Accepted Manuscript

Genome-wide association studies identify multiple genetic loci influencing eyebrow color variation in Europeans

Fuduan Peng, Gu Zhu, Pirro G. Hysi, Ryan J. Eller, Yan Chen, Yi Li, Merel A. Hamer, Changqing Zeng, Racquel L. Hopkins, Case L. Jacobus, Paige L. Wallace, André G. Uitterlinden, M. Arfan Ikram, Tamar Nijsten, David L. Duffy, Sarah E. Medland, Timothy D. Spector, Susan Walsh, Nicholas G. Martin, Fan Liu, Manfred Kayser, on behalf of the International Visible Trait Genetics Consortium

PII: S0022-202X(19)30037-5

DOI: https://doi.org/10.1016/j.jid.2018.12.029

Reference: JID 1765

To appear in: *The Journal of Investigative Dermatology*

Received Date: 3 September 2018

Revised Date: 29 November 2018

Accepted Date: 12 December 2018

Please cite this article as: Peng F, Zhu G, Hysi PG, Eller RJ, Chen Y, Li Y, Hamer MA, Zeng C, Hopkins RL, Jacobus CL, Wallace PL, Uitterlinden AG, Ikram MA, Nijsten T, Duffy DL, Medland SE, Spector TD, Walsh S, Martin NG, Liu F, Kayser M, on behalf of the International Visible Trait Genetics Consortium, Genome-wide association studies identify multiple genetic loci influencing eyebrow color variation in Europeans, *The Journal of Investigative Dermatology* (2019), doi: https://doi.org/10.1016/j.jid.2018.12.029.

Genome-wide association studies identify multiple genetic loci influencing eyebrow color variation in Europeans

Fuduan Peng[1,#], Gu Zhu[2,#], Pirro G. Hysi[3,#], Ryan J. Eller[4], Yan Chen[1], Yi Li[1], Merel A. Hamer[5], Changqing Zeng[1], Racquel L. Hopkins[4], Case L. Jacobus[4], Paige L. Wallace[4], André G. Uitterlinden[6,7], M. Arfan Ikram[7], Tamar Nijsten[5], David L. Duffy[2], Sarah E. Medland[2], Timothy D. Spector[3], Susan Walsh[4], Nicholas G. Martin[2], Fan Liu[1,8,*,$], Manfred Kayser[8,*,$] on behalf of the International Visible Trait Genetics Consortium

[1]CAS Key Laboratory of Genomic and Precision Medicine, Beijing Institute of Genomic, University of Chinese Academy of Sciences, Chinese Academy of Sciences, Beijing, P.R. China

[2]QIMR Berghofer Medical Research Institute, Brisbane, Australia

[3]Department of Twin Research and Genetic Epidemiology, King's College London, London, United Kingdom

[4]Department of Biology, Indiana University-Purdue University Indianapolis, Indianapolis, IN, U.S.A.

[5]Department of Dermatology, Erasmus MC University Medical Center Rotterdam, Rotterdam, The Netherlands

[6]Department of Internal Medicine, Erasmus MC University Medical Center Rotterdam, Rotterdam, The Netherlands

[7]Department of Epidemiology, Erasmus MC University Medical Center Rotterdam, Rotterdam, The Netherlands

[8]Department of Genetic Identification, Erasmus MC University Medical Center Rotterdam, Rotterdam, The Netherlands

[#,*] these authors, respectively, contributed equally to this work

[$] correspondence: Fan Liu (f.liu@erasmusmc.nl; liufan@big.ac.cn) or Manfred Kayser (m.kayser@erasmusmc.nl)

Running title: GWAS for eyebrow color

Keywords: eyebrows, eyebrow color; GWAS; human pigmentation, human appearance

TO THE EDITOR,

Eyebrow color shows a high degree of variation in Europeans. Although no heritability estimate has yet been reported, eyebrow color may share a large genetic component with scalp hair color, which has an estimated heritability of up to 90% (Lin et al., 2015). Although phenotypic relationship between eyebrow and scalp hair color clearly exists, such a correlation is not perfect, suggesting the existence of overlapping and unique genetic components for both traits. While previous genome-wide association studies (GWASs) on human eye (Kayser et al., 2008, Liu et al., 2010, Sulem et al., 2008, Sulem et al., 2007), scalp hair (Han et al., 2008, Hysi et al., 2018, Sulem et al., 2007), and skin color (Han et al., 2008, Liu et al., 2015, Sulem et al., 2007, Visconti et al., 2018) have identified multiple DNA variants, no GWAS for eyebrow color has been reported as of yet.

The cohorts related studies included in the current study have been approved by the Medical Ethics Committee of the Erasmus MC, the St. Thomas' Hospital Local Research Ethics Committee, the QIMR Berghofer Human Research Ethics Committee, and the Indiana University Internal Review Board, respectively; and all participants provided written informed consent under protocols reviewed by the corresponding institutions.

The discovery stage meta-analysis of three GWASs for eyebrow color included a total of 6,513 European individuals from three cohorts, the Rotterdam Study (RS, n=3,114, mean age 68.48±9.34, 53.6% female), the TwinsUK study (n=1,038, mean

age 59.47±9.50, all females), and the Queensland Institute of Medical Research study (QIMR, n=2,361, mean age 16.43±0.80, 54.0% female, **Table S1**). Eyebrow color was graded into four broad ordinal categories i.e., red, blond, brown, and black using photo-numeric scales (**Table S1**). Detailed phenotype evaluation is provided in **Table S2-S6** and in Supplementary Materials.

The discovery stage meta-analysis of three GWASs identified a total of 355 SNPs at 6 distinct genetic loci showing genome-wide significant association with eyebrow color ($p<5\times10^{-8}$, **Figure 1 & S1, Table S7**). Among these 6 loci, one locus (10q22.2: *C10orf11*) had not been previously associated with any other human pigmentation trait; the top-associated SNP (rs11001536, β=-0.21, $p=3.16\times10^{-8}$, **Table S7**) is an intronic DNA variant in *C10orf11* (**Figure S2**). The remaining 5 loci have been repeatedly reported with genome-wide significant association with human eye, scalp hair, and/or skin color. These include 15q13.1 (rs7494942, β=0.22, $p=6.36\times10^{-58}$ for *HERC2* and rs4778237, β=0.17, $p=7.01\times10^{-31}$ for *OCA2*), 16q24.3 (*MC1R* rs75570604, β=-0.25, $p=9.88\times10^{-52}$), 14q32.12 (*SLC24A4* rs12883151, β=0.10, $p=4.44\times10^{-27}$), 20q11.22 (*ASIP* rs6059655, β=-0.11, $p=4.60\times10^{-10}$) and 5p13.2 (*SLC45A2* rs16891982, β=0.18, $p=2.60\times10^{-8}$) (**Table S7**).

The replication was conducted in 2,054 individuals of European origin from an additional United States cohort (US, mean age 25.75±11.39, 68% female, **Table S1**). Five loci highlighted in our discovery GWAS meta-analysis (*SCL45A2*, *SCL24A4*, *HERC2* & *OCA2*, *MC1R*, and *ASIP*) have been successfully replicated in US ($p<0.05$,

Table S7), and showed consistent allele effects in all four cohorts (**Figure S3**). The significant association for rs11001536 in *C10orf11* highlighted in the discovery GWAS was not replicated in US. This SNP was nominally significant in RS (β=-0.23, p=2.16×10^{-6}) and in TwinsUK (β=-0.27, p=4.97×10^{-3}), but non-significant in QIMR (β=-0.09, p=0.26) and in US (β=0.07, p=0.60) (**Table S7**). Notably, both cohorts that showed significant association consist of elderlies (RS & TwinsUK), while the two datasets not showing significant association were obtained from adolescents (QIMR & US). This suggests that the eyebrow color effect of *C10orf11* may be age-dependent, which warrants further investigation in future studies. The light eyebrow color associated G-allele had a relatively low frequency in Europeans (f=0.02) but more frequent in Asians with darker eyebrows (**Figure S4**), potentially explained by different LD structures between these populations. Previous studies suggest that *C10orf11* is an effective melanocyte-differentiation gene, which is known to cause the Oculocutaneous Albinism (OCA) 7 phenotype via a rare nonsense mutation c.580C>T (p.Arg194*, rs587776952) (Gronskov et al., 2013). An additional genome-wide meta-analysis in all four cohorts did not reveal any additional genome-wide significant loci (**Figure S5**).

Among the 138 SNPs from a recently published GWAS meta-analysis on scalp hair color involving almost 300,000 Europeans (Hysi et al., 2018), 7 SNPs showed significant association after the correction for multiple testing (adjusted p<4.67×10^{-4}), including 15q13.1 *HERC2* rs12913832 (p=1.12×10^{-48}), 16q24.3 *MC1R* rs1805007

(p=1.25×10^{-47}), 14q32.12 *SLC24A4* rs17184180 (p=1.67×10^{-26}), 20q11.22 *ASIP* rs6059655 (p=4.60×10^{-10}), 5p13.2 *SLC45A2* rs16891982 (p=2.60×10^{-8}), 6p25.3 *IRF4* rs12203592 (p=3.47×10^{-6}), and 1q32.1 *DSTYK* rs2369633 (p=5.03×10^{-5}) (**Table S8**). The first 6 SNPs all have known effect on human pigmentation traits. The last SNP was only recently identified in association with hair color (Hysi et al., 2018).

In a subset of RS (n=1,656), we compared the 8 top-associated SNPs at the 7 genetic loci that were highlighted with significant eyebrow color association in our GWAS and candidate gene study (**Table S9**). In general, the contributions of the 8 SNPs to scalp hair color variation were slightly larger than their impact on eyebrow color variation (**Figure S6**). *MC1R* rs1805007 showed much larger contribution to scalp hair color than eyebrow color, likely explained by an aging effect on red eyebrow color since the scalp hair color information in RS was ascertained from questionnaire on "hair color when young". *SLC45A2* rs16891982 was the only DNA variant of the 8 tested that showed slightly larger contribution to eyebrow color than to scalp hair color. These results suggest that the prediction accuracy for eyebrow color should be at a similar level to that for hair color in the same sample set under equal phenotype accuracy.

An eyebrow color prediction model was trained in 3,114 RS subjects and validated in 779 independent RS subjects not included in the GWAS. Red eyebrow was excluded from the prediction analysis since none of these individuals had the phenotype. A model including 25 SNPs achieved prediction accuracies expressed as

AUC at 0.701 (95% CI: 0.621-0.781) for blond, 0.620 (95% CI: 0.576-0.658) for brown, and 0.674 (95% CI: 0.633-0.709) for black eyebrows (**Figure 2a & 2b & S7, Table S10 & S11**). The AUC values reported here for eyebrow color are lower compared to those previously reported for scalp hair color with the 22-SNP HIrisPlex model, which ranged between 0.75 and 0.92 for the four scalp hair color categories used (Walsh et al., 2014). This discrepancy can be explained by our dataset lacking four important rare *MC1R* SNPs (which are used in the HIrisPlex model), and an aging effect that decreases phenotype quality, particularly for red color. With the mid-ranged accuracy level, our 25-SNP model provided highly confident prediction results for ~7% of the validation set, e.g. those with high (>0.80) or low (<0.20) prediction probabilities of certain eyebrow color type (**Figure 2c**). Applying this model to 2,504 subjects from the 1000 Genomes Project revealed that prediction outcomes were generally consistent with knowledge about the global distribution of eyebrow color variation (**Figure 2d & S8**). More detailed prediction results are provided in the **Supplementary Materials**.

In conclusion, this first eyebrow color GWAS in Europeans highlighted 6 genome-wide significant genetic loci harboring 6 well-known pigmentation genes including *ASIP*, *HERC2*, *MC1R*, *OCA2*, *SLC24A4*, *SLC45A2*, and a gene to our knowledge previously unreported, *C10orf11*.The finding at *C10orf11* warrants further investigations in European subjects with different age distributions. A candidate gene study suggested the involvement of two additional known pigmentation genes *DSTYK*

and *IRF4* in human eyebrow color. This DNA-based eyebrow color prediction model is useful in future forensic applications.

Conflict of interests

The authors declare that they have no competing interests.

Acknowledgements

This work has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 740580 (VISAGE Project). FL is additionally supported by National Key R&D Program of China (2017YFC0803501), National Natural Science Foundation of China (91651507) and "The Thousand Talents Plan for Young Professionals". MK is additionally supported by Erasmus MC University Medical Center Rotterdam.

The generation of the GWAS datasets of the Rotterdam Study was supported by the Netherlands Organisation of Scientific Research NOW Investments (nr. 175.010.2005.011, 911-03-012), the Genetic Laboratory of the Department of Internal Medicine, Erasmus MC, the Research Institute for Diseases in the Elderly (014-93-015; RIDE2), the Netherlands Genomics Initiative (NGI)/Netherlands Organisation for Scientific Research (NWO) Netherlands Consortium for Healthy Aging (NCHA), project nr. 050-060-810. The Rotterdam Study is funded by Erasmus Medical Center and Erasmus University, Rotterdam, Netherlands Organization for the

Health Research and Development (ZonMw), the Research Institute for Diseases in the Elderly (RIDE), the Ministry of Education, Culture and Science, the Ministry for Health, Welfare and Sports, the European Commission (DG XII), and the Municipality of Rotterdam.

The TwinsUK study is funded by the Wellcome Trust, Medical Research Council, European Union (FP7/2007-2013), the National Institute for Health Research (NIHR)-funded BioResource, Clinical Research Facility and Biomedical Research Centre based at Guy's and St Thomas' NHS Foundation Trust in partnership with King's College London.

QIMR gratefully acknowledges the participants and their families. We would like to thank Kerrie McAloney for data collection and Scott Gordon for data management. We acknowledge funding by Australian National Health and Medical Research Council (NHMRC) grants 241944, 339462, 389927, 389875, 389891, 389892, 389938, 442915, 442981, 496739, 552485, 552498 and Australian Research Council grants A7960034, A79906588, A79801419, DP0770096, DP0212016, DP0343921 for building and maintaining the adolescent twin family resource through which samples were collected. SEM is supported by an NHMRC fellowship APP1103623.

All information and work pertaining to the US cohort was generated under funding by the United States National Institute of Justice (US-NIJ); grant number 2014-DN-BX-K031. We thank NIJ and all individuals who participated in this collection.

Reference

Gronskov K, Dooley CM, Ostergaard E, Kelsh RN, Hansen L, Levesque MP, et al. Mutations in c10orf11, a melanocyte-differentiation gene, cause autosomal-recessive albinism. American journal of human genetics 2013;92(3):415-21.

Han J, Kraft P, Nan H, Guo Q, Chen C, Qureshi A, et al. A genome-wide association study identifies novel alleles associated with hair color and skin pigmentation. PLoS genetics 2008;4(5):e1000074.

Hysi PG, Valdes AM, Liu F, Furlotte NA, Evans DM, Bataille V, et al. Genome-wide association meta-analysis of individuals of European ancestry identifies new loci explaining a substantial fraction of hair color variation and heritability. Nature genetics 2018;50(5):652-6.

Kayser M, Liu F, Janssens AC, Rivadeneira F, Lao O, van Duijn K, et al. Three genome-wide association studies and a linkage analysis identify HERC2 as a human iris color gene. American journal of human genetics 2008;82(2):411-23.

Lin BD, Mbarek H, Willemsen G, Dolan CV, Fedko IO, Abdellaoui A, et al. Heritability and Genome-Wide Association Studies for Hair Color in a Dutch Twin Family Based Sample. Genes 2015;6(3):559-76.

Liu F, Visser M, Duffy DL, Hysi PG, Jacobs LC, Lao O, et al. Genetics of skin color variation in Europeans: genome-wide association studies with functional follow-up. Human genetics 2015;134(8):823-35.

Liu F, Wollstein A, Hysi PG, Ankra-Badu GA, Spector TD, Park D, et al. Digital quantification of human eye color highlights genetic association of three new loci. PLoS genetics 2010;6(5):e1000934.

Sulem P, Gudbjartsson DF, Stacey SN, Helgason A, Rafnar T, Jakobsdottir M, et al. Two newly identified genetic determinants of pigmentation in Europeans. Nature genetics 2008;40(7):835-7.

Sulem P, Gudbjartsson DF, Stacey SN, Helgason A, Rafnar T, Magnusson KP, et al. Genetic determinants of hair, eye and skin pigmentation in Europeans. Nature genetics 2007;39(12):1443-52.

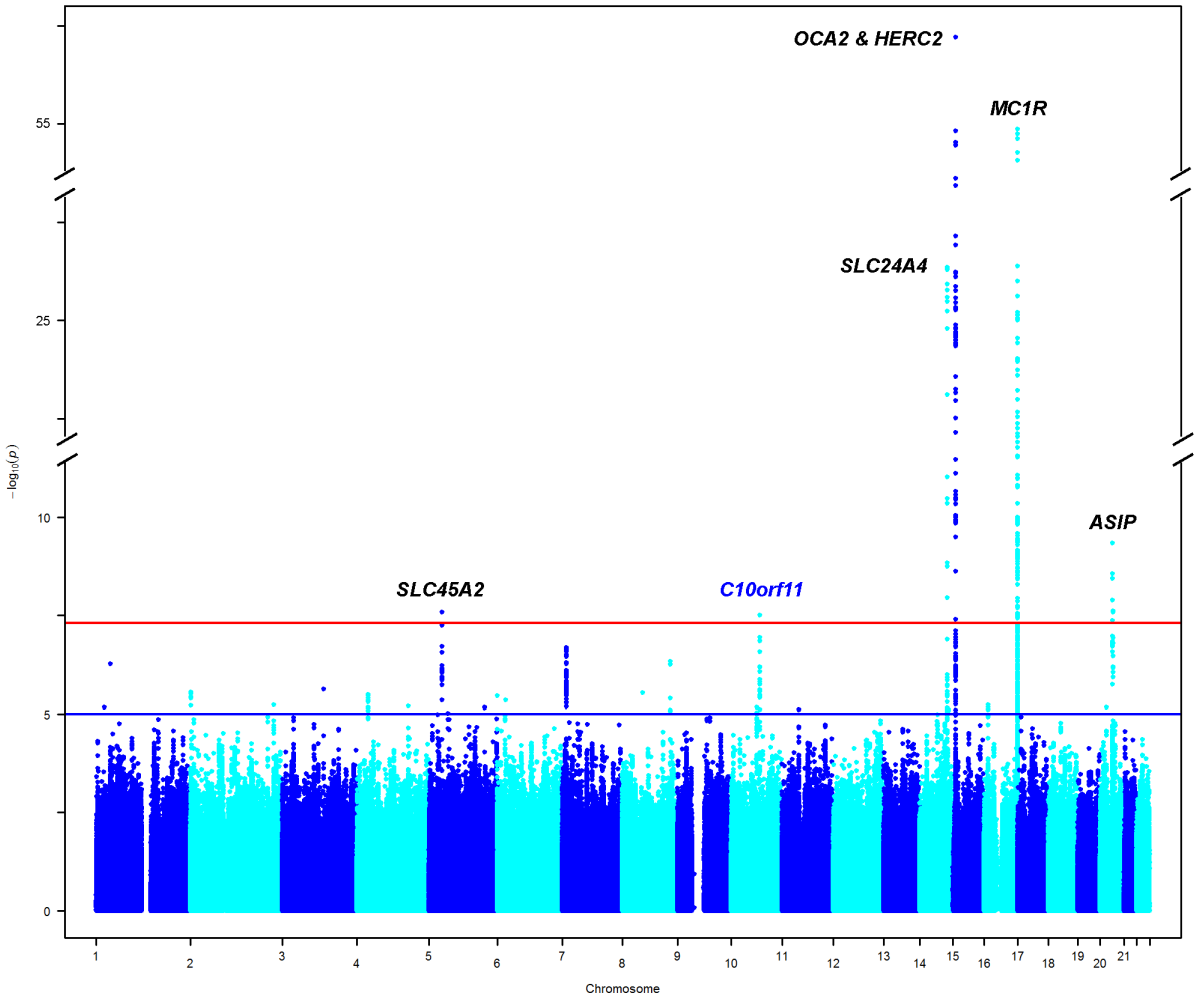
Visconti A, Duffy DL, Liu F, Zhu G, Wu W, Chen Y, et al. Genome-wide association study in 176,678 Europeans reveals genetic loci for tanning response to sun exposure. Nature communications 2018;9(1):1684.

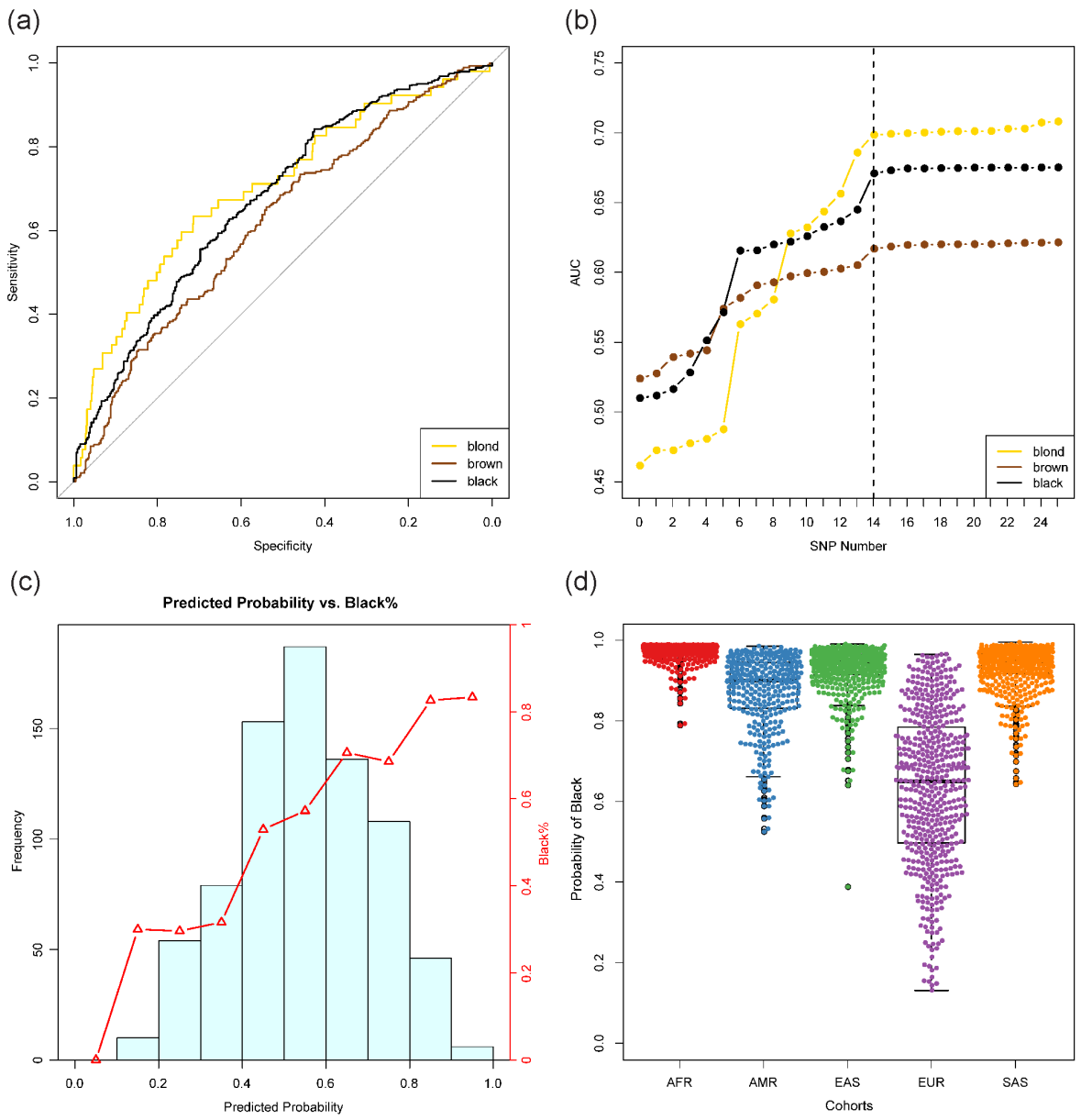
Walsh S, Chaitanya L, Clarisse L, Wirken L, Draus-Barini J, Kovatsi L, et al. Developmental validation of the HIrisPlex system: DNA-based eye and hair colour prediction for forensic and anthropological usage. Forensic science international Genetics 2014;9:150-61.

Figure legends

Figure 1. Manhattan plot of the discovery stage meta-analysis results for human eyebrow color from three European GWASs (RS, TwinsUK and QIMR, total N=6,513). The −log10 P values for association were plotted for each SNP according to chromosomal positions (genome assembly GRCh37.p13). Previously known pigmentation genes were marked using black text and gene to our knowledge previously unreported were highlighted in blue. The red and blue lines, respectively, correspond to the thresholds for genome-wide significance ($p=5\times10^{-8}$) and suggestive significance ($p=1\times10^{-5}$).

Figure 2. Genetic predictions of eyebrow color. **(a)** Illustration of the prediction performance of the set of 25 SNPs for eyebrow color prediction model using ROC curves with AUC estimates in independent individuals (N=779). **(b)** AUC was plotted against the number of SNPs included in the multinomial logistic model. The top 14-SNP annotation and prediction ranks are provided in **Table S11**. **(c)** Frequency (left y-axis) of predicted probability (x-axis) for black eyebrow and percentage (right y-axis) of non-black eyebrow in all 779 RS samples. **(d)** The distribution of predicted black eyebrow color probabilities for 2,504 subjects from the 1000-Genomes Project panel; samples are grouped according to the 5 continents they originate from: AFR-Africans, AMR-Native Americans, EAS-East Asians, EUR-Europeans, and SAS-South Asians.




(a)
Sensitivity
Specificity
blond
brown
black
(b)
AUC
SNP Number
blond
brown
black
(c)
Predicted Probability vs. Black%
Frequency
Black%
Predicted Probability
(d)
Probability of Black
AFR
AMR
EAS
EUR
SAS
Cohorts