Prediction of eye, hair and skin colour in Latin Americans

Sagnik Palmal, Kaustubh Adhikari, Javier Mendoza-Revilla, Macarena Fuentes-Guajardo, Caio Cesar Silva de Cerqueira, Betty Bonfante, Juan Camilo Chacón-Duque, Anood Sohail, Malena Hurtado, Valeria Villegas, et al.

▶ **To cite this version:**

Sagnik Palmal, Kaustubh Adhikari, Javier Mendoza-Revilla, Macarena Fuentes-Guajardo, Caio Cesar Silva de Cerqueira, et al.. Prediction of eye, hair and skin colour in Latin Americans. Forensic Science International: Genetics , 2021, 53, pp.102517. 10.1016/j.fsigen.2021.102517 . hal-03268294

HAL Id: hal-03268294
https://hal.science/hal-03268294v1

Submitted on 24 Apr 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

Version of Record: https://www.sciencedirect.com/science/article/pii/S1872497321000557
Manuscript_2d55652bf234e9b131ae8f4efb5c8054

PREDICTION OF EYE, HAIR AND SKIN COLOUR IN LATIN AMERICANS

Sagnik Palmal[1*], Kaustubh Adhikari[2,3*], Javier Mendoza-Revilla[4,5], Macarena Fuentes-Guajardo[6], Caio C. Silva de Cerqueira[7], Betty Bonfante[1], Juan Camilo Chacón-Duque[8], Anood Sohail[9], Malena Hurtado[4], Valeria Villegas[4], Vanessa Granja[4], Claudia Jaramillo[9], William Arias[10], Rodrigo Barquera Lozano[11,12], Paola Everardo-Martínez[11], Jorge Gómez-Valdés[11], Hugo Villamil-Ramírez[13], Tábita Hünemeier[14], Virginia Ramallo[15,16], Maria-Laura Parolin[17], Rolando Gonzalez-José[16], Lavinia Schüler-Faccini[15], Maria-Cátira Bortolini[15], Victor Acuña-Alonzo[11], Samuel Canizales-Quinteros[13], Carla Gallo[4], Giovanni Poletti[4], Gabriel Bedoya[10], Francisco Rothhammer[18,19], David Balding[3,20], Pierre Faux[1#], Andrés Ruiz-Linares[21,1,3#]

[1] UMR 7268 ADES, CNRS, Aix-Marseille Université, EFS, Faculté de Médecine Timone, Marseille, 13005, France.
[2] School of Mathematics and Statistics, Faculty of Science, Technology, Engineering and Mathematics, The Open University, Milton Keynes, MK7 6AA, United Kingdom.
[3] Department of Genetics, Evolution and Environment, and UCL Genetics Institute, University College London, London, WC1E 6BT, UK
[4]Laboratorios de Investigación y Desarrollo, Facultad de Ciencias y Filosofía, Universidad Peruana Cayetano Heredia, Lima, 31, Perú.
[5] Unit of Human Evolutionary Genetics, Institut Pasteur, Paris, 75015, France.
[6] Departamento de Tecnología Médica, Facultad de Ciencias de la Salud, Universidad de Tarapacá, Arica, 1000000, Chile.
[7] Scientific Police of São Paulo State, Ourinhos-SP, 19900-109, Brazil.
[8] Division of Vertebrates and Anthropology, Department of Earth Sciences, Natural History Museum, London, SW7 5BD, UK.
[9] Department of Biotechnology, Kinnaird College for Women, 93 - Jail Road, Lahore, 54000, Pakistan.
[10] GENMOL (Genética Molecular), Universidad de Antioquia, Medellín, 5001000, Colombia.
[11] National Institute of Anthropology and History, Mexico City, 6600, Mexico.
[12] Department of Archaeogenetics, Max Planck Institute for the Science of Human History (MPI-SHH), Jena, 07745, Germany.
[13] Unidad de Genomica de Poblaciones Aplicada a la Salud, Facultad de Química, UNAM-Instituto Nacional de Medicina Genómica, Mexico City, 4510, Mexico.
[14] Departamento de Genética e Biologia Evolutiva, Instituto de Biociências, Universidade de São Paulo, São Paulo, SP, 05508-090, Brazil.
[15] Departamento de Genética, Universidade Federal do Rio Grande do Sul, Porto Alegre, 90040-060, Brasil.
[16] Instituto Patagónico de Ciencias Sociales y Humanas, Centro Nacional Patagónico, CONICET, Puerto Madryn, U9129ACD, Argentina.
[17] Instituto de Diversidad y Evolución Austral (IDEAus), Centro Nacional Patagónico, CONICET, Puerto Madryn, Argentina.
[18] Instituto de Alta Investigación, Universidad de Tarapacá, Arica, Arica 1000000, Chile.
[19] Programa de Genetica Humana, ICBM, Facultad de Medicina, Universidad de Chile, Santiago, Arica 1000000, Chile.
[20] Melbourne Integrative Genomics, Schools of BioSciences and Mathematics & Statistics, University of Melbourne, Melbourne, VIC 3010, Australia.
[21] Ministry of Education Key Laboratory of Contemporary Anthropology and Collaborative Innovation Center of Genetics and Development, School of Life Sciences and Human Phenome Institute, Fudan University, Yangpu District, Shanghai, China.

© 2021 published by Elsevier. This manuscript is made available under the CC BY NC user license
https://creativecommons.org/licenses/by-nc/4.0/

*Contributed equally to this work. #Co-supervised this work
Corresponding authors: andresruiz@fudan.edu.cn; pierre.faux@univ-amu.fr

Acknowledgments

We thank the volunteers for their enthusiastic support for this research. We also thank Alvaro Alvarado, Mónica Ballesteros Romero, Ricardo Cebrecos, Miguel Ángel Contreras Sieck, Francisco de Ávila Becerril, Joyce De la Piedra, María Teresa Del Solar, William Flores, Martha Granados Riveros, Rosilene Paim, Ricardo Gunski, Ana Angélica Leal Barbosa, Sergeant João Felisberto Menezes Cavalheiro, Major Eugênio Correa de Souza Junior, Wendy Hart, Ilich Jafet Moreno, Paola León-Mimila, Francisco Quispealaya, Diana Rogel Diaz, Ruth Rojas, and Vanessa Sarabia, for assistance with volunteer recruitment, sample processing and data entry. We are very grateful to the institutions that allowed the use of their facilities for the assessment of volunteers, including: Escuela Nacional de Antropología e Historia and Universidad Nacional Autónoma de México (México); Universidade Federal do Rio Grande do Sul (Brazil); 13° Companhia de Comunicações Mecanizada do Exército Brasileiro (Brazil); Pontificia Universidad Católica del Perú, Universidad de Lima and Universidad Nacional Mayor de San Marcos (Perú). We also acknowledge the Centre de Calcul Intensif d'Aix-Marseille for granting access to high-performance computing resources. We thank Manfred Kayser for assistance with Online HIrisPlex-S.

Funding: Work leading to this publication was funded by grants from: the Leverhulme Trust (F/07 134/DF), BBSRC (BB/I021213/1), the Excellence Initiative of Aix-Marseille University - A*MIDEX (a French "Investissements d'Avenir" programme), Universidad de Antioquia (CODI sostenibilidad de grupos 2013- 2014 and MASO 2013-2014), the National Natural Science Foundation of China (#31771393), the Scientific and Technology Committee

of Shanghai Municipality (18490750300), Ministry of Science and Technology of China (2020YFE0201600), Shanghai Municipal Science and Technology Major Project (2017SHZDZX01) and the 111 Project (B13016), Conselho Nacional de Desenvolvimento Científico e Tecnológico, Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul (Apoio a Núcleos de Excelência Program), Fundação de Aperfeiçoamento de Pessoal de Nível Superior.

Competing interests: The authors declare that they have no competing interests.

Abstract

Here we evaluate the accuracy of prediction for eye, hair and skin pigmentation in a dataset of > 6,500 individuals from Mexico, Colombia, Peru, Chile and Brazil (including genome-wide SNP data and quantitative/categorical pigmentation phenotypes - the CANDELA dataset CAN). We evaluated accuracy in relation to different analytical methods and various phenotypic predictors. As expected from statistical principles, we observe that quantitative traits are more sensitive to changes in the prediction models than categorical traits. We find that Random Forest or Linear Regression are generally the best performing methods. We also compare the prediction accuracy of SNP sets defined in the CAN dataset (including 56, 101 and 120 SNPs for eye, hair and skin colour prediction, respectively) to the well-established HIrisPlex-S SNP set (including 6, 22 and 36 SNPs for eye, hair and skin colour prediction respectively). When training prediction models on the CAN data, we observe remarkably similar performances for HIrisPlex-S and the larger CAN SNP sets for the prediction of hair (categorical) and eye (both categorical and quantitative), while the CAN sets outperform HIrisPlex-S for quantitative, but not for categorical skin pigmentation prediction. The performance of HIrisPlex-S, when models are trained in a world-wide sample (although consisting of 80% Europeans, https://hirisplex.erasmusmc.nl), is lower relative to training in the CAN data (particularly for hair and skin colour). Altogether, our observations are consistent with common variation of eye and hair colour having a relatively simple genetic architecture, which is well captured by HIrisPlex-S, even in admixed Latin Americans (with partial European ancestry). By contrast, since skin pigmentation is a more polygenic trait, accuracy is more sensitive to prediction SNP set size, although here this effect was only apparent for a quantitative measure of skin pigmentation. Our results support the use of HIrisPlex-S in the prediction of categorical pigmentation

traits for forensic purposes in Latin America, while illustrating the impact of training datasets on its accuracy.

Keywords: DNA phenotyping, eye-colour, hair-colour, skin-colour, pigmentation prediction, admixture, Latin Americans

Introduction

There is growing interest in the use of genetic data for the prediction of physical appearance, particularly in forensic, historical and paleo-anthropological studies[1–3]. Strong impetus for these studies has been provided by Genome Wide Association Studies (GWAS) of traits such as eye, hair and skin pigmentation, which show variation within and between continental populations. In the case of eye and hair colour, a large variation range is seen in Europeans [4–9]**, while other continents have limited variation, essentially within the narrow brown-black range** [8,10–12]**. By contrast, variation in skin colour in Europeans has a narrower range than in other continents. In terms of characterizing the genetic basis of variation in pigmentation traits, so far, the great majority of GWASs have been performed in Europeans[13–18], although recent GWAS in non-Europeans are enabling the identification of additional variants associated with pigmentation variation outside Europe**[12,19]**, particularly for skin colour.**

Early studies on prediction of pigmentation traits, exploiting GWAS findings, focused on eye colour [20–23] **and these analyses were subsequently extended to hair [24,25] and skin [26]. As a result, sets of SNPs have now been proposed for the simultaneous prediction of eye, hair and skin colour[27,28]. These SNP sets have been shown to have high prediction accuracy in European samples and there is now interest in evaluating the performance of these tools in non-European populations, as well as in populations of mixed continental ancestry. Latin Americans represent one of the largest recently admixed populations world-wide. The history of Latin America has involved extensive admixture, mostly between Native Americans, Europeans and sub-Saharan Africans. Consistent with its partly European ancestry, a recent GWAS for pigmentation traits in Latin Americans in the CANDELA cohort detected phenotypic**

effects for a number of loci previously identified in Europeans[12]. In addition to these, novel pigmentation SNPs with genome-wide significant association were also identified in that study. These included SNPs polymorphic only in East Asians and Native Americans, consistent with the independent evolution of skin pigmentation in West and East Eurasia[7,29,30]. The admixed ancestry of Latin America and the finding in the region of pigmentation variants not present in Europeans emphasizes the need to evaluate the accuracy of tools currently available for prediction of pigmentation traits in this population.

Here we aimed to evaluate the accuracy of prediction of pigmentation traits in a large Latin American dataset from Mexico, Colombia, Peru, Chile and Brazil [12,31–36] **characterized for eye colour, comparing methods, model predictors, and training datasets. We compared results from the widely-used HIrisPlex-S with sets of SNPs selected from the CANDELA data (and including a larger number of SNPs than HIrisPlex-S). We find that, when trained in the CANDELA data, the HIrisPlex-S SNP set has a performance similar to the CAN SNPs in the prediction of eye and hair colour, but its performance is somewhat lower for skin colour. This observation is consistent with the polygenicity of skin colour, relative to hair and (particularly) eye colour and the large variation in skin colour across the world. The work presented here sets the stage for the optimization of tools for the prediction of pigmentation traits across Latin America, for forensic purposes.**

Materials and Methods

Study sample: phenotypes, genetic data and covariates

We analyzed data previously studied by the CANDELA consortium for GWAS of pigmentation traits[12,31–34]. The consortium gathered genetic and phenotypic data from

over 6,500 individuals recruited in five Latin American countries: Mexico (N=~1,200), Colombia (N=~1,700), Peru(N=~1,230), Chile (N=~1,730) and Brazil (N=~630).

Pigmentation traits evaluated directly on the research subjects consists of (A) hair pigmentation (recorded in four categories: 1-red/reddish, 2-blond, 3-dark blond/light brown or 4-brown/black. However, due to their very low frequency (<0.6%), individuals in the 'red/reddish' category were not included here), (B) eye colour, recorded as five ordered categories: 1-blue/grey, 2-honey, 3-green, 4-light brown, 5-dark brown/black. For increasing consistency with previous publications[23,26,37–39], here we recoded these data into just three categories: 1-Blue/Grey, 2-Intermediate (honey or green) and 3-Brown/Black (light brown or dark brown/black) and (C) a quantitative measure of skin pigmentation from an area unexposed to sunlight (the Melanin Index MI, obtained by reflectometry). We also had available additional measures of iris pigmentation, extracted from digital photographs, using the HCL colour space (Hue, Chroma and Luminance). Hue being an angle (recorded in arc degree), we linearized this trait with cosine and shifted the angle by 15° in order to maximize the number of samples in the range [0,180°]; hence the trait considered is cos(Hue+15). The frequency distribution for these traits in the CANDELA dataset is shown in Supplementary Figure S1.

To enable comparison with previous studies on categorical skin colour [26,37], we converted the quantitative MI values into a three-level categorical trait (Fair, Intermediate and Dark skin colour). For this, we used the individual genetic ancestry estimates in the CANDELA dataset to select individuals with estimated 100% European ancestry (N=70) and individuals with African ancestry higher than both the European and Native estimates (i.e. >39% African ancestry; N=23). Based on the MI distribution in these two groups, we defined MI values of 33 and 47 as thresholds for three skin colour categories: Fair (MI <33; N= 2,506), Intermediate (MI 33-47; N=3,840) and Dark

(MI>47; N=180) (Supplementary figure S2). These thresholds are in line with values obtained in a previous study of Brazilians[40].

The genetic data consisted of ~9 million genotypes, ~700k of which were obtained experimentally by genotyping Illumina's Omni Express chip, the remainder obtained by imputation as described in Adhikari et al[12]. We applied several filters to the CANDELA dataset prior to the trait prediction analyses. Firstly, we retained only individuals aged 18 to 45. Secondly, we removed 8 pairs of individuals whose pairwise probability of IBD was estimated close to 1, to discard potential sample mix-ups (hence, 16 individuals removed), and individuals whose estimated African ancestry was more than European and native ancestry estimations (23 individuals), as those were considered as genetic outliers and thus excluded. Finally, we excluded all individuals with missing data on any of the covariates (age, sex, BMI). Note that BMI was considered as a covariate since we found it significantly correlated to some pigmentation phenotypes; that correlation is most likely a confounding effect of continental genetic ancestry. The final sample size used in the analyses was: 6,495 for hair colour, 6,526 for MI, 6,529 for categorical eye colour and 5,738 for quantitative eye colour traits – Hue, Chroma and Luminance. These three eye colour phenotypes constitute the bicone colour space model – HCL scale for human perception of eye colours (previously explained by Adhikari et al[12]).

Pigmentation SNP sets used for prediction

We used two sets of SNPs for the prediction analyses. Firstly, we devised "CANDELA" (CAN) SNP sets for prediction of each pigmentation trait (E-eye; H-hair; S-skin) based on results from a GWAS conducted in the CANDELA sample[12]. To pre-select SNPs for each trait, we used the following protocol: (1) selection of all SNPs with GWAS association *p*-values $<10^{-5}$, (2) grouping SNPs in high LD and (3) for each SNP group,

selection of the SNP with highest predictive power (as the most significant SNP is not necessarily the most predictive [41]). We thus pre-selected 1,471, 207 and 701 SNPs for skin, hair and eye pigmentation prediction, respectively. **For each trait, the preselected SNPs were ranked based on decreasing conditional predictive power and R^2 of prediction models computed sequentially, each time adding a SNP from the ranked list. We set a limit for the number of SNPs included in a set when $R^2(i) >= max(R^2)*.999$**. Details of the approach used for SNP selection are provided in Supplementary method S1, and the resulting CAN-E, CAN-S and CAN-H SNP sets are described in the Results section.

Secondly, as a benchmark, we used HIrisPlex-S, a SNP set that has been developing over the years for the prediction of eye, hair and skin pigmentation [20,24,27] HIrisPlex-S currently includes 41 SNPs, of which 6/22/36 are relevant respectively for prediction of eye/hair/skin pigmentation. Of the 41 SNPs included in HIrisPlex-S 22 SNPs were directly genotyped in the CANDELA samples, the remaining having been imputed. We only retained SNPs with MAF >= 1% in the CANDELA data. This led to seven HIrisPlex-S SNPs being excluded, thus reducing the set used here to 34 SNPs (6/16/32 of these being used for eye/hair/skin prediction, respectively, Supplementary table S1). Due to the low MAF frequency in the CANDELA data, the exclusion of these SNPs has a negligible impact on prediction accuracy. The lack of informativity of the 7 discarded SNPs in the CANDELA dataset is underlined by the fact that they are all located in the *MC1R* region, and are important mainly for red hair prediction, a trait nearly absent in the CANDELA data (the few red/reddish individuals with red hair were removed due to their low frequency). Of the 34 SNPs retained for the analyses, 13 had imputed genotypes, all with high imputation quality metrics (IMPUTE2's INFO > 0.8 – see Supplementary table S1).

Overall, for all the imputed SNPs used in the analyses (either from CAN or HIrisPlex SNP sets), the average imputation quality metric (INFO) was high, 0.942. In addition, we verified the accuracy of imputed genotypes for these SNPs by comparing with two independently sequenced datasets. A set of Native American samples collected and chip genotyped for a previous study [35] were sequenced at high coverage and variants filtered. We calculated the concordance for these samples as the proportion of imputed genotypes that match the sequence data exactly. The average concordance was high for these SNPs, 98.8%. For another set of sequenced European samples, the average concordance for these SNPs were equally high at 98.5%.

Prediction methods and models evaluated

A broad array of statistical methods have been employed in the literature to predict pigmentation traits, such as (multiple) linear[42,43] or (multinomial) logistic regression[23,24], decision trees[42,44], neural networks[42,45], and naïve Bayes classifiers[37,46,47]. Each method has its advantages and disadvantages, and are better suited for certain types of traits, e.g. linear regression for quantitative traits[42,43] and logistic regression for categorical traits[23,24].

The overall strategy we used for performing pigmentation prediction in the CANDELA dataset is shown in Figure 1. Linear Regression (LR) or Multinomial Logistic Regression (MLR) were used as the reference methods for quantitative or categorical traits, respectively. These two methods were used to evaluate three prediction models, incorporating an increasing number of predictors (Fig. 1A):

1- Using only non-genetic covariates as predictors:

$$y \sim Age + Sex + BMI \qquad \text{(equation 1)}$$

2- Incorporating genetic ancestry to model 1:

$$y \sim Age + Sex + BMI + EUR_{ancestry} + AFR_{ancestry} \quad \text{(equation 2)}$$

Here we included as predictors the estimates of European and African ancestry (obtained by unsupervised admixture estimation on genome-wide data). Native American ancestry was omitted so as to avoid collinearity (since the three continental ancestries sum to 1).

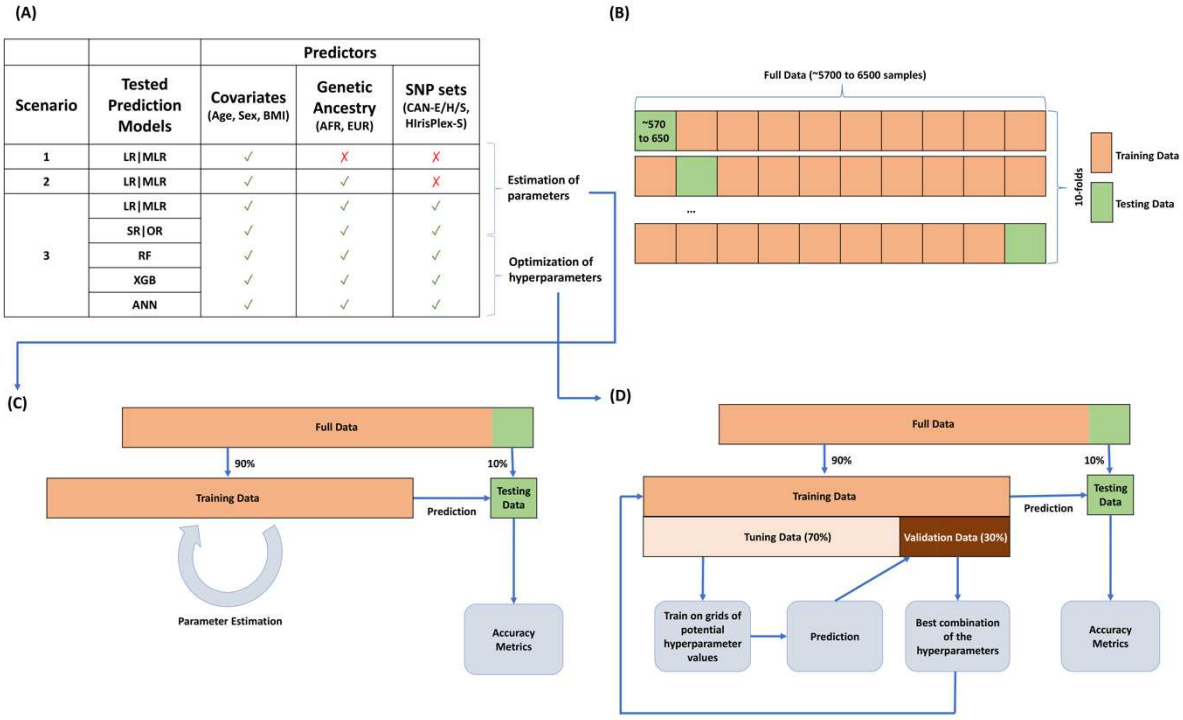
3- Incorporating pigmentation SNPs to model 2:

$$y \sim Age + Sex + BMI + EUR_{ancestry} + AFR_{ancestry} + \sum_{j \in SNPsets} SNP_j, \text{ (equation 3)}$$

where SNPsets refers to SNPs included either in the HIrisPlex-S or CAN sets defined above.

For the third (full) model, in addition to regression, the following statistical and machine-learning methods were used, in order to evaluate their relative performance for prediction: Random Forest (RF), Extreme Gradient Boosting (XGB), Artificial Neural Network (ANN), Ordinal regression (OR) and Stepwise regression (SR). We provide more details on their implementation in Supplementary method S2.

Figure 1. Study overview. (A) Models tested, predictors used and prediction methods: multinomial regression (MLR), linear regression (LR), ordinal regression (OR), stepwise regression (SR), random forests (RF), extreme gradient boosting (XGB) and artificial neural network (ANN). **(B)** 10-fold cross-validation: the full data is randomly split into 10 equally-sized data sub-groups. For each of the 10 sub-groups, the estimation of model parameters **(C)** or optimization of model hyperparameters **(D)** was performed on a pool of the nine remaining sub-groups.



Evaluation of prediction accuracy

To measure prediction accuracy in quantitative traits, we used the coefficient of determination (R^2, the proportion of phenotype variance that is explained by the model, measured as $1 - SS_{res}/SS_{tot}$, where SS_{res} and SS_{tot} respectively stand for the residual and total sum of squares). For categorical traits, we used a metric denoted "accuracy" that is the proportion of correctly classified individuals in confusion matrices. For these traits, we also computed the Area Under the ROC Curve (AUC – ranging from 0.5 to 1) as well as the expected accuracies from two benchmark strategies (either using the categories' frequency – PropStrat, or always guessing the most frequent category – maxP; see Supplementary Method S3) for comparison.

Accuracy of prediction was evaluated using 10-fold cross validation (10-fold CV) (Fig. 1B). For this, the full dataset is split into 10 approximately equal subsets based on a stratified sampling on the trait (so that trait distribution is similar across subsets). Each of the $1/10^{th}$ subsets is used as test data for evaluating prediction accuracy of methods trained using

the other 9/10th of the data. For regression methods (MLR, LR, SR, OR), coefficient parameters are estimated in the training data (Fig. 1C). For Machine Learning models, the training data is further split into a tuning data (70% of the training data) and validation data (30% of the training data). The parameter space for these methods are tuned creating a grid of all possible combinations of the hyperparameters and the particular combination producing best result on the validation data is selected as the set of optimal combination for the whole training data (see Supplementary Method S2). **Each one of the ten folds is taken as the test data in turn, while the rest nine folds are used for training, producing** 10 different train-test data combinations and the hyperparameters are tuned based on that subsequently gives rise to 10 prediction accuracy results (Fig. 1D). These measures of goodness of fit are used in a boxplot, or the average of them is used as a single prediction accuracy metric of the method. This helps us in avoiding inflation in the results and the predictions are more robust to small changes in the data.

Comparison with prediction accuracy from HIrisPlex-S-online

Several studies have examined the prediction accuracy of HIrisPlex-S for categorical pigmentation traits using an online tool (https://hirisplex.erasmusmc.nl, referred to here as HIrisPlex-S-Online). This implementation uses MLR prediction models trained in a reference dataset comprising individuals with various continental origins (80% from Europe, 16% from North-America – including individuals of European, African and Asian ancestry – and 4% from Africa and Oceania)[20,24,27]. For Eye colour HIrisPlex-S-Online predicts eye colour categories corresponding to those used in the CANDELA dataset (Blue/Grey; Intermediate and Brown/Black). For hair colour, HIrisPlex-S-Online predicts the categories Red, Blond, Brown and Black. The last three correspond to hair categories common in the CANDELA dataset. However, due to their

low frequency, individuals in the CANDELA dataset with Red hair were removed prior to these analyses. We therefore adjusted the prediction probabilities estimated by HIrisPlex-S-Online in the CANDELA dataset, for Blond, Brown and Black by the probability of Red hair colour. Finally, HIrisPlex-S-Online predicts five categories of Skin colour (modified from the Fitzpatrick scale): Very pale, Pale, Intermediate, Dark and Dark-Black. To allow comparison with the three-category predictions from models trained in the CANDELA data, we followed Walsh et al. 2017 in merging the prediction probabilities for Very pale, Pale and Intermediate into one category (Light, corresponding to the "White" category of Fitzpatrick). The Dark and Dark-Black of HIrisPlex-S-Online correspond to the Brown and Black categories of the Fitzpatrick scale (Walsh et al. 2017). This three-level categorization thus matches the one defined in the CANDELA dataset based on the transformation of MI values (described above)

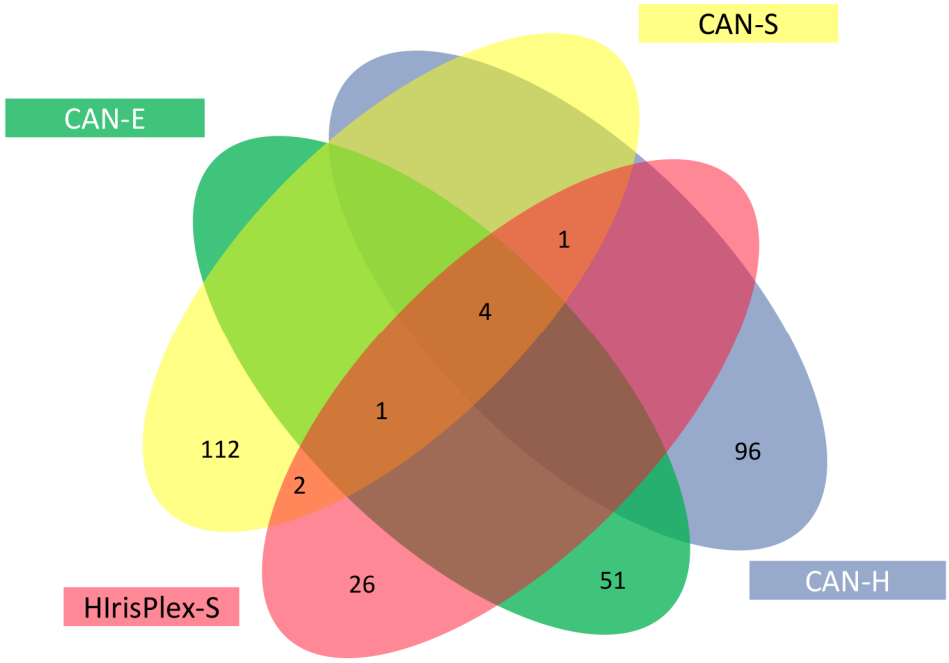
Prediction of MI in Native American individuals of unknown phenotype

Genotype and geo-localization data for 117 Native American individuals (with an estimated >99% Native American ancestry) from 17 Native American populations were available from a previous study[35]. We analysed genotypes for these 'pure' Native American individuals, predicted their MI, and regressed these predicted values on the amount of solar radiation at the site of population sampling. We trained two RF models (one with CAN-S and one with HIrisPlex-S) using 550 CANDELA individuals with >= 80% native ancestry and using sex as the only covariate. Solar radiation levels were defined as insolation incident on a horizontal surface (in kWh/m^2/day) as reported in the NASA Surface meteorology and Solar Energy (SSE) Web site (https://eosweb.larc.nasa.gov/sse/) (data previously used in[12]).

Results

The CANDELA dataset analyzed here consisted of individual genome-wide SNP genotypes and pigmentation traits. In that dataset eye colour was recorded both as categorical and quantitative variables, hair colour as a categorical variable, and skin colour as a quantitative variable (the Melanin Index). In what follows we compare the performance of HIrisPlex-S with SNP sets devised here for the prediction of eye (CAN-E), hair (CAN-H) and skin (CAN-S) from summary statistics of a pigmentation GWAS performed in the CANDELA sample (Materials and Methods and Supplementary method S1). The sets devised here consist of : 56 (CAN-E), 101 (CAN-H) and 120 (CAN-S) pigmentation-associated SNPs and are detailed in Supplementary table S2. Consistent with the genetic correlation of eye, hair and skin pigmentation, some SNPs are shared across the CAN-E/H/S sets, as well as with the HIrisPlex-S set. The overlap between these four SNP sets is shown in Figure 2.

Figure 2. Overlap between SNP sets used for prediction of pigmentation traits. CAN-E, CAN- S and CAN-H refer, respectively, to SNP sets designed here for the prediction of eye, skin and hair pigmentation, based on a GWAS performed in the CANDELA sample[12]. HIrisPlex-S is a SNP assay developed for simultaneous Eye, Hair and Skin colour prediction[27]. Numbers refer to SNPs shared between the SNP sets.

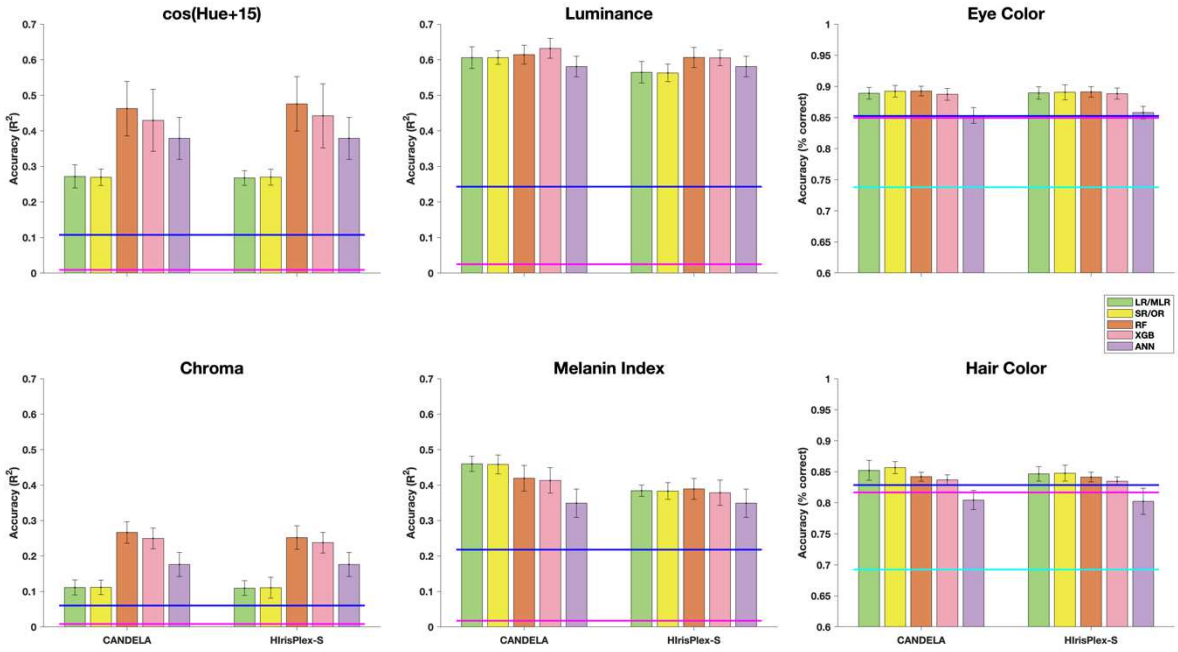


Prediction Accuracy in relation to models, methods and pigmentation SNP sets

Figure 3 presents the accuracy of prediction for various phenotypes of eye, hair and skin colour. For categorical traits, the baseline model (i.e. including only non-genetic predictors: age, sex and BMI), reaches 84.9% and 81.7% accuracy (proportion of correctly classified individuals) for eye and hair colour, respectively. That level of accuracy is actually also reached by always guessing the phenotype to be the most frequent category (maxP strategy, see Supplementary table S3 and Supplementary Method S2). This high accuracy obtained by a deterministic strategy probably relates to the highly skewed trait distribution in the CANDELA individuals: ~ 82% having black/dark brown hair and 85% having Brown/Black eyes. Alternately, randomly guessing the phenotypes based solely on the frequency of the traits (PropStrat; cyan line in Figure 3) also yields good levels of accuracy (~74% and ~69% for categorical eye and hair colour, respectively).

Figure 3. Prediction accuracy in relation to models, methods and pigmentation SNP sets. For continuous traits (Hue (transformed), Luminance, Chroma and Melanin Index; top and middle panels) we used R^2 as measure of prediction accuracy. For categorical traits (Eye

and Hair colour; bottom panels) accuracy is the proportion of correctly classified individuals. Magenta and blue lines indicate the accuracy obtained when only non-genetic predictors or non-genetic + genomic ancestry are included in regression models, respectively. For categorical traits, the performance of a random guessing strategy (PropStrat) was also evaluated (cyan line). For these traits, the average accuracy of the deterministic maxP strategy is numerically the same as the accuracy obtained when only non-genetic predictors are used (magenta line), hence is not shown separately in this figure. For the full prediction model (non-genetic predictors + genetic ancestry + pigmentation SNPs) the performance of regression and four additional prediction methods was evaluated (bars are coloured: green = LR/MLR; yellow = SR/OR; brown = RF; pink = XGB; purple = ANN). Detailed numerical values are given in Supplementary Table S3. The pigmentation SNP set incorporated in the prediction models is indicated at the bottom of the plots.



It is important to keep the maxP strategy in context when assessing prediction performance for categorical traits, since it represents how skewed the trait distribution is – a binary trait with a frequency distribution of 90% and 10% of the two categories will have 90% accuracy under the simplest maxP strategy (even though its sensitivity will be 0 for the rare category; see Supplementary Method S3). Thus, a skew in the trait distribution causes an upward shift in accuracy of prediction methods, especially those methods which are biased to the most frequent category, making them appear better-performing than they actually are. The PropStrat strategy is comparatively less biased as it gives proportional weight to the rare

category (hence non-zero sensitivity for this category), and thus has lower accuracy than the maxP strategy. It is therefore a better benchmark to compare the performance of other strategies for assessing their gain in accuracy. Conversely, a comparison of those strategies to the maxP benchmark better represents their relative change in incorrect classification rather than the relative gain in correct classification (i.e. gain in accuracy).

Although the accuracies of these basic strategies are already high due to our skewed trait distributions, adding genetic ancestry to the model has a further impact, especially for hair colour: it decreases the proportion of error by ~6% (from 18.3% of error to 17.1% - detailed numbers in Supplementary table S3). Then the further addition of pigmentation SNPs has an even larger effect: the remaining proportion of errors decreases by another ~16% and ~27% respectively for hair and eye colour. It is also noticeable that the gain in prediction brought by SNPs relatively to that brought by genetic ancestries is much larger for eye colour (~14*x*) than for hair colour (~3*x*).

For continuous traits, we observe a large increase in prediction accuracy (R^2) when genetic ancestry is incorporated in regression models, relative to the baseline (including only non-genetic predictors). Furthermore, when pigmentation SNP sets are incorporated in the regression model, accuracy usually more than doubles over that obtained with genetic ancestry plus non-genetic covariates (green bar versus blue line in Figure 3). When using this full prediction model, lowest LR prediction accuracy was observed for Chroma (R^2 ~0.12) and highest for Luminance (R^2 ~0.58), two quantitative estimates of eye colour variation.

Comparing different prediction methods for the full model (i.e. incorporating all predictors) we do not observe large differences in performance, with the exception of a relatively lower accuracy of regression for Hue and Chroma. For those two traits, RF markedly outperforms regression methods, more than doubling the accuracy of LR in the case of Chroma. For the two other continuous traits (Luminance and Melanin Index),

regression models are as effective as machine learning models (Figure 3). We also note that those rank similarly throughout categorical and continuous traits: RF is almost always better than the other tree-based model (extreme gradient boosting; XGB) and artificial neural networks (ANN) always underperform compared to tree-based models.

Regarding the two SNP sets tested, we observe little difference between them in prediction accuracy across traits and methods (despite the number of SNPs being considerably larger in the CAN sets than in HIrisPlex-S), except for the skin Melanin Index. For that trait, CAN-S consistently outperforms HIrisPlex-S, particularly with regression methods.

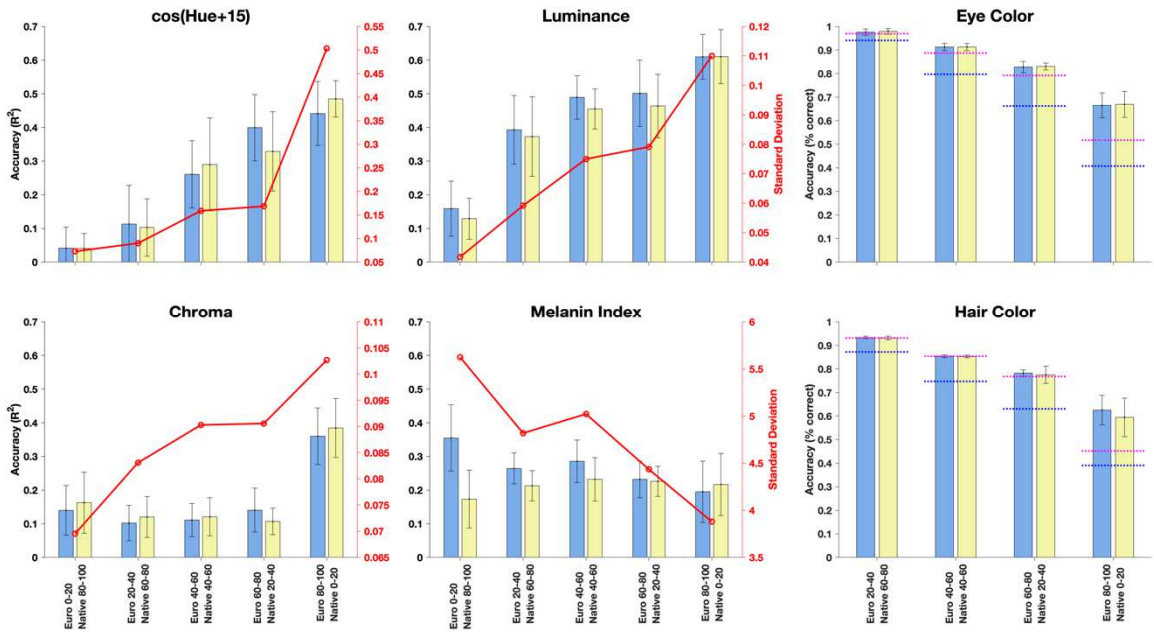
Prediction accuracy at varying levels of European/Native American ancestry

Since a substantial fraction of individuals in the CANDELA sample have minimal African ancestry, we sought to evaluate prediction accuracy specifically for varying levels of European/Native American ancestry in the CANDELA sample. To this aim, we pooled individuals with negligible African ancestry in ~20% ancestry bins (so that the smallest pool size included at least 570 individuals) and examined prediction accuracy in the pools (see Supplementary table S4). Furthermore, for categorical traits, we ensured that at least two trait categories, each with >20 individuals, were observed in the pools, which led to withdraw the most Native-American pool. We assessed prediction accuracy using RF, a full model (i.e. equation 3 but without genetic ancestry as predictor) including only pigmentation SNPs having >1% MAF in the pool of individuals being tested.

For the categorical traits, there is a drop in prediction accuracy (from ~95% to ~70%) at increasing European ancestry (Figure 4 and Supplementary table S5). However, as European ancestry increases there is greater accuracy relative to random guessing based on trait frequency, probably reflecting the trait being less variable at higher Native American

ancestry levels. For the quantitative eye pigmentation variables (particularly H and L, Figure 4), as the percentage of European ancestry increases there is a trend for an increase in trait variation (red line in Figure 4) and also in prediction R^2. For skin pigmentation (MI) we observe an opposite trend in trait variability in relation to ancestry, relative to hair/eye colour: variation in MI decreases at increasing European ancestry. There is also a trend towards an increase in the performance of the CAN-S SNP set at decreasing European ancestry: in individuals with <20% European ancestry CAN-S has an accuracy that is nearly twice that observed for HIrisPlex-S. Although CAN-S tends to outperform HIrisPlex-S in most comparisons, it is only for MI that such a large difference in performance was observed. In summary, across all pigmentation traits we observe a gain in prediction accuracy for the Native American/European ancestry bins showing greater phenotypic diversity, with CAN SNPs markedly outperforming HIrisPlex-S only for MI in individuals with low (<20%) European ancestry.

Figure 4. Prediction accuracy for individuals with varying Native American/European admixture. Prediction was assessed in individual bins varying ~20% in admixture (bottom axis; for eye and hair pigmentation <20 individuals with >80% Native Ancestry were available in each trait category, thus preventing estimation of prediction accuracy). Coloured bars indicate accuracy obtained with Random Forest models using non-genetic + pigmentation SNP sets as predictors (R^2 being used as prediction measure for quantitative traits: Hue, Luminance, Chroma and Melanin index). Blue bars indicate the CAN-E/H/S SNP sets. Yellow bars indicate the HIrisPlex-S set. The standard deviation of the quantitative traits in each ancestry bin is indicated as a red line. For the categorical traits (Eye and Hair Colour), accuracy (proportion of correctly classified individuals) is used as the metric for prediction measures. Accuracy obtained without genetic predictors using a guessing strategy is indicated with a horizontal blue line for Proportional Strategy (random guessing) and magenta line for maxP (deterministic guessing). Detailed numerical values are given in Supplementary Table S4 and S5.



Prediction accuracy in CANDELA relative to other population samples

Table 1 compares published HIrisPlex-S-Online prediction accuracy estimates with those we obtained here with the CAN and HIrisPlex-S SNP sets using our implementation of MLR models trained in the CANDELA data and three-level colour categories for eyes, hair and skin (as described in Material and Methods). Prediction accuracy estimates for eye colour have been reported for HIrisPlex-S-Online in a Latin American sample (including 99 individuals from Venezuela and Brazil) and in a European sample[23,27,28]. The light (blue/grey) and dark (Brown/Black) colour categories have similar prediction accuracies across studies (~90-93%), except for the prediction of light eye colour reported for HIrisPlex-S-Online in the Venezuelan/Brazilian sample, where accuracy is lower (85%). The main difference in eye-colour prediction across studies lies in the intermediate category. No intermediate eye colours were predicted by HIrisPlex-S-Online in the Venezuelan/Brazilian sample. Our predictions **for the intermediate category** in the CANDELA sample have higher accuracy than **that** reported **for this category** for HIrisPlex-S-Online in Europeans, both when the prediction model was trained in the CANDELA data

or in the reference HIrisPlex-S data (respectively 89% and 85%, versus 73% in Europeans). Prediction accuracy, in the CANDELA sample, of the CAN-E and HIrisPlex-S SNP sets was identical.

Estimates for hair colour prediction accuracy using HIrisPlex-S-Online have been reported for a European sample[39]. Prediction accuracy estimates obtained here for the CANDELA sample are higher than reported in Europeans for all hair colours, except in the case of intermediate hair-colour (i.e. brown) predicted with HIrisPlex-S-Online. The highest hair-colour prediction accuracy was consistently obtained with the CAN-H SNP set, although the difference relative to HIrisPlex-S trained in the CANDELA data is marginal.

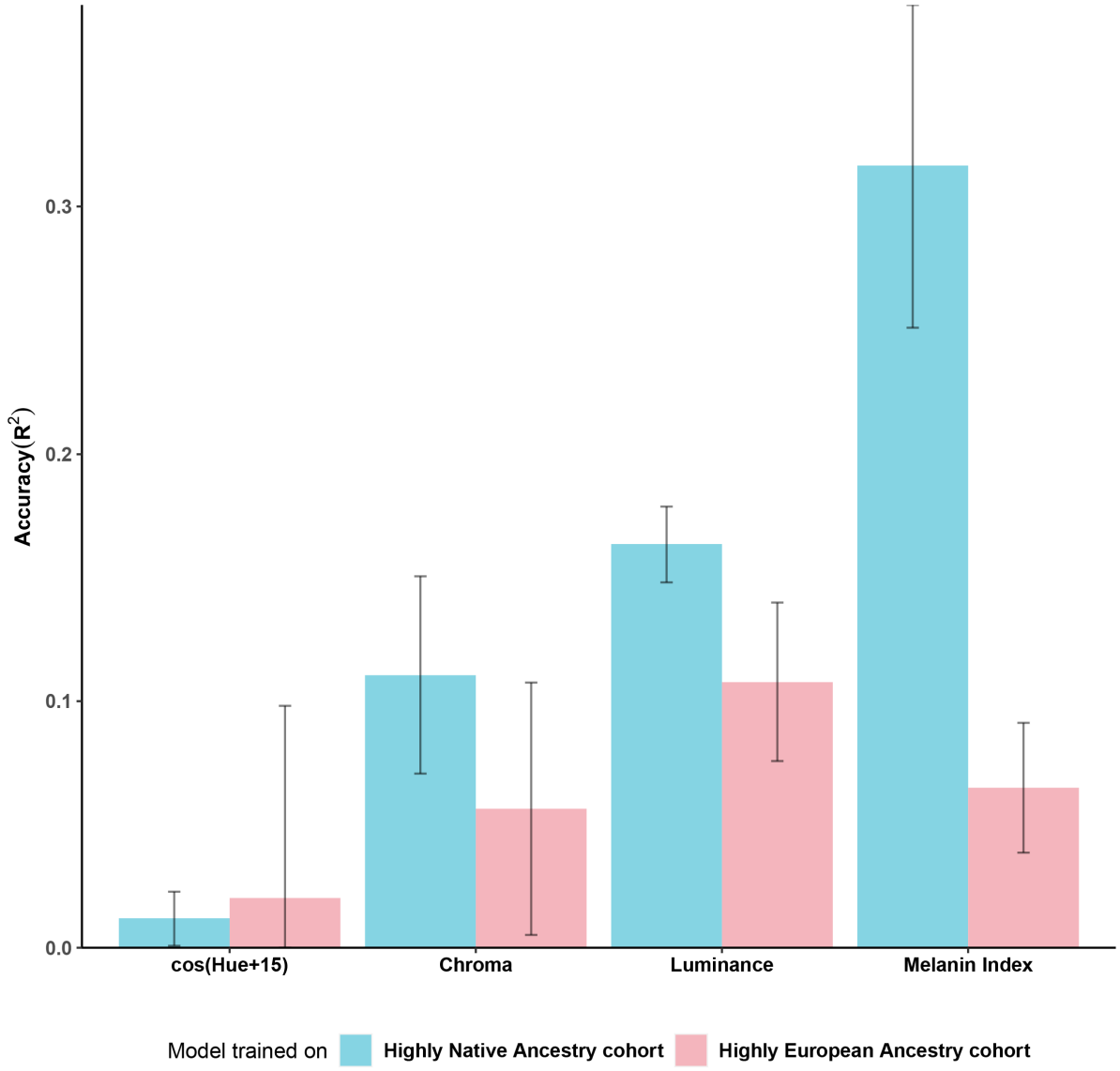
Concerning skin colour, we observe that predictions from HIrisPlex-S-Online have markedly lower accuracy in the CANDELA sample than reported for a world-wide sample. Model training in the CANDELA dataset increases prediction accuracy substantially, both for HIrisPlex-S and CAN-S (with CAN-S SNP set marginally outperforming HIrisPlex-S), although the accuracy values obtained for both sets are still below those reported for HIrisPlex-S-Online.

Portability of models for pigmentation prediction in individuals with high Native Ancestry

Considering the impact of training datasets in the performance of HIrisPlex-S (Table 1), we specifically examined the portability of RF models developed in two training datasets with extreme differences in ancestry (extracted from the CANDELA sample – see Supplementary figure S3): (i) a highly European training dataset (European ancestry >= 80% and Native American ancestry < 20%) and (ii) a highly Native training dataset (European ancestry < 20% and Native American ancestry >= 80%). We examined the performance of the resulting prediction models in a subset of the highly Native test dataset (Figure 5) in a

cross-validation scheme. We observe that models developed in the highly Native training dataset have a better performance than those developed in the highly European training dataset for Chrome, Luminance and MI. The most striking difference in performance is seen for MI, where the model trained with highly Native data has a prediction accuracy ~6 times that of the model trained in highly European data (Figure 5). Hue is the only trait for which the model trained in the highly European dataset **very slightly** outperforms the model trained in the highly Native dataset, but prediction accuracy in this case is extremely low (<2%), **and the confidence intervals substantially overlap**.

Figure 5. Portability of prediction models trained in highly European/Native American cohorts. For each continuous trait (cos(H+15), C, L and Melanin Index) we compare prediction accuracy on the same test data (a highly Native ancestry cohort). The prediction models were trained either on a highly Native (Blue) or highly European (Pink) cohort established from the CANDELA sample. For testing, we created equally-sized 4-folds for each pool of individuals. We built the RF models using three of the folds and evaluated prediction accuracy in the left-out fold from the Native ancestry cohort (see Supplementary figure S3).

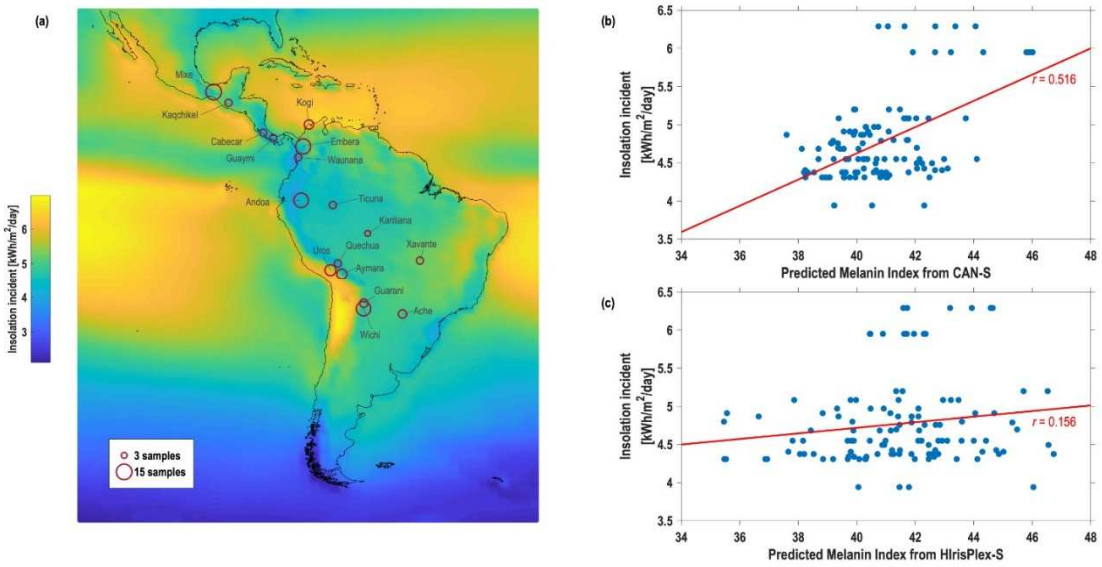


Prediction of skin pigmentation in Native Americans

We examined prediction performance of the CAN-S and HIrisPlex-S sets in a highly Native dataset independent of CANDELA by predicting MI in 117 individuals from 17 Native American populations[35]. As above, we trained RF prediction models using CANDELA individuals with >= 80% native ancestry. Since performance could not be measured directly in this dataset (due to the lack of phenotypic data), we examined the correlation of predicted skin pigmentation (MI) with solar radiation levels at the site of population sampling (Figure 6). Previous surveys of skin pigmentation in native populations

from across the world have found a correlation between skin pigmentation and solar radiation[48], an observation that has been interpreted as the result of selection throughout human evolution. In the Native Americans examined here we obtained correlations of 0.516 (p-value 2 x 10^{-9}) and 0.156 (p-value 0.1) with the CAN-S and HIrisPlex-S SNP sets, respectively (Figure 6).

Figure 6. Solar radiation levels and skin pigmentation (MI) predicted with CAN-S and HIrisplex-S SNP sets. (a) Annual average of insolation incident on a horizontal surface (kWh/m^2/day - data from NASA Surface meteorology and Solar Energy, 2008) and location of the Native American population sampled. The predicted MI and solar radiation levels at the sampling site for 117 individuals from 17 Native American populations is shown in **(b)** CAN-S and **(c)** HIrisPlex-S.



Discussion

Our comparison of different methods for predicting pigmentation traits agree with previous studies[23,42] **in finding that regression and RF generally outperform other approaches. Owing to its tree-based structure, RF implicitly models underlying**

interaction between SNPs. Since it has been shown that epistasis (SNP-SNP interactions) occurs to variable degrees for pigmentation traits [4,5,49], it is possible that prediction accuracy could be further increased by specifically allowing for interaction between SNPs. The difference in accuracy between tree-based methods relative to additive linear/logistic models could also shed light on the genetic architecture of these trait. For instance, it is tempting to hypothesize that SNP interactions may have a more substantial impact on the genetic architecture of Hue and Chroma [12,50], especially considering that these traits are less linear in nature (e.g. Hue is an angle, i.e. a circular trait)[49]. Furthermore, differences in accuracy are larger for linear compared to tree-based models (up to +0.19 gain in R^2 – see Supplementary table S3) than they are for HIrisPlex-S compared to CAN-E (up to +0.014 gain in R^2), whereas these two SNP sets only have 5 SNPs in common. We might therefore expect that relevant interactions would be limited to a handful of SNPs, consistent with proposals of significant interactions only between major pigmentation SNPs[4,5,12]. By contrast, skin pigmentation has been proposed to have greater additive polygenicity[4,6]. Consistent with this, linear models outperform tree-based methods for skin pigmentation prediction and the number of SNPs used in the model improves accuracy (Figure 3).

Previous studies[12,32,51,52] have shown that genetic ancestry correlates with pigmentation, probably as a result of the variable frequency of various pigmentation-associated alleles across populations. Here we observe that continental genetic ancestry has considerable predictive power (Figure 3, Supplementary table S3) and that the SNPs used for pigmentation prediction rank among the most correlated to continental ancestry components (see Supplementary table S2). Inclusion in the prediction models of SNPs selected based on pigmentation GWAS results further increases predictive

power, especially so for quantitative traits. We find that, for MI and the eye colour measurements, pigmentation-associated SNPs add, on an average, twice the prediction accuracy of that brought in by the genetic ancestry. The increase in prediction accuracy provided by pigmentation SNPs is less pronounced for categorical traits, but still, inclusion of these SNPs in the models reduces the proportion of incorrect classifications by a larger amount than does genetic ancestry, especially for eye colour.

The differences in predictive power that we observe between categorical and quantitative pigmentation traits is partly the result of the intrinsically lower statistical informativeness of discrete relative to continuous variables. This is accentuated here by the fact that categorical eye and hair colour have a highly skewed distribution in the CANDELA sample: 82% of individuals in this sample are assigned to the darkest category for eye and 85% for hair colour. The highly skewed distribution of these traits in the CANDELA sample results from lightly pigmented eyes and hair being essentially Western Eurasian traits [4–9]. That is, the occurrence of lightly pigmented hair and eyes in the CANDELA sample reflects the partly European ancestry of Latin Americans. This is consistent with the HIrisPlex-S SNP set (built using a world-wide sample, though consisting of 80% Europeans) and the CAN (E and H) SNP sets performing about equally for the prediction of eye and hair colour (despite the CAN sets including a much larger number of SNPs), and matching what has been reported in the literature [4,46,53].

Contrasting with eye and hair colour, differences in the prediction accuracy of skin colour appear to also influenced by the different genetic architecture of skin pigmentation, reflecting the world-wide variation that is observed for this trait[54]**. Our analysis of prediction along a gradient of Native American-European ancestry shows the highest gain in accuracy for the ancestry bins with the greatest phenotypic diversity:**

the highest European bin for eye/hair colour and the highest Native American bin for skin colour. In the bin with lowest European ancestry, there is hardly any gain in prediction accuracy for eye/hair colour over the deterministic maxP strategy, as almost all individuals in that bin are in the highly pigmented category. By contrast, this bin has the highest variation in skin pigmentation, and also shows the highest accuracy for the CAN-S (strongly outperforming HIrisPlex-S in this ancestry bin, Figure 4). Although this could be partly the result of model training, Figure 3 shows that the difference in performance is larger for MI than for eye and hair colour. A similar trend is observed in Table 1: there is a greater difference between the two HIrisPlex-S results for skin colour than for eye and hair colour. These observations point to skin pigmentation prediction in Latin Americans being impacted by the genetic architecture of this trait in non-European populations. Our finding of a stronger correlation of predicted MI with solar radiation levels in Native Americans for the CAN-S set, relative to HIrisPlex-S, is also consistent with this, and with literature reporting comparatively poor portability of European-based skin pigmentation prediction models in non-European populations[4,46,53].

Although quantitative variables are intrinsically more informative than categorical ones, forensic applications are mostly interested in the prediction of discrete categories, often just two (e.g. blue v. non-blue eyes) or three (light, intermediate or dark pigmentation). In that setting we find that for the CANDELA dataset there is remarkably little difference in performance between HIrisPlex-S and the much larger CAN SNP sets, particularly for eye and hair colour. As discussed above, this is likely to relate to the lower informativity of categorical traits, to which several other factors could be contributing, such as HIrisPlex-S capturing particularly well the genetic architecture of eye and hair colour, and having been optimized for the prediction of

categorical traits. However, our analyses indicate that use of the online implementation of HIrisPlex-S for the prediction of pigmentation traits in Latin Americans should be performed with caution. This likely relates mostly to the reference population set used for training of the prediction model implemented in the online HIrisPlex-S tool, not being sufficiently representative of the diversity of Latin-American populations.

In conclusion, our analyses of the large CANDELA pigmentation data underline the impact on prediction accuracy of greater polygenicity of skin, compared to eye and hair pigmentation. As expected from statistical principles, the effect of this greater polygenicity on prediction accuracy is more manifest for quantitative (MI) than for categorical skin colour. Prediction methods more sensitive to this polygenicity (e.g. regression) can therefore improve their accuracy for MI, through the inclusion of additional genetic predictors. This could be of considerable interest for certain evolutionary studies, as shown in Figure 6[55]**. However, for forensic applications, in which predictions of interest mostly relate to discrete categories, increasing the number of genetic predictors does not appear to provide much benefit, even for skin colour. The HIrisPlex-S SNP set, optimised precisely for this type of application, already provides excellent prediction accuracy. However, use of HIrisPlex-S for forensic studies in Latin America should carefully consider training of prediction models in reference datasets that more closely match the genetic diversity of the region.**

References

[1] T.E. King, G.G. Fortes, P. Balaresque, M.G. Thomas, D. Balding, P.M. Delser, R. Neumann, W. Parson, M. Knapp, S. Walsh, L. Tonasso, J. Holt, M. Kayser, J. Appleby, P. Forster, D. Ekserdjian, M. Hofreiter, K. Schürer, Identification of the remains of King Richard III, Nat Commun. 5 (2014) 5631. https://doi.org/10.1038/ncomms6631.

[2] I. Olalde, M.E. Allentoft, F. Sánchez-Quinto, G. Santpere, C.W.K. Chiang, M. DeGiorgio, J. Prado-Martinez, J.A. Rodríguez, S. Rasmussen, J. Quilez, O. Ramírez, U.M. Marigorta, M. Fernández-Callejo, M.E. Prada, J.M.V. Encinas, R. Nielsen, M.G. Netea, J. Novembre, R.A. Sturm, P. Sabeti, T. Marquès-Bonet, A. Navarro, E.

Willerslev, C. Lalueza-Fox, Derived immune and ancestral pigmentation alleles in a 7,000-year-old Mesolithic European, Nature. 507 (2014) 225–228. https://doi.org/10.1038/nature12960.
[3] C. Lalueza-Fox, H. Rompler, D. Caramelli, C. Staubert, G. Catalano, D. Hughes, N. Rohland, E. Pilli, L. Longo, S. Condemi, M. de la Rasilla, J. Fortea, A. Rosas, M. Stoneking, T. Schoneberg, J. Bertranpetit, M. Hofreiter, A Melanocortin 1 Receptor Allele Suggests Varying Pigmentation Among Neanderthals, Science. 318 (2007) 1453–1455. https://doi.org/10.1126/science.1147417.
[4] E. Pośpiech, A. Wojas-Pelc, S. Walsh, F. Liu, H. Maeda, T. Ishikawa, M. Skowron, M. Kayser, W. Branicki, The common occurrence of epistasis in the determination of human pigmentation and its impact on DNA-based pigmentation phenotype prediction, Forensic Science International: Genetics. 11 (2014) 64–72. https://doi.org/10.1016/j.fsigen.2014.01.012.
[5] E.E. Quillen, H.L. Norton, E.J. Parra, F. Lona-Durazo, K.C. Ang, F.M. Illiescu, L.N. Pearson, M.D. Shriver, T. Lasisi, O. Gokcumen, I. Starr, Y. Lin, A.R. Martin, N.G. Jablonski, Shades of complexity: New perspectives on the evolution and genetic architecture of human skin, Am J Phys Anthropol. 168 (2019) 4–26. https://doi.org/10.1002/ajpa.23737.
[6] A.R. Martin, M. Lin, J.M. Granka, J.W. Myrick, X. Liu, A. Sockell, E.G. Atkinson, C.J. Werely, M. Möller, M.S. Sandhu, D.M. Kingsley, E.G. Hoal, X. Liu, M.J. Daly, M.W. Feldman, C.R. Gignoux, C.D. Bustamante, B.M. Henn, An Unexpectedly Complex Architecture for Skin Pigmentation in Africans, Cell. 171 (2017) 1340-1353.e14. https://doi.org/10.1016/j.cell.2017.11.015.
[7] L. Deng, S. Xu, Adaptation of human skin color in various populations, Hereditas. 155 (2018) 1. https://doi.org/10.1186/s41065-017-0036-2.
[8] H.L. Norton, M. Edwards, S. Krithika, M. Johnson, E.A. Werren, E.J. Parra, Quantitative assessment of skin, hair, and iris variation in a diverse sample of individuals and associated genetic variation: QUANTITATIVE ASSESSMENT OF PIGMENTARY PHENOTYPE, Am. J. Phys. Anthropol. 160 (2016) 570–581. https://doi.org/10.1002/ajpa.22861.
[9] M. Jonnalagadda, M.A. Faizan, S. Ozarkar, R. Ashma, S. Kulkarni, H.L. Norton, E. Parra, A Genome-Wide Association Study of Skin and Iris Pigmentation among Individuals of South Asian Ancestry, Genome Biology and Evolution. 11 (2019) 1066–1076. https://doi.org/10.1093/gbe/evz057.
[10] M. Edwards, D. Cha, S. Krithika, M. Johnson, G. Cook, E.J. Parra, 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 (2016) 141–162. https://doi.org/10.1111/pcmr.12435.
[11] L. Rawofi, M. Edwards, S. Krithika, P. Le, D. Cha, Z. Yang, Y. Ma, J. Wang, B. Su, L. Jin, H.L. Norton, E.J. Parra, Genome-wide association study of pigmentary traits (skin and iris color) in individuals of East Asian ancestry, PeerJ. 5 (2017) e3951. https://doi.org/10.7717/peerj.3951.
[12] K. Adhikari, J. Mendoza-Revilla, A. Sohail, M. Fuentes-Guajardo, J. Lampert, J.C. Chacón-Duque, M. Hurtado, V. Villegas, V. Granja, V. Acuña-Alonzo, C. Jaramillo, W. Arias, R.B. Lozano, P. Everardo, J. Gómez-Valdés, H. Villamil-Ramírez, C.C. Silva de Cerqueira, T. Hunemeier, V. Ramallo, L. Schuler-Faccini, F.M. Salzano, R. Gonzalez-José, M.-C. Bortolini, S. Canizales-Quinteros, C. Gallo, G. Poletti, G. Bedoya, F. Rothhammer, D.J. Tobin, M. Fumagalli, D. Balding, A. Ruiz-Linares, A GWAS in Latin Americans highlights the convergent evolution of lighter skin pigmentation in Eurasia, Nat Commun. 10 (2019) 358. https://doi.org/10.1038/s41467-018-08147-0.

[13] M. Kayser, F. Liu, A.C.J.W. Janssens, F. Rivadeneira, O. Lao, K. van Duijn, M. Vermeulen, P. Arp, M.M. Jhamai, W.F.J. van IJcken, J.T. den Dunnen, S. Heath, D. Zelenika, D.D.G. Despriet, C.C.W. Klaver, J.R. Vingerling, P.T.V.M. de Jong, A. Hofman, Y.S. Aulchenko, A.G. Uitterlinden, B.A. Oostra, C.M. van Duijn, Three Genome-wide Association Studies and a Linkage Analysis Identify HERC2 as a Human Iris Color Gene, The American Journal of Human Genetics. 82 (2008) 411–423. https://doi.org/10.1016/j.ajhg.2007.10.003.

[14] P. Sulem, D.F. Gudbjartsson, S.N. Stacey, A. Helgason, T. Rafnar, K.P. Magnusson, A. Manolescu, A. Karason, A. Palsson, G. Thorleifsson, M. Jakobsdottir, S. Steinberg, S. Pálsson, F. Jonasson, B. Sigurgeirsson, K. Thorisdottir, R. Ragnarsson, K.R. Benediktsdottir, K.K. Aben, L.A. Kiemeney, J.H. Olafsson, J. Gulcher, A. Kong, U. Thorsteinsdottir, K. Stefansson, Genetic determinants of hair, eye and skin pigmentation in Europeans, Nat Genet. 39 (2007) 1443–1452. https://doi.org/10.1038/ng.2007.13.

[15] P. Sulem, D.F. Gudbjartsson, S.N. Stacey, A. Helgason, T. Rafnar, M. Jakobsdottir, S. Steinberg, S.A. Gudjonsson, A. Palsson, G. Thorleifsson, S. Pálsson, B. Sigurgeirsson, K. Thorisdottir, R. Ragnarsson, K.R. Benediktsdottir, K.K. Aben, S.H. Vermeulen, A.M. Goldstein, M.A. Tucker, L.A. Kiemeney, J.H. Olafsson, J. Gulcher, A. Kong, U. Thorsteinsdottir, K. Stefansson, Two newly identified genetic determinants of pigmentation in Europeans, Nat Genet. 40 (2008) 835–837. https://doi.org/10.1038/ng.160.

[16] S.I. Candille, D.M. Absher, S. Beleza, M. Bauchet, B. McEvoy, N.A. Garrison, J.Z. Li, R.M. Myers, G.S. Barsh, H. Tang, M.D. Shriver, Genome-Wide Association Studies of Quantitatively Measured Skin, Hair, and Eye Pigmentation in Four European Populations, PLoS ONE. 7 (2012) e48294. https://doi.org/10.1371/journal.pone.0048294.

[17] F. Liu, M. Visser, D.L. Duffy, P.G. Hysi, L.C. Jacobs, O. Lao, K. Zhong, S. Walsh, L. Chaitanya, A. Wollstein, G. Zhu, G.W. Montgomery, A.K. Henders, M. Mangino, D. Glass, V. Bataille, R.A. Sturm, F. Rivadeneira, A. Hofman, W.F.J. van IJcken, A.G. Uitterlinden, R.-J.T.S. Palstra, T.D. Spector, N.G. Martin, T.E.C. Nijsten, M. Kayser, Genetics of skin color variation in Europeans: genome-wide association studies with functional follow-up, Hum Genet. 134 (2015) 823–835. https://doi.org/10.1007/s00439-015-1559-0.

[18] The International Visible Trait Genetics Consortium, P.G. Hysi, A.M. Valdes, F. Liu, N.A. Furlotte, D.M. Evans, V. Bataille, A. Visconti, G. Hemani, G. McMahon, S.M. Ring, G.D. Smith, D.L. Duffy, G. Zhu, S.D. Gordon, S.E. Medland, B.D. Lin, G. Willemsen, J. Jan Hottenga, D. Vuckovic, G. Girotto, I. Gandin, C. Sala, M.P. Concas, M. Brumat, P. Gasparini, D. Toniolo, M. Cocca, A. Robino, S. Yazar, A.W. Hewitt, Y. Chen, C. Zeng, A.G. Uitterlinden, M.A. Ikram, M.A. Hamer, C.M. van Duijn, T. Nijsten, D.A. Mackey, M. Falchi, D.I. Boomsma, N.G. Martin, D.A. Hinds, M. Kayser, T.D. Spector, Genome-wide association meta-analysis of individuals of European ancestry identifies new loci explaining a substantial fraction of hair color variation and heritability, Nat Genet. 50 (2018) 652–656. https://doi.org/10.1038/s41588-018-0100-5.

[19] N.G. Crawford, D.E. Kelly, M.E.B. Hansen, M.H. Beltrame, S. Fan, S.L. Bowman, E. Jewett, A. Ranciaro, S. Thompson, Y. Lo, S.P. Pfeifer, J.D. Jensen, M.C. Campbell, W. Beggs, F. Hormozdiari, S.W. Mpoloka, G.G. Mokone, T. Nyambo, D.W. Meskel, G. Belay, J. Haut, NISC Comparative Sequencing Program, H. Rothschild, L. Zon, Y. Zhou, M.A. Kovacs, M. Xu, T. Zhang, K. Bishop, J. Sinclair, C. Rivas, E. Elliot, J. Choi, S.A. Li, B. Hicks, S. Burgess, C. Abnet, D.E. Watkins-Chow, E. Oceana, Y.S. Song, E. Eskin, K.M. Brown, M.S. Marks, S.K. Loftus, W.J. Pavan, M. Yeager, S.

Chanock, S.A. Tishkoff, Loci associated with skin pigmentation identified in African populations, Science. 358 (2017) eaan8433. https://doi.org/10.1126/science.aan8433.
[20] 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 Science International: Genetics. 5 (2011) 170–180. https://doi.org/10.1016/j.fsigen.2010.02.004.
[21] S. Walsh, A. Wollstein, F. Liu, U. Chakravarthy, M. Rahu, J.H. Seland, G. Soubrane, L. Tomazzoli, 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 Science International: Genetics. 6 (2012) 330–340. https://doi.org/10.1016/j.fsigen.2011.07.009.
[22] J. Purps, M. Geppert, M. Nagy, L. Roewer, Evaluation of the IrisPlex eye colour prediction tool in a German population sample, Forensic Science International: Genetics Supplement Series. 3 (2011) e202–e203. https://doi.org/10.1016/j.fsigss.2011.08.101.
[23] 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, Current Biology. 19 (2009) R192–R193. https://doi.org/10.1016/j.cub.2009.01.027.
[24] 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 Science International: Genetics. 7 (2013) 98–115. https://doi.org/10.1016/j.fsigen.2012.07.005.
[25] 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 Science International: Genetics. 9 (2014) 150–161. https://doi.org/10.1016/j.fsigen.2013.12.006.
[26] 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.
[27] L. Chaitanya, K. Breslin, S. Zuñiga, 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 Science International: Genetics. 35 (2018) 123–135. https://doi.org/10.1016/j.fsigen.2018.04.004.
[28] K. Breslin, B. Wills, A. Ralf, M. Ventayol Garcia, M. Kukla-Bartoszek, E. Pospiech, A. Freire-Aradas, C. Xavier, S. Ingold, M. de La Puente, K.J. van der Gaag, N. Herrick, C. Haas, W. Parson, C. Phillips, T. Sijen, W. Branicki, S. Walsh, M. Kayser, HIrisPlex-S system for eye, hair, and skin color prediction from DNA: Massively parallel sequencing solutions for two common forensically used platforms, Forensic Science International: Genetics. 43 (2019) 102152. https://doi.org/10.1016/j.fsigen.2019.102152.
[29] H.L. Norton, R.A. Kittles, E. Parra, P. McKeigue, X. Mao, K. Cheng, V.A. Canfield, D.G. Bradley, B. McEvoy, M.D. Shriver, Genetic Evidence for the Convergent Evolution of Light Skin in Europeans and East Asians, Molecular Biology and Evolution. 24 (2006) 710–722. https://doi.org/10.1093/molbev/msl203.
[30] M. Edwards, A. Bigham, J. Tan, S. Li, A. Gozdzik, K. Ross, L. Jin, E.J. Parra, Association of the OCA2 Polymorphism His615Arg with Melanin Content in East Asian Populations: Further Evidence of Convergent Evolution of Skin Pigmentation, PLoS Genet. 6 (2010) e1000867. https://doi.org/10.1371/journal.pgen.1000867.

[31] K. Adhikari, J. Mendoza-Revilla, J.C. Chacón-Duque, M. Fuentes-Guajardo, A. Ruiz-Linares, Admixture in Latin America, Current Opinion in Genetics & Development. 41 (2016) 106–114. https://doi.org/10.1016/j.gde.2016.09.003.
[32] A. Ruiz-Linares, K. Adhikari, V. Acuña-Alonzo, M. Quinto-Sanchez, C. Jaramillo, W. Arias, M. Fuentes, M. Pizarro, P. Everardo, F. de Avila, J. Gómez-Valdés, P. León-Mimila, T. Hunemeier, V. Ramallo, C.C. Silva de Cerqueira, M.-W. Burley, E. Konca, M.Z. de Oliveira, M.R. Veronez, M. Rubio-Codina, O. Attanasio, S. Gibbon, N. Ray, C. Gallo, G. Poletti, J. Rosique, L. Schuler-Faccini, F.M. Salzano, M.-C. Bortolini, S. Canizales-Quinteros, F. Rothhammer, G. Bedoya, D. Balding, R. Gonzalez-José, Admixture in Latin America: Geographic Structure, Phenotypic Diversity and Self-Perception of Ancestry Based on 7,342 Individuals, PLoS Genet. 10 (2014) e1004572. https://doi.org/10.1371/journal.pgen.1004572.
[33] K. Adhikari, M. Fuentes-Guajardo, M. Quinto-Sánchez, J. Mendoza-Revilla, J. Camilo Chacón-Duque, V. Acuña-Alonzo, C. Jaramillo, W. Arias, R.B. Lozano, G.M. Pérez, J. Gómez-Valdés, H. Villamil-Ramírez, T. Hunemeier, V. Ramallo, C.C. Silva de Cerqueira, M. Hurtado, V. Villegas, V. Granja, C. Gallo, G. Poletti, L. Schuler-Faccini, F.M. Salzano, M.-C. Bortolini, S. Canizales-Quinteros, M. Cheeseman, J. Rosique, G. Bedoya, F. Rothhammer, D. Headon, R. González-José, D. Balding, A. Ruiz-Linares, A genome-wide association scan implicates DCHS2, RUNX2, GLI3, PAX1 and EDAR in human facial variation, Nat Commun. 7 (2016) 11616. https://doi.org/10.1038/ncomms11616.
[34] K. Adhikari, T. Fontanil, S. Cal, J. Mendoza-Revilla, M. Fuentes-Guajardo, J.-C. Chacón-Duque, F. Al-Saadi, J.A. Johansson, M. Quinto-Sanchez, V. Acuña-Alonzo, C. Jaramillo, W. Arias, R. Barquera Lozano, G. Macín Pérez, J. Gómez-Valdés, H. Villamil-Ramírez, T. Hunemeier, V. Ramallo, C.C. Silva de Cerqueira, M. Hurtado, V. Villegas, V. Granja, C. Gallo, G. Poletti, L. Schuler-Faccini, F.M. Salzano, M.-C. Bortolini, S. Canizales-Quinteros, F. Rothhammer, G. Bedoya, R. Gonzalez-José, D. Headon, C. López-Otín, D.J. Tobin, D. Balding, A. Ruiz-Linares, A genome-wide association scan in admixed Latin Americans identifies loci influencing facial and scalp hair features, Nat Commun. 7 (2016) 10815. https://doi.org/10.1038/ncomms10815.
[35] J.-C. Chacón-Duque, K. Adhikari, M. Fuentes-Guajardo, J. Mendoza-Revilla, V. Acuña-Alonzo, R. Barquera, M. Quinto-Sánchez, J. Gómez-Valdés, P. Everardo Martínez, H. Villamil-Ramírez, T. Hünemeier, V. Ramallo, C.C. Silva de Cerqueira, M. Hurtado, V. Villegas, V. Granja, M. Villena, R. Vásquez, E. Llop, J.R. Sandoval, A.A. Salazar-Granara, M.-L. Parolin, K. Sandoval, R.I. Peñaloza-Espinosa, H. Rangel-Villalobos, C.A. Winkler, W. Klitz, C. Bravi, J. Molina, D. Corach, R. Barrantes, V. Gomes, C. Resende, L. Gusmão, A. Amorim, Y. Xue, J.-M. Dugoujon, P. Moral, R. González-José, L. Schuler-Faccini, F.M. Salzano, M.-C. Bortolini, S. Canizales-Quinteros, G. Poletti, C. Gallo, G. Bedoya, F. Rothhammer, D. Balding, G. Hellenthal, A. Ruiz-Linares, Latin Americans show wide-spread Converso ancestry and imprint of local Native ancestry on physical appearance, Nat Commun. 9 (2018) 5388. https://doi.org/10.1038/s41467-018-07748-z.
[36] K. Adhikari, G. Reales, A.J.P. Smith, E. Konka, J. Palmen, M. Quinto-Sanchez, V. Acuña-Alonzo, C. Jaramillo, W. Arias, M. Fuentes, M. Pizarro, R. Barquera Lozano, G. Macín Pérez, J. Gómez-Valdés, H. Villamil-Ramírez, T. Hunemeier, V. Ramallo, C.C. Silva de Cerqueira, M. Hurtado, V. Villegas, V. Granja, C. Gallo, G. Poletti, L. Schuler-Faccini, F.M. Salzano, M.-C. Bortolini, S. Canizales-Quinteros, F. Rothhammer, G. Bedoya, R. Calderón, J. Rosique, M. Cheeseman, M.F. Bhutta, S.E. Humphries, R. Gonzalez-José, D. Headon, D. Balding, A. Ruiz-Linares, A genome-wide association

study identifies multiple loci for variation in human ear morphology, Nat Commun. 6 (2015) 7500. https://doi.org/10.1038/ncomms8500.
[37] O. Maroñas, C. Phillips, J. Söchtig, A. Gomez-Tato, R. Cruz, J. Alvarez-Dios, M.C. de Cal, Y. Ruiz, M. Fondevila, Á. Carracedo, M.V. Lareu, Development of a forensic skin colour predictive test, Forensic Science International: Genetics. 13 (2014) 34–44. https://doi.org/10.1016/j.fsigen.2014.06.017.
[38] A. Freire-Aradas, Y. Ruiz, C. Phillips, O. Maroñas, J. Söchtig, A.G. Tato, J.Á. Dios, M.C. de Cal, V.N. Silbiger, A.D. Luchessi, A.D. Luchessi, M.A. Chiurillo, Á. Carracedo, M.V. Lareu, Exploring iris colour prediction and ancestry inference in admixed populations of South America, Forensic Science International: Genetics. 13 (2014) 3–9. https://doi.org/10.1016/j.fsigen.2014.06.007.
[39] 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.
[40] T.K.M. Leite, R.M.C. Fonseca, N.M. de França, E.J. Parra, R.W. Pereira, Genomic Ancestry, Self-Reported "Color" and Quantitative Measures of Skin Pigmentation in Brazilian Admixed Siblings, PLoS ONE. 6 (2011) e27162. https://doi.org/10.1371/journal.pone.0027162.
[41] A. Lo, H. Chernoff, T. Zheng, S.-H. Lo, Why significant variables aren't automatically good predictors, Proc Natl Acad Sci USA. 112 (2015) 13892–13897. https://doi.org/10.1073/pnas.1518285112.
[42] K. Zaorska, P. Zawierucha, M. Nowicki, Prediction of skin color, tanning and freckling from DNA in Polish population: linear regression, random forest and neural network approaches, Hum Genet. 138 (2019) 635–647. https://doi.org/10.1007/s00439-019-02012-w.
[43] R.K. Valenzuela, M.S. Henderson, M.H. Walsh, N.A. Garrison, J.T. Kelch, O. Cohen-Barak, D.T. Erickson, F. John Meaney, J. Bruce Walsh, K.C. Cheng, S. Ito, K. Wakamatsu, T. Frudakis, M. Thomas, M.H. Brilliant, Predicting Phenotype from Genotype: Normal Pigmentation, Journal of Forensic Sciences. 55 (2010) 315–322. https://doi.org/10.1111/j.1556-4029.2009.01317.x.
[44] O. Spichenok, Z.M. Budimlija, A.A. Mitchell, A. Jenny, L. Kovacevic, D. Marjanovic, T. Caragine, M. Prinz, E. Wurmbach, Prediction of eye and skin color in diverse populations using seven SNPs, Forensic Science International: Genetics. 5 (2011) 472–478. https://doi.org/10.1016/j.fsigen.2010.10.005.
[45] B. Alipanahi, P. Fontanillas, 23andMe Resarch Team, S. Pitts, R. Gentleman, Pigmentor—Accurate prediction of multiple pigmentation phenotypes, (2017).
[46] L. Yun, Y. Gu, H. Rajeevan, K.K. Kidd, Application of six IrisPlex SNPs and comparison of two eye color prediction systems in diverse Eurasia populations, Int J Legal Med. 128 (2014) 447–453. https://doi.org/10.1007/s00414-013-0953-1.
[47] 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 Science International: Genetics. 7 (2013) 28–40. https://doi.org/10.1016/j.fsigen.2012.05.009.
[48] N.G. Jablonski, G. Chaplin, The evolution of human skin coloration, Journal of Human Evolution. 39 (2000) 57–106. https://doi.org/10.1006/jhev.2000.0403.
[49] E. Pośpiech, J. Draus-Barini, T. Kupiec, A. Wojas-Pelc, W. Branicki, Gene–gene interactions contribute to eye colour variation in humans, J Hum Genet. 56 (2011) 447–455. https://doi.org/10.1038/jhg.2011.38.
[50] A. Wollstein, S. Walsh, F. Liu, U. Chakravarthy, M. Rahu, J.H. Seland, G. Soubrane, L. Tomazzoli, F. Topouzis, J.R. Vingerling, J. Vioque, S. Böhringer, A.E. Fletcher, M.

Kayser, Novel quantitative pigmentation phenotyping enhances genetic association, epistasis, and prediction of human eye colour, Sci Rep. 7 (2017) 43359. https://doi.org/10.1038/srep43359.
[51] T.N. Frudakis, Molecular photofitting: predicting ancestry and phenotype using DNA, Elsevier/Academic Press, Amsterdam ; Boston, 2008.
[52] C.C.S. de Cerqueira, T. Hünemeier, J. Gomez-Valdés, V. Ramallo, C.D. Volasko-Krause, A.A.L. Barbosa, P. Vargas-Pinilla, R.C. Dornelles, D. Longo, F. Rothhammer, G. Bedoya, S. Canizales-Quinteros, V. Acuña-Alonzo, C. Gallo, G. Poletti, R. González-José, F.M. Salzano, S.M. Callegari-Jacques, L. Schuler-Faccini, A. Ruiz-Linares, M. Cátira Bortolini, for CANDELA (Consortium for the Analysis of the Diversity and Evolution of Latin America), Implications of the Admixture Process in Skin Color Molecular Assessment, PLoS ONE. 9 (2014) e96886. https://doi.org/10.1371/journal.pone.0096886.
[53] A. Pneuman, Z.M. Budimlija, T. Caragine, M. Prinz, E. Wurmbach, Verification of eye and skin color predictors in various populations, Legal Medicine. 14 (2012) 78–83. https://doi.org/10.1016/j.legalmed.2011.12.005.
[54] G. Hudjashov, R. Villems, T. Kivisild, Global Patterns of Diversity and Selection in Human Tyrosinase Gene, PLoS ONE. 8 (2013) e74307. https://doi.org/10.1371/journal.pone.0074307.
[55] D. Ju, I. Mathieson, The evolution of skin pigmentation-associated variation in West Eurasia, Proc Natl Acad Sci USA. 118 (2021) e2009227118. https://doi.org/10.1073/pnas.2009227118.

Data availability:. Summary statistics from the GWAS analyses on which the three CAN SNP sets used in this study were established the have been previously deposited at GWAS central (http://www.gwascentral.org/study/HGVST3308).

Tables

Table 1. Overall Accuracy (%) and trait AUC (%) for categorical eye (A), hair (B) and skin (C) colour obtained here and in other studies (% frequency of trait in sample is shown in parentheses).

(A)

Sample	CANDELA N=6,529				Venezuela/Brazil[b] N = 99		Europeans[d] N=2,364	
SNP set		CAN-E[a]	HIrisPlex-S[a]	HIrisPlex-S-Online				
Overall Accuracy		89	89	87				83
1. Blue/grey	(3)	93	93	93	(12)	85	(68)	91
2. Intermediate	(13)	89	89	85	(23)	NA	(10)	73
3. Brown/black	(85)	92	92	90	(64)	91	(23)	93

(B)

Sample	CANDELA N=6,495				Europeans[c] N = 385	
SNP set		CAN-H[a]	HIrisPlex-S[a]	HIrisPlex-S-Online		
Overall Accuracy		85	84	56		71
1.Red	(0)	NA	NA	NA	(25)	90
2. Blond	(3)	94	92	90	(54)	75
3. Brown	(16)	82	81	65	(9)	72
4. Black	(82)	87	85	80	(12)	78

(C)

Sample	CANDELA N=6,526				World-wide N=2,025[e]	
SNP set		CAN-S[a]	HIrisPlex-S[a]	HIrisPlex-S-Online		
Overall Accuracy		73	72	26		87
1. Fair	(38)	83	81	78	(92)	97
2. Intermediate	(59)	77	77	71	(3)	83
3. Dark	(3)	86	84	76	(5)	96

HIrisPlex-S-Online: https://hirisplex.erasmusmc.nl

[a] Including only SNPs as predictors and with MLR as prediction method (as in HIrisPlex-S-Online).

[b] Obtained with HIrisPlex as described in Freire-Aradas et al. 2014[38].

[c] Obtained with HIrisPlex-S as described in Branicki et al 2011[39], including individuals with red hair.

[d] Obtained from the specificity reported in Liu et al 2009[23]

[e] Prediction values reported in Walsh et al 2017[26]