



OPEN

A new approach to broaden the range of eye colour identifiable by IrisPlex in DNA phenotyping

Ersilia Papparazzo^{1,7}, Anzor Gozalishvili^{2,3,7}, Vincenzo Lagani^{4,5}, Silvana Geracitano¹, Alessia Bauleo⁶, Elena Falcone⁶, Giuseppe Passarino¹ & Alberto Montesanto¹✉

IrisPlex system represents the most popular model for eye colour prediction. Based on six polymorphisms this model provides very accurate predictions that strongly depend on the definition of eye colour phenotypes. The aim of the present study was to introduce a new approach to improve eye colour prediction using the well-validated IrisPlex system. A sample of 238 individuals from a Southern Italian population was collected and for each of them a high-resolution image of eye was obtained. By quantifying eye colour variation into CIELAB space several clustering algorithms were applied for eye colour classification. Predictions with the IrisPlex model were obtained using eye colour categories defined by both visual inspection and clustering algorithms. IrisPlex system predicted blue and brown eye colour with high accuracy while it was inefficient in the prediction of intermediate eye colour. Clustering-based eye colour resulted in a significantly increased accuracy of the model especially for brown eyes. Our results confirm the validity of the IrisPlex system for forensic purposes. Although the quantitative approach here proposed for eye colour definition slightly improves its prediction accuracy, further research is still required to improve the model particularly for the intermediate eye colour prediction.

Forensic DNA Phenotyping (FDP) is an emerging field of forensic genetics aimed at prediction of externally visible characteristics (EVC) of unknown sample donors directly from biological materials found at the crime scene. This approach is expected to provide clues helping investigators reduce/prioritize their list of suspects and make police investigations more rapid, efficient and less expensive¹⁻³. While forensic genetic research is searching for additional phenotypic characteristics for predicting human appearance, those related to the pigmentations (eye, skin and hair colour) are today among the ones best characterized and validated⁴. In this context, eye colour is the best investigated phenotype for forensic genetic applications. In fact, a lot of genetic variants have been successfully identified in relation with iris pigmentation⁵⁻⁹. Some of these variants constitute the so-called IrisPlex system that to date represents the most popular model for eye colour prediction¹⁰. This system is based on the analysis of six Single Nucleotide Polymorphisms (SNP) located in six different genes: rs12913832 (*HERC2*), rs1800407 (*OCA2*), rs12896399 (*SLC24A4*), rs16891982 (*SLC45A2*), rs1393350 (*TYR*) and rs12203592 (*IRF4*). The IrisPlex model is based on a multinomial logistic regression model by which each individual is classified as being brown, blue or intermediate^{10,11}. The parameters of such a model were initially estimated using phenotype and genotype data from 3804 Dutch individuals. In particular, genetic data are modelled in an additive fashion (number of minor alleles in the genotype) and the highest probability of all 3 categories was taken as the predicted iris colour of that individual. Using this model, very accurate prediction values were obtained for brown and blue eyes, while the prediction of intermediate colour is less precise. There have been several attempts to refine the IrisPlex system to improve its predictive value. These were based on both an increased number of analysed genetic variants and a different statistical modelling strategy¹²⁻¹⁴. However, despite these precautions, these alternative systems did not obtain the desired effects since recent data showed that the IrisPlex system still was the best performing model for eye colour prediction¹⁵. Eye colour is usually described qualitatively using subjective and visually defined phenotype categories. This discretization approach oversimplifies the quantitative nature of the

¹Department of Biology, Ecology and Earth Sciences, University of Calabria, 87036 Rende, Italy. ²Toptal, LLC, 2810 N. Church St. #36879, Wilmington, DE 19802-4447, USA. ³Ivane Javakhishvili Tbilisi State University, 0162 Tbilisi, Georgia. ⁴Institute of Chemical Biology, Ilia State University, 0162 Tbilisi, Georgia. ⁵Biological and Environmental Sciences and Engineering Division (BESE), King Abdullah University of Science and Technology KAUST, Thuwal 23952, Saudi Arabia. ⁶BIOGENET, Medical and Forensic Genetics Laboratory, 87100 Cosenza, ASP, Italy. ⁷These authors contributed equally: Ersilia Papparazzo and Anzor Gozalishvili. ✉email: alberto.montesanto@unical.it

SNP	Alleles	MAF	HWE
HERC2-rs12913832	A/G	31.3	0.030
OCA2-rs1800407	G/A	4.9	1.000
SLC45A2-rs16891982	G/C	15.4	1.000
TYR-rs1393350	G/A	17.7	1.000
SLC24A4-rs12896399	G/T	26.6	0.423
IRF4-rs12203592	C/T	6.0	0.604

Table 1. Minor allele frequency (MAF) for each SNP, along with Hardy–Weinberg Equilibrium (HWE) p-value.

Algorithm	Data preprocessing	Number of clusters	Silhouette value	Adjusted rand index
K-means	Original	4	0.407	0.332
BIRCH	Original	4	0.389	0.315
K-means	Normalized	4	0.372	0.381
K-medoids	Normalized	4	0.342	0.381
SC	Normalized	4	0.309	0.396

Table 2. selection of solutions from the clustering analysis. For each solution, the respective clustering algorithm, whether the data were normalized or used in the original CIELAB values, the number of clusters, as well as the silhouette and adjusted Rand index value are reported. Full list in Supplementary Table 1.

trait causing an inevitably loss of information². For this reason, several authors proposed quantitative measurements of iris colour^{16–19}. This strategy not only allowed in the past years the identification of new genetic variants, but also the determination of a genetic model able to explain about 50% of quantitative eye colour variation¹⁷. Anyhow, the introduction of these measurements requires a methodology able to capture eye/hair colour in its fully continuous spectrum as accurately as possible² since current models for eye colour prediction, such as the IrisPlex system, are not able to handle this kind of data.

The aim of this present study is to introduce a new quantitative approach for eye colour prediction using the well-validated IrisPlex system and high-resolution digital images and genotype data from 238 individuals from a Southern Italian population. To this purpose, several alternative iris colour categorizations were evaluated and inserted within the frame of the IrisPlex model for improving its classification accuracy.

Results

Table 1 reports the minor allele frequencies for each SNP in the analysed sample together with the p-values of test of departure from Hardy–Weinberg equilibrium (HWE). All polymorphisms complied with HWE except rs12913832 located within the *HERC2* gene.

Eye colour categorization. The visual inspection produced the following eye colour distribution in the analysed sample: 29 blue (3 blue-grey and 26 sky-blue), 55 intermediate (34 chestnut-green and 21 green), and 154 brown (52 light brown and 102 dark brown).

Eye colour quantification using clustering algorithms. In order to obtain an objective eye colour classification, several clustering algorithms were applied on the CIELAB parameters. Table 2 reports the clustering solutions with the highest Silhouette index and four different clusters (see Supplementary Table 1 for the full list of explored clustering solutions).

We select the best clustering model based on a Pareto-optimal criterion; solutions that were top-ranked in either silhouette or adjusted rand index were deemed the optimal ones (see Fig. 1). According to this criterion, k-means with both original and normalized data, and SC with normalized data were chosen for subsequent analyses.

Alluvial plots (Fig. 2, Supplementary Figs. 1 and 2) show the distribution of the three-category classification of the IrisPlex model (blue, intermediate and brown) across a more detailed initial visual classification (sky-blue, grey-blue, green, chestnut-green, light-brown and dark-brown) and the groups produced by the selected clustering algorithms.

We then labelled each cluster according to the prevalence of the colour flows into the cluster itself. For all clustering solutions, cluster 1 was labelled as blue, cluster 3 as intermediate and both cluster 0 and 2 as brown. In general, all clustering results allowed to distinguish between a light and a dark intermediate colour (cluster 3 and cluster 0, respectively).

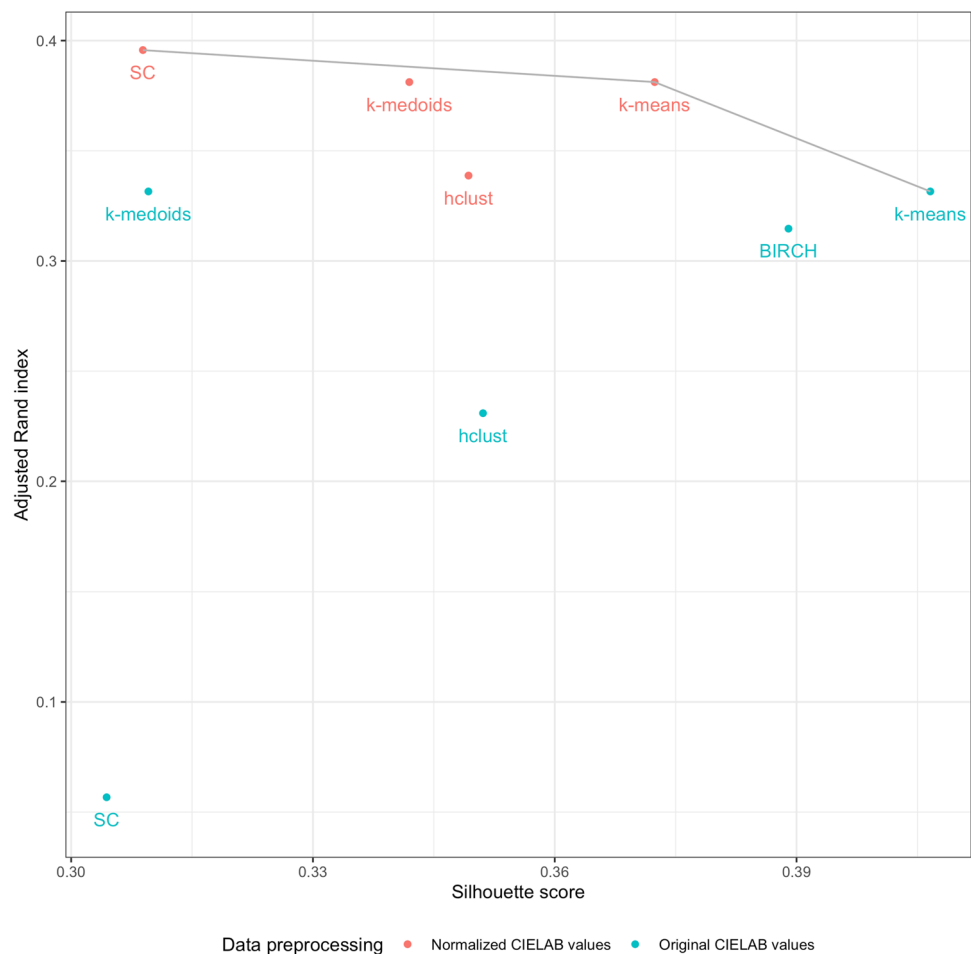


Figure 1. clustering performances. Each point represents a clustering solution, with the x-axis reporting the corresponding silhouette score, the y-axis the adjusted Rand index, and the colour indicating whether the CIELAB values were normalized. The Pareto front is represented as a grey line connecting the Pareto-optimal solutions (k-means solutions as well as SC with original CIELAB values).

Contrasting IrisPlex predictions against eye-colour labels obtained by visual inspection and clustering analysis.

Table 3 reports the overall accuracy obtained by the IrisPlex model on our cohort, according to different levels of thresholding and different eye colour definition. IrisPlex performances generally improve with higher threshold values. Most relevantly, it is clearly visible that the overall accuracy increases when the eye-colour labels defined by the k-means clustering algorithm are considered, with the original CIELAB values (not normalized) giving the best results.

Since the best-performing clustering solution was the k-means on the original (non-normalized) data, all subsequent analyses were performed based on eye-colour defined on the basis of such an algorithm.

Figure 1 shows the number of correct, incorrect, and undefined predictions at each threshold value and for (a) the eye-colour defined by visual inspection, (b) eye-colour defined through k-means clustering on the original (non-normalized) CIELAB values. The histograms indicate that applying a threshold improves the overall performance of the model because mostly incorrect predictions are turned into inconclusive ones. In other words, low confidence predictions are most likely incorrect, and excluding them from the evaluation increases the overall model performance.

In order to investigate this increase in accuracy, Fig. 4 dissects the model predictions according to eye colour and classification threshold. The eye-colour classification obtained by the clustering analysis provided performances in terms of accuracy higher than those obtained using eye-colour classification by visual inspection. In particular, the clustering analysis reclassified as brown a substantial number (29) of samples labelled as intermediate by the visual inspection, and this reclassification agrees with the IrisPlex which classifies these same samples as brown as well. Notably, the clustering analysis operates exclusively on the CIELAB values, while the IrisPlex solely analyses the genomic data, thus these two independent sources of information agree on this reclassification.

Regarding the effect of thresholding, it can be observed that increasing the threshold to 0.7 redefined as undefined the brown eyes that are incorrectly predicted as blue. Brown eyes became inconclusive by 3.2% (5 out of 154) for eye colour defined by visual inspection and 7.1% (13 out of 183) for the clustering-based approach, respectively. Blue eyes predicted as brown were reduced by 40% (2 out of 5) for eye colour defined by visual

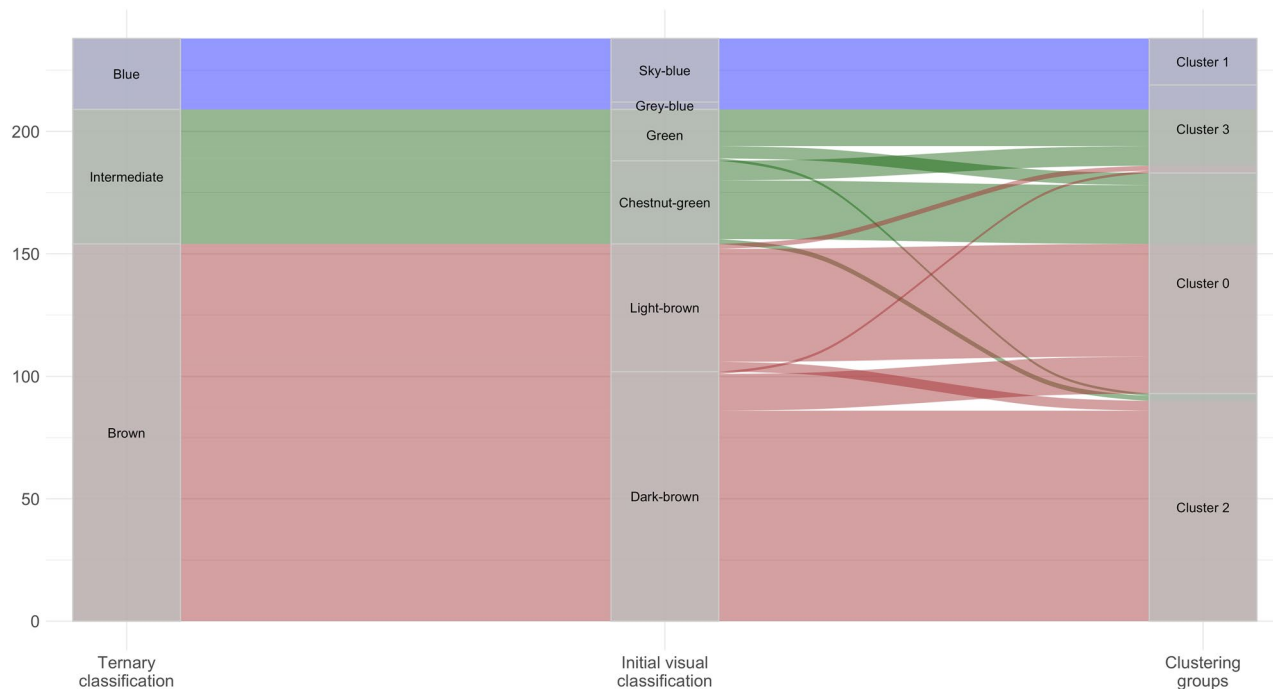


Figure 2. Alluvial plot showing sample distributions across different classifications. From left to right: ternary classification as used in the IrisPlex system, initial visual classification, and clustering groups as defined by k-means applied on the original CIELAB values.

Threshold	Visual classification	K-means (not normalized)	K-means (normalized)	SC (normalized)
0.7	0.795	0.865	0.851	0.763
0.5	0.765	0.848	0.835	0.730
0	0.744	0.832	0.811	0.706

Table 3. IrisPlex accuracy obtained for different levels of thresholding (rows) and different eye colour definition (columns).

Colour	Threshold	Visual inspection eye-colour					Clustering-based eye-colour (k-means on non-normalized data)				
		ACC	Sensitivity	Specificity	PPV	NPV	ACC	Sensitivity	Specificity	PPV	NPV
Blue	0	0.962	0.828	0.981	0.857	0.976	0.937	0.842	0.945	0.571	0.986
Blue	0.5	0.970	0.885	0.980	0.852	0.985	0.948	0.941	0.948	0.593	0.995
Blue	0.7	0.972	0.880	0.984	0.880	0.984	0.953	0.941	0.955	0.640	0.995
Brown	0	0.756	0.994	0.321	0.729	0.964	0.878	0.995	0.491	0.867	0.964
Brown	0.5	0.778	0.994	0.342	0.754	0.963	0.891	0.994	0.520	0.882	0.963
Brown	0.7	0.809	1.000	0.379	0.784	1.000	0.907	1.000	0.556	0.895	1.000
Intermediate	0	0.000	0.000	1.000	0.000	0.769	0.000	0.000	1.000	0.000	0.849
Intermediate	0.5	0.000	0.000	1.000	0.000	0.783	0.000	0.000	1.000	0.000	0.857
Intermediate	0.7	0.000	0.000	1.000	0.000	0.809	0.000	0.000	1.000	0.000	0.870

Table 4. Detailed performance metrics by eye colour and threshold. Results shown both for eye-colour defined through visual inspection and clustering algorithms.

inspection and 66.7% (2 out of 3) for the clustering approach. Intermediate eyes were never predicted as intermediate but incorrect predictions as brown decreased by 26.9% (14 out of 52) for eye-colour defined by visual inspection; for eye-colour defined by clustering analysis, applying a 0.7 threshold reduces the number of incorrect predictions by 22.2% (8 out of 36 samples becomes undefined).

In Table 4 the classification metrics for each colour category and threshold value are reported, both for the eye-colour defined by visual inspection and clustering analysis. It is clearly visible that all the performance metrics

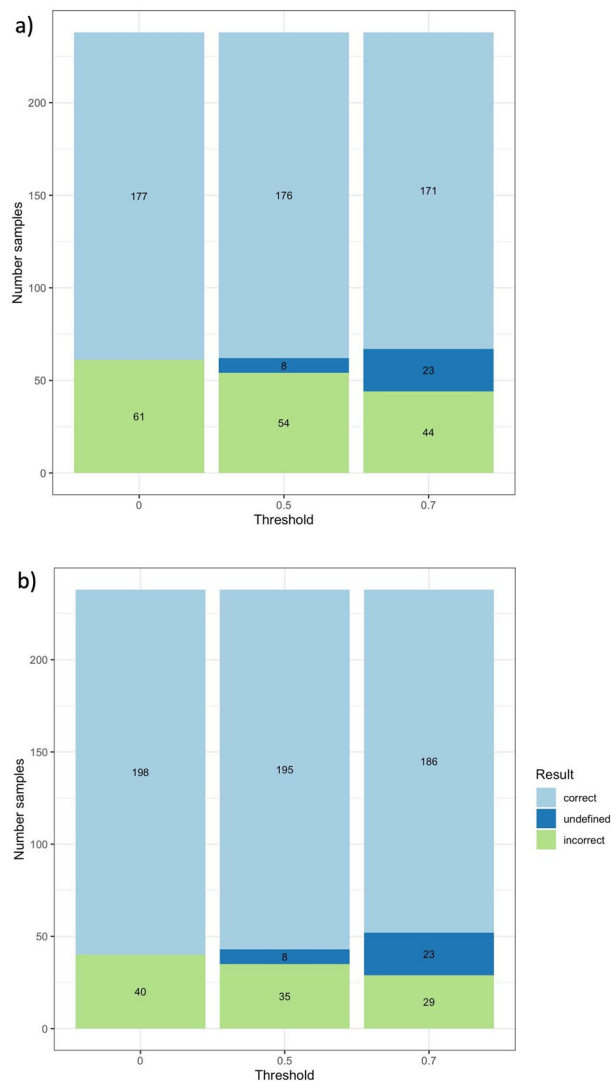


Figure 3. IrisPlex results. Each panel shows the number of correct, incorrect and undefined predictions at each threshold value. Particularly, panel (a) shows the results obtained with the eye-colour defined by visual inspection, and panel (b) with the eye-colour defined through k-means clustering on the original (non-normalized) CIELAB values.

were improved by applying a threshold, as shown also in Figs. 3 and 4. Using the eye classification provided by the clustering analysis clearly improves the specificity for the brown category, mainly due to reclassification as brown of several intermediate samples recognized as brown also by the IrisPlex model. We also observe a decrease in the specificity and PPV for the blue colour, due to the reclassification of 8 samples from blue (visual inspection) to intermediate and classified as blue by IrisPlex.

The ternary plots in Fig. 5 show the probabilities produced by the IrisPlex system. We highlight the IrisPlex model difficulties in separating the intermediate category from brown. Basically, no threshold can well separate brown and intermediate examples. Intermediate samples fell in both blue and brown sector almost equally when the eye colour was defined through clustering analysis (panel b), while they were mostly concentrated in the brown section in the case of the labels defined by visual inspection.

These plots also clearly underline the samples deemed as intermediate by visual inspection that become brown according to the clustering analysis (points switching from green to brown in the bottom left corner of the panel b) as well as the blue samples (visual inspection) turning into intermediate (clustering analysis) in the top corner of the two triangles.

Discussion

In the present study the efficacy of the IrisPlex model for eye colour prediction was analyzed in 238 individuals of Italian ancestry to evaluate their possible applicability as a tool of DNA intelligence in forensic investigations. Our results confirm the previous findings from several different populations showing once again that the IrisPlex system predicts blue and brown eye colour with high accuracy while it is inefficient in the prediction of

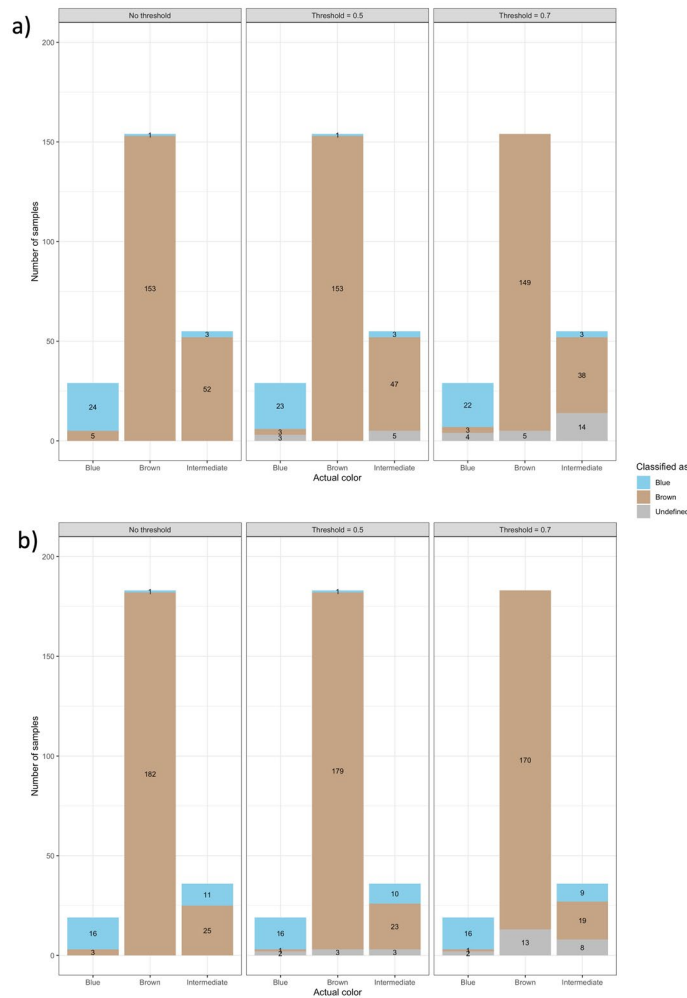


Figure 4. IrisPlex results dissected by eye-colour and threshold. Each panel shows blue, brown and intermediate eyes are classified. Panel (a) shows the results obtained with the eye-colour defined by visual inspection, and panel (b) with the eye-colour defined through k-means clustering on the original (non-normalized) CIELAB values.

intermediate eye colour^{20–25}. Indeed, the accuracy values for blue and brown eye colour categories in our sample were very high and equal to 0.972 and 0.809, respectively, while no one intermediate eye colour was correctly predicted as previously reported in another Italian sample¹⁵.

Here, we quantified continuous eye colour variation into CIELAB colour space using high-resolution digital full-eye photographs following the procedure reported in Edwards¹⁹. Clustering algorithms applied on the CIELAB parameters allowed us to obtain a standardized and objective measurement of eye colour, as well as, a better and more precise definition of the phenotype under study. Slightly improved results were obtained when this clustering-based approach was used for eye colour classification. In particular, using several clustering algorithms applied on quantitative measurements of iris colour, we obtained an improved classification performance especially for the clustering-based brown category.

The clustering-based approach here proposed, likewise other similar quantitative approaches for eye colour definition, may also be exploited as a standardized and objective measurement of eye colour useful also because it makes possible to directly compare results from different studies. In fact, one of the most important limitations affecting the development of a genetic model for eye colour prediction is the definition of the phenotype. Subjective interpretations of eye colour, by oversimplifying the quantitative nature of the trait and causing an inevitably loss of information, makes it difficult to compare and validate the results obtained in different populations and this also affects the classification performance of the adopted model.

There have been several attempts to refine the IrisPlex system to improve its predictive value mainly focused on the increase in the number of genetic variants^{12–14}. This approach did not obtain the desired effects since the IrisPlex system still represents the best performing model for eye colour prediction. Within this context, another very promising approach seems to be the inclusion of epigenetic markers. In fact, several authors observed that the hect domain and RCC1-like domain 2 (*HERC2*) rs12913832 variation, the marker of the IrisPlex system with the highest discrimination power, is located in an enhancer element that regulates the expression of *OCA2* gene⁷. In addition, it was also shown that *OCA2* expression was reduced in lightly pigmented melanocytes with the

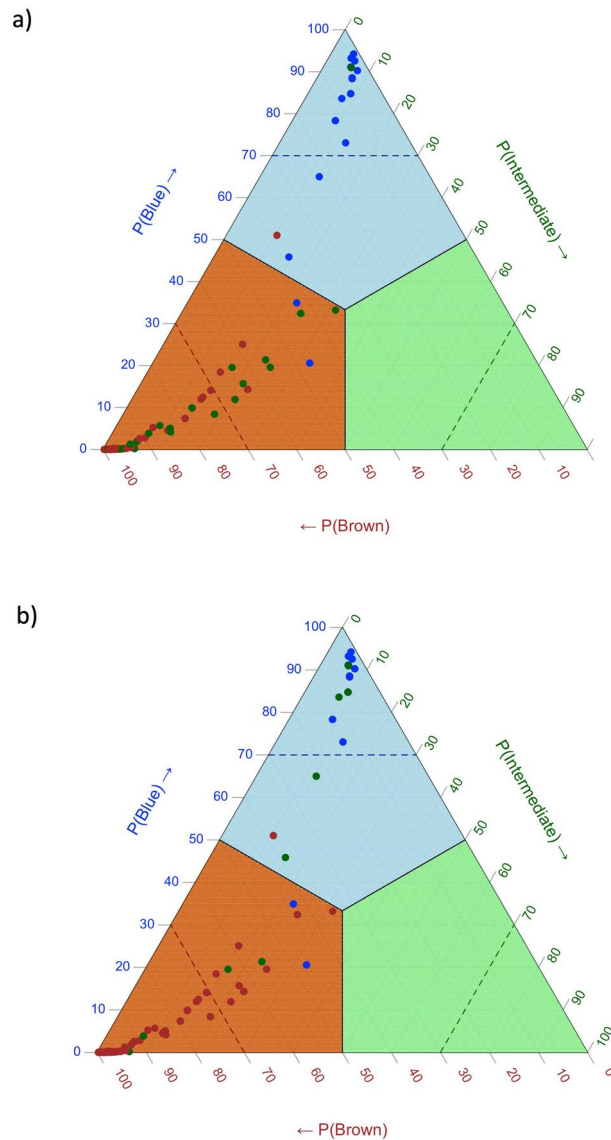


Figure 5. Ternary plots representing, for all samples (dots), the probabilities provided by the IrisPlex system. Panel (a) shows the results obtained with the eye-colour defined by visual inspection, and panel (b) with the eye-colour defined through k-means clustering on the original (non-normalized) CIELAB values.

rs12913832-G variant with respect to darkly pigmented melanocytes with the A allele^{7,26}. In agreement with this observation, the inclusion of epigenetic markers in the IrisPlex model might be useful to improve its prediction accuracy and in particular for the non-blue and non-brown eye colours.

The aim of this work was to test the predictive capabilities of the IrisPlex system, using eye colour definitions based both on visual inspection and on quantitative approach (clustering). Consequently, we based our attention to the clustering solutions in which three or more groups were identified, discarding clustering solutions identifying only two eye colours, since testing the IrisPlex predictions on these solutions would have been problematic. However, an interesting study carried out by Meyer et al. clearly showed that the perception of intermediate eye colour varies greatly among individuals, and this represents the main reason why using only two categories of eye colour (blue and brown) provides better results than a three-category system (blue, intermediate, and brown)²³. In line with these results, the Section of Forensic Genetics in Denmark recently began offering eye colour prediction to the police using two categories of eye colour (blue and brown) through the analysis of rs12913832 variability. All these lines of evidence, together with our results, suggest that the current definition of eye colour based on visual inspection should either be re-defined on the basis of more quantitative criteria or should be dropped all together in favour of a two-colour definition. Although the quantitative approach here proposed for eye colour definition improves the prediction accuracy of IrisPlex system, further research is still required to improve the model performance particularly for the non-blue and non-brown eye colour prediction.

Methods

Sample. The present study was carried out at the Department Biology, Ecology and Earth Sciences of the University of Calabria within a recruitment campaign focused on students and staff of the University between November 2018 and October 2019. 238 individuals (72 men and 166 women) were recruited. Trained staff members administered a brief and standardized questionnaire in order to obtain information regarding the socio-demographic data. During the interview, eye images using a professional camera were obtained and buccal swabs were collected as source of DNA. Written informed consent was obtained from all recruited individuals. The study was approved by the Ethics Committee of University of Calabria (Prot. NP-5942018) and met the criteria of the Helsinki declaration.

Digital photographs. Photographs were taken at a distance of approximately 10 cm of each individual's left iris under similar light conditions with a Nikon P300 with 100 mm f/1.8 NIKKOR Optical Zoom Lens, ISO 800. A coaxial biometric illuminator was used to deliver a constant and uniform source of light to each iris at 5,500 K (D55 illuminant).

Classification of eye colour by visual inspection of digital photographs. Iris colour was classified qualitatively by human visual identification as already described in other studies^{15,20,21,25}. Briefly, each eye image was graded independently by 2 different observers who classified eye colours into four categories: blue (including blue-grey and sky-blue), green (including green, and green with brown iris ring), chestnut-green (including peripheral green central brown, brown with some peripheral green) and brown (including light brown and dark brown). In order to keep the three-category classification of the IrisPlex model and to ensure consistency across studies, we mapped green and chestnut-green categories to intermediate category. Note that these two categories correspond to light intermediate and dark intermediate classes described in other studies^{15,22,25}. A third observer was consulted to resolve inconsistencies through majority-voting and to assess the final eye colour of each volunteer. Overall, 91% (217/238) of the classifications showed complete agreement between the 2 observers. Of the 21 remaining discrepancies, 18 were between light brown and chestnut-green, finally classified 17 as light brown, one as chestnut-green; the remaining discrepancy was between sky-blue and green, finally classified as green.

Quantitative eye colour. Image processing was based on the procedure reported in Edwards and colleagues using the dedicated webtool¹⁹. In brief, after the scleral, pupillary and collarette boundaries are defined, the application automatically extracts a measurement of average eye colour starting from a 60° angle wedge taken from the left side of the iris. The web application also isolates the portion of the wedge that represents the ciliary zone and the portion of the wedge that represents the pupillary zone. At the end of this procedure, for each iris image, the average RGB value of the entire wedge, the ciliary and the pupillary zones are obtained. The obtained RGB values are then converted into in CIE 1976 L*a*b* (CIELAB) colour space. In this colour space, the L* coordinate represents the lightness dimension and ranges from 0 to 100, with 0 being black and 100 being white. The red/green colours are represented along the a* coordinate, with green at negative a* values and red at positive a* values. The yellow/blue colours are represented along the b* coordinate, with blue at negative b* values and yellow at positive b* values.

Although several automated methods have been developed to facilitate the isolation of the iris from photographs of the eye^{17,18,25}, the method here adopted as reported in Edwards et al¹⁹, appears to be superior as it allows to manually define the boundaries of the iris and to separate the eye into different regions. Since the left quadrant of the iris was least likely to be obstructed by eyelashes and eyelids, it would bias the colour of the iris towards the pupillary region, we selected a wedge to represent iris colour instead of the entire iris.

Classification of eye colour using an unsupervised machine learning approach. In order to make eye colour categorization process more objective, a cluster analysis approach based on the coordinates in CIELAB space was carried out. To this purpose, several clustering algorithms were experimented, including Affinity Propagation (AP)²⁷, Balanced Iterative Reducing and Clustering using Hierarchies (BIRCH)²⁸, Density-Based Spatial Clustering of Applications with Noise (DBSCAN)²⁹, hierarchical clustering (hclust)³⁰, k-means³¹, k-medoids³², k-modes³³, mean-shift³⁴, Ordering Points To Identify the Clustering Structure (OPTICS)³⁵, and Spectral Clustering (SC)³⁶. The settings adopted for each of the algorithms is indicated in Supplementary Table 1. Each clustering algorithm was applied on the original CIELAB values as well as on normalized values. The Euclidean distance was used in conjunction with all the methods requiring a distance metric. Preliminary analyses with a distance metric specifically designed for the CIELAB space, namely the CIEDE2000³⁷, produced results comparable with the ones obtained with the Euclidean distance. Thus, we decided to only use the latter, simpler metric rather than CIEDE2000. The optimal clustering solution was chosen according to the silhouette criterion³⁸, while the agreement of each clustering solution with the categorization obtained through visual inspection was assessed through the adjusted rand index³⁹. It should be noticed that among the clustering solutions identified by cluster analysis, since the IrisPlex model was developed for the prediction of three eye colour categories, we evaluated only the solutions providing at least three groups. In particular, solutions with four groups were taken into account only because we condensed two clusters in a single intermediate category.

Genetic markers. Genetic profiling was carried out on the DNA extracted from buccal swab samples by analysing the genetic polymorphisms included in the IrisPlex¹⁰. Genotyping was performed using TaqMan genotyping assays following manufacturer's instruction and 10 ng of DNA mixed with the TaqMan Genotyping Master Mix (Thermo Fisher Scientific).

The IrisPlex model. From a statistical point of view the IrisPlex system exploits a multinomial logistic regression model by which each individual is classified as being brown, blue or intermediate based on the three obtained prediction probabilities¹⁰. The parameters of such a model were estimated using phenotype and genotype data modeled in an additive fashion (number of minor alleles in the genotype). Prediction with the IrisPlex model were obtained using the dedicated webtool (<https://hirisplex.erasmusmc.nl/>). As suggested by the authors, the predicted colour was the one with a probability higher than the threshold of 0.7. Individuals with all the colour probabilities under 0.7 were marked as “undefined”. Additionally, we also applied a threshold of 0.5. When no threshold was applied, the predictions were assigned to the colour with the absolute highest probability. In this last case, individuals that obtained equal probabilities for multiple (two or three) colour categories were classified as intermediate.

Data availability

The dataset generated during and/or analysed during the current study are not publicly available due to ethical concerns but is available from the corresponding author on reasonable request.

Received: 30 March 2022; Accepted: 21 July 2022

Published online: 27 July 2022

References

- Kaysner, M. Forensic DNA phenotyping: Predicting human appearance from crime scene material for investigative purposes. *Forensic Sci. Int. Genet.* **18**, 33–48. <https://doi.org/10.1016/j.fsigen.2015.02.003> (2015).
- Kaysner, M. & de Knijff, P. Improving human forensics through advances in genetics, genomics and molecular biology. *Nat. Rev. Genet.* **12**, 179–192. <https://doi.org/10.1038/nrg2952> (2011).
- Kaysner, M. & Schneider, P. M. DNA-based prediction of human externally visible characteristics in forensics: Motivations, scientific challenges, and ethical considerations. *Forensic Sci. Int. Genet.* **3**, 154–161. <https://doi.org/10.1016/j.fsigen.2009.01.012> (2009).
- Chaitanya, L. *et al.* The HIRISplex-S system for eye, hair and skin colour prediction from DNA: Introduction and forensic developmental validation. *Forensic Sci. Int. Genet.* **35**, 123–135. <https://doi.org/10.1016/j.fsigen.2018.04.004> (2018).
- Han, J. *et al.* A genome-wide association study identifies novel alleles associated with hair color and skin pigmentation. *PLoS Genet.* **4**, e1000074. <https://doi.org/10.1371/journal.pgen.1000074> (2008).
- Sulem, P. *et al.* Genetic determinants of hair, eye and skin pigmentation in Europeans. *Nat. Genet.* **39**, 1443–1452. <https://doi.org/10.1038/ng.2007.13> (2007).
- Visser, M., Kaysner, M. & Palstra, R.-J. HERC2 rs12913832 modulates human pigmentation by attenuating chromatin-loop formation between a long-range enhancer and the OCA2 promoter. *Genome Res.* **22**, 446–455. <https://doi.org/10.1101/gr.128652.111> (2012).
- Simcoe, M. *et al.* Genome-wide association study in almost 195,000 individuals identifies 50 previously unidentified genetic loci for eye color. *Sci. Adv.* **7**, eabd61239. <https://doi.org/10.1126/sciadv.abd1239> (2020).
- Suarez, P., Baumer, K. & Hall, D. Further insight into the global variability of the OCA2-HERC2 locus for human pigmentation from multiallelic markers.
- Walsh, S. *et al.* IrisPlex: A sensitive DNA tool for accurate prediction of blue and brown eye colour in the absence of ancestry information. *Forensic Sci. Int. Genet.* **5**, 170–180. <https://doi.org/10.1016/j.fsigen.2010.02.004> (2011).
- Pospiech, E. *et al.* The common occurrence of epistasis in the determination of human pigmentation and its impact on DNA-based pigmentation phenotype prediction. *Forensic Sci. Int. Genet.* **11**, 64–72. <https://doi.org/10.1016/j.fsigen.2014.01.012> (2014).
- Ruiz, Y. *et al.* Further development of forensic eye color predictive tests. *Forensic Sci. Int. Genet.* **7**, 28–40. <https://doi.org/10.1016/j.fsigen.2012.05.009> (2013).
- Spichenok, O. *et al.* Prediction of eye and skin color in diverse populations using seven SNPs. *Forensic Sci. Int. Genet.* **5**, 472–478. <https://doi.org/10.1016/j.fsigen.2010.10.005> (2011).
- Hart, K. L. *et al.* Improved eye- and skin-color prediction based on 8 SNPs. *Croat. Med. J.* **54**, 248–256. <https://doi.org/10.3325/cmj.2013.54.248> (2013).
- Salvoro, C. *et al.* Performance of four models for eye color prediction in an Italian population sample. *Forensic Sci. Int. Genet.* **40**, 192–200. <https://doi.org/10.1016/j.fsigen.2019.03.008> (2019).
- Andersen, J. D. *et al.* Genetic analyses of the human eye colours using a novel objective method for eye colour classification. *Forensic Sci. Int. Genet.* **7**, 508–515. <https://doi.org/10.1016/j.fsigen.2013.05.003> (2013).
- Liu, F. *et al.* Digital quantification of human eye color highlights genetic association of three new loci. *PLoS Genet.* **6**, e1000934. <https://doi.org/10.1371/journal.pgen.1000934> (2010).
- Wollstein, A. *et al.* Novel quantitative pigmentation phenotyping enhances genetic association, epistasis, and prediction of human eye colour. *Sci. Rep.* <https://doi.org/10.1038/srep43359> (2017).
- Edwards, M. *et al.* Iris pigmentation as a quantitative trait: Variation in populations of European, East Asian and South Asian ancestry and association with candidate gene polymorphisms. *Pigment Cell Melanoma Res.* **29**, 141–162. <https://doi.org/10.1111/pcmr.12435> (2016).
- Dario, P. *et al.* Assessment of IrisPlex-based multiplex for eye and skin color prediction with application to a Portuguese population. *Int. J. Legal Med.* **129**, 1191–1200. <https://doi.org/10.1007/s00414-015-1248-5> (2015).
- Dembinski, G. M. & Picard, C. J. Evaluation of the IrisPlex DNA-based eye color prediction assay in a United States population. *Forensic Sci. Int. Genet.* **9**, 111–117. <https://doi.org/10.1016/j.fsigen.2013.12.003> (2014).
- Kastelic, V., Pospiech, E., Draus-Barini, J., Branicki, W. & Drobnic, K. Prediction of eye color in the Slovenian population using the IrisPlex SNPs. *Croat Med. J.* **54**, 381–386. <https://doi.org/10.3325/cmj.2013.54.381> (2013).
- Meyer, O. S., Borsting, C. & Andersen, J. D. Perception of blue and brown eye colours for forensic DNA phenotyping. *Forensic Sci. Int. Genet. Suppl. Ser.* **7**, 476–477. <https://doi.org/10.1016/j.figss.2019.10.057> (2019).
- Meyer, O. S. *et al.* Prediction of eye colour in Scandinavians using the EyeColour 11 (EC11). *SNP Set.* **12**, 821 (2021).
- Pietroni, C. *et al.* The effect of gender on eye colour variation in European populations and an evaluation of the IrisPlex prediction model. *Forensic Sci. Int. Genet.* **11**, 1–6. <https://doi.org/10.1016/j.fsigen.2014.02.002> (2014).
- Eiberg, H. *et al.* 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**, 177–187. <https://doi.org/10.1007/s00439-007-0460-x> (2008).
- Frey, B. J. & Dueck, D. Clustering by passing messages between data points. *Science* **315**, 972–976. <https://doi.org/10.1126/science.1136800> (2007).
- Zhang, T., Ramakrishnan, R. & Livny, M. in *Proceedings of the 1996 ACM SIGMOD international conference on Management of data* 103–114 (Association for Computing Machinery, Montreal, Quebec, Canada, 1996).

29. Schubert, E., Sander, J., Ester, M., Kriegel, H. P. & Xu, X. DBSCAN Revisited, Revisited: Why and How You Should (Still) Use DBSCAN. *42*, 19. <https://doi.org/10.1145/3068335> (2017).
30. Ward, J. H. Hierarchical grouping to optimize an objective function. *J. Am. Stat. Assoc.* **58**, 236–244. <https://doi.org/10.1080/01621459.1963.10500845> (1963).
31. Arthur, D. & Vassilvitskii, S. in *Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms* 1027–1035 (Society for Industrial and Applied Mathematics, New Orleans, Louisiana, 2007).
32. Park, H.-S. & Jun, C.-H. A simple and fast algorithm for K-medoids clustering. *Expert Syst. Appl.* **36**, 3336–3341. <https://doi.org/10.1016/j.eswa.2008.01.039> (2009).
33. Huang, Z. Extensions to the k-Means algorithm for clustering large data sets with categorical values. *Data Min. Knowl. Disc.* **2**, 283–304. <https://doi.org/10.1023/A:1009769707641> (1998).
34. Comaniciu, D. & Meer, P. Mean shift: A robust approach toward feature space analysis. *IEEE Trans. Pattern Anal. Mach. Intell.* **24**, 603–619. <https://doi.org/10.1109/34.1000236> (2002).
35. Ankerst, M., Breunig, M. M., Kriegel, H.-P. & Sander, J. OPTICS: Ordering points to identify the clustering structure. *Science* **28**, 49–60. <https://doi.org/10.1145/304181.304187> (1999).
36. Jianbo, S. & Malik, J. Normalized cuts and image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **22**, 888–905. <https://doi.org/10.1109/34.868688> (2000).
37. Sharma, G., Wu, W. & Dalal, E. N. The CIEDE2000 color-difference formula: Implementation notes, supplementary test data, and mathematical observations. *Color. Res. Appl.* **30**, 21–30. <https://doi.org/10.1002/col.20070> (2005).
38. Rousseeuw, P. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Science* **20**, 53–65. [https://doi.org/10.1016/0377-0427\(87\)90125-7](https://doi.org/10.1016/0377-0427(87)90125-7) (1987).
39. Steinley, D. Properties of the Hubert-Arabie adjusted Rand index. *Psychol. Methods* **9**, 386–396. <https://doi.org/10.1037/1082-989X.9.3.386> (2004).

Author contributions

A.M. and G.P.: study design. E.P., S.G., E.F. and A.B.: sample recruitment, genetic analyses and data collection. A.M., V.L. and A.G.: data analyses. The Ms was initially drafted by E.P. and A.M. and then finalized by all authors. All authors contributed to the article and approved the submitted version.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-17208-w>.

Correspondence and requests for materials should be addressed to A.M.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022