppl. Ser. 7* (2019) 476–477, <https://doi.org/10.1016/J.FSIGSS.2019.10.057>.
- [32] S. Walsh, F. Liu, K.N. Ballantyne, M. van Oven, O. Lao, M. Kayser, IrisPlex: A sensitive DNA tool for accurate prediction of blue and brown eye colour in the absence of ancestry information, *Forensic Sci. Int. Genet.* 5 (2011) 170–180, <https://doi.org/10.1016/J.FSIGEN.2010.02.004>.
- [33] R Core Team, R: A Lang. Environ. Stat. Comput. (2019). <https://www.r-project.org/>.
- [34] T. Sing, O. Sander, N. Beerenwinkel, T. Lengauer, ROCr: visualizing classifier performance in R, *Bioinformatics* 21 (2005) 3940–3941, <https://doi.org/10.1093/bioinformatics/bti623>.
- [35] C. Weihs, U. Ligges, K. Luebke, N. Raabe, klaR Analyzing German Business Cycles, in: D. Baier, R. Decker, L. Schmidt-Thieme (Eds.), *Data Anal. Decis. Support*, Springer-Verlag, 2005, pp. 335–343, [https://doi.org/10.1007/3-540-28397-8\\_36](https://doi.org/10.1007/3-540-28397-8_36).
- [36] O.S. Meyer, N.M. Salvo, A. Kjørbye, M. Kjersem, M.M. Andersen, E. Sørensen, H. Ullum, K. Janssen, N. Morling, C. Børsting, G.-H. Olsen, J.D. Andersen, Prediction of eye colour in Scandinavians using the eyeColour 11 (EC11) SNP set, in: *Genes*, 12, Basel, 2021, p. 821, <https://doi.org/10.3390/GENES12060821>.
- [37] J.D. Churchill, N.M.M. Novroski, J.L. King, L.H. Seah, B. Budowle, Population and performance analyses of four major populations with illumina's fgx forensic genomics system, *Forensic Sci. Int. Genet.* 30 (2017) 81–92, <https://doi.org/10.1016/j.fsiggen.2017.06.004>.
- [38] A.L. Silvia, N. Shugarts, J. Smith, A preliminary assessment of the ForenSeq™ FGx System: next generation sequencing of an STR and SNP multiplex, *Int. J. Leg. Med.* 131 (2017) 73–86, <https://doi.org/10.1007/s00414-016-1457-6>.
- [39] C. Hussing, R. Bytyci, C. Huber, N. Morling, C. Børsting, The Danish STR sequence database: duplicate typing of 363 Danes with the ForenSeq™ DNA Signature Prep Kit, *Int. J. Leg. Med.* 133 (2019) 325–334, <https://doi.org/10.1007/s00414-018-1854-0>.
- [40] C. Pietroni, J.D. Andersen, P. Johansen, M.M. Andersen, S. Harder, R. Paulsen, C. Børsting, N. Morling, The effect of gender on eye colour variation in European populations and an evaluation of the IrisPlex prediction model, *Forensic Sci. Int. Genet.* 11 (2014) 1–6, <https://doi.org/10.1016/J.FSIGEN.2014.02.002>.
- [41] H. Eiberg, J. Troelsen, M. Nielsen, A. Mikkelsen, J. Mengel-From, K.W. Kjaer, L. Hansen, Blue eye color in humans may be caused by a perfectly associated founder mutation in a regulatory element located within the HERC2 gene inhibiting OCA2 expression, *Hum. Genet.* 123 (2008) 177–187, <https://doi.org/10.1007/s00439-007-0460-x>.
- [42] O.S. Meyer, M.M.B. Lunn, S.L. Garcia, A.B. Kjørbye, N. Morling, C. Børsting, J. D. Andersen, Association between brown eye colour in rs12913832:GG individuals and SNPs in TYR, TYRP1, and SLC24A4, *PLoS One* 15 (2020), e0239131, <https://doi.org/10.1371/journal.pone.0239131>.
- [43] M.N. Shapurenko, S.I. Vakula, A.V. Kandratsiuk, I.G. Gudievskaya, M. V. Shinkevich, A.U. Luhauniou, S.R. Borovko, L.N. Marchenko, A.V. Kilchevsky, HERC2 (rs12913832) and OCA2 (rs1800407) genes polymorphisms in relation to iris color variation in Belarusian population, *Forensic Sci. Int. Genet. Suppl. Ser. 7* (2019) 331–332, <https://doi.org/10.1016/J.FSIGSS.2019.09.127>.
- [44] J.D. Andersen, C. Pietroni, P. Johansen, M.M. Andersen, V. Pereira, C. Børsting, N. Morling, Importance of nonsynonymous OCA2 variants in human eye color prediction, *Mol. Genet. Genom. Med.* 4 (2016) 420–430, <https://doi.org/10.1002/mgg3.213>.
- [45] C. Phillips, W. Parson, B. Lundsberg, C. Santos, A. Freire-Aradas, M. Torres, M. Eduardoff, C. Børsting, P. Johansen, M. Fondevila, N. Morling, P. Schneider, Á. Carracedo, M.V. Lareu, Building a forensic ancestry panel from the ground up: The EUROFORGEN global AIM-SNP set, *Forensic Sci. Int. Genet.* 11 (2014) 13–25, <https://doi.org/10.1016/J.FSIGEN.2014.02.012>.
- [46] K.K. Kidd, W.C. Speed, A.J. Pakstis, M.R. Furtado, R. Fang, A. Madbouly, M. Maiers, M. Middha, F.R. Friedlaender, J.R. Kidd, Progress toward an efficient panel of SNPs for ancestry inference, *Forensic Sci. Int. Genet.* 10 (2014) 23–32, <https://doi.org/10.1016/J.FSIGEN.2014.01.002>.
- [47] L. Le, I.E. Escobar, T. Ho, A.J. Lefkovich, E. Latteri, K.D. Haltaufderhyde, M. K. Dennis, L. Plowright, E.V. Sviderskaya, D.C. Bennett, E. Oancea, M.S. Marks, SLC45A2 protein stability and regulation of melanosome pH determine melanocyte pigmentation, *Mol. Biol. Cell.* 31 (2020) 2687–2702, <https://doi.org/10.1091/mbc.E20-03-0200>.
- [48] R.A. Sturm, M. Larsson, Genetics of human iris colour and patterns, *Pigment Cell Melanoma Res* 22 (2009) 544–562, <https://doi.org/10.1111/j.1755-148X.2009.00606.x>.
- [49] C. Martinez-Cadenas, M. Peña-Chilet, M. Ibarrola-Villava, G. Ribas, Gender is a major factor explaining discrepancies in eye colour prediction based on HERC2/OCA2 genotype and the IrisPlex model, *Forensic Sci. Int. Genet.* 7 (2013) 453–460, <https://doi.org/10.1016/J.FSIGEN.2013.03.007>.
- [50] C. Martinez-Cadenas, M. Peña-Chilet, M.J. Llorca-Cardeñosa, R. Cervera, M. Ibarrola-Villava, G. Ribas, Gender and eye colour prediction discrepancies: a reply to criticisms, *Forensic Sci. Int. Genet.* 9 (2014) e7–e9, <https://doi.org/10.1016/J.FSIGEN.2013.10.002>.
- [51] K. Junker, A. Staadig, M. Sidstedt, A. Tillmar, J. Hedman, Phenotype prediction accuracy – A Swedish perspective, *Forensic Sci. Int. Genet. Suppl. Ser. 7* (2019) 384–386, <https://doi.org/10.1016/J.FSIGSS.2019.10.022>.
- [52] M. Kukla-Bartoszek, E. Pošpiech, M. Spólnicka, J. Karłowska-Pik, D. Strapagiel, E. Żądzińska, I. Rosset, M. Sobalska-Kwapis, M. Stomka, S. Walsh, M. Kayser, A. Sitek, W. Branicki, Investigating the impact of age-dependent hair colour darkening during childhood on DNA-based hair colour prediction with the HIRisPlex system, *Forensic Sci. Int. Genet.* 36 (2018) 26–33, <https://doi.org/10.1016/j.fsiggen.2018.06.007>.
- [53] 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, in: *Genes*, 11, Basel, 2020, pp. 1–24, <https://doi.org/10.3390/genes11121398>.