MISS YINGJIE CHEN (Orcid ID : 0000-0002-1537-131X)
MR SAGNIK PALMAL (Orcid ID : 0000-0003-1318-7444)
DR J. CAMILO CHACÓN-DUQUE (Orcid ID : 0000-0003-0715-1947)
DR DESMOND JOHN TOBIN (Orcid ID : 0000-0003-4566-9392)
PROFESSOR SIJIA WANG (Orcid ID : 0000-0001-6961-7867)
DR PIERRE FAUX (Orcid ID : 0000-0003-4486-6327)

Article type : Original Article

A GWAS identifies novel gene associations with facial skin wrinkling and mole count in Latin-Americans

Y. Chen,[1]* M. André,[2,3,4]* K. Adhikari,[5,22] M. Blin,[2] B. Bonfante,[2] J. Mendoza-Revilla,[6,7] M. Fuentes-Guajardo,[8] S. Palmal,[2] J.C. Chacón-Duque,[9] M. Hurtado,[6] V. Villegas,[6] V. Granja,[6] C. Jaramillo,[10] W. Arias,[10] R.B. Lozano,[11,12] P. Everardo-Martínez,[11] J. Gómez-Valdés,[11] H. Villamil-Ramírez,[13] C.C.S. de Cerqueira,[14] T. Hünemeier,[15] V. Ramallo,[16,17] R. Gonzalez-José,[17] L. Schüler-Faccini,[16] M.-C. Bortolini,[16] V. Acuña-Alonzo,[11] S. Canizales-Quinteros,[13] C. Gallo,[6] G. Poletti,[6] G. Bedoya,[10] F. Rothhammer,[18] D. Balding,[19] D.J. Tobin,[20] S. Wang,[21] P. Faux[2†] and A. Ruiz-Linares[1,2,22†]

[1] 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.

[2] UMR 7268 ADES, CNRS, Aix-Marseille Université, EFS, Faculté de médecine Timone, Marseille, 13005, France.

[3] Estonian Biocentre, Institute of Genomics, University of Tartu, 51010 Tartu, Tartumaa, Estonia

[4] Department of Evolutionary Biology, Institute of Molecular and Cell Biology, University of Tartu, 51010 Tartu, Tartumaa, Estonia

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/BJD.20436

[5] School of Mathematics and Statistics, Faculty of Science, Technology, Engineering and Mathematics, The Open University, Milton Keynes, MK7 6AA, United Kingdom.
[6] Laboratorios de Investigación y Desarrollo, Facultad de Ciencias y Filosofía, Universidad Peruana Cayetano Heredia, Lima, 31, Perú.
[7] Unit of Human Evolutionary Genetics, Institut Pasteur, Paris, 75015, France.
[8] Departamento de Tecnología Médica, Facultad de Ciencias de la Salud, Universidad de Tarapacá, Arica, 1000000, Chile.
[9] Division of Vertebrates and Anthropology, Department of Earth Sciences, Natural History Museum, London, SW7 5BD, UK.
[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] Scientific Police of São Paulo State, Ourinhos-SP, 19900-109, Brazil.
[15] Departamento de Genética e Biologia Evolutiva, Instituto de Biociências, Universidade de São Paulo, São Paulo, SP, 05508-090, Brazil.
[16] Departamento de Genética, Universidade Federal do Rio Grande do Sul, Porto Alegre, 90040-060, Brasil.
[17] Instituto Patagónico de Ciencias Sociales y Humanas, Centro Nacional Patagónico, CONICET, Puerto Madryn, U9129ACD, Argentina.
[18] Instituto de Alta Investigación, Universidad de Tarapacá, Arica, Arica 1000000, Chile.
[19] Melbourne Integrative Genomics, Schools of BioSciences and Mathematics & Statistics, University of Melbourne, Melbourne, VIC 3010, Australia.
[20] The Charles Institute of Dermatology, University College Dublin, Ireland
[21] Chinese Academy of Sciences Key Laboratory of Computational Biology, Shanghai Institutes for Biological Sciences, University of the Chinese Academy of Sciences, Chinese Academy of Sciences, Shanghai 200031, China
[22] Department of Genetics, Evolution and Environment, and UCL Genetics Institute, University College London, London, WC1E 6BT, UK.

*Contributed equally to this work.

†Co-supervised this work

Corresponding author: Andrés Ruiz-Linares

Email: andresruiz@fudan.edu.cn; pierre.faux@univ-amu.fr

What is already known about this topic?

- A few loci have been associated with skin ageing features in Europeans but most of these loci have not been replicated in independent studies
- Previous studies have mainly focused on Europeans.

What does this study add?

- We replicate the association of *MC1R, IRF4* and *SLC45A2* with skin wrinkling in Latin Americans with Native American, European and African ancestry.
- We identify new wrinkling loci, including *VAV3*.
- We report a novel association with mole count: *SLC45A2*, a known melanoma-associated gene.

Abstract

Background: Genome-wide association studies (GWAS) have identified genes influencing skin **ageing** and mole count in Europeans but little is known about the relevance of these (or other genes) in non-Europeans.

Objective: To conduct a GWAS for facial skin ageing and mole count in adults < 40 years old, of mixed European, Native American and African ancestry, recruited in Latin America.

Methods: Skin **ageing** and mole count scores were obtained from facial photographs of >6,000 individuals. After quality control checks, three wrinkling traits and mole count were retained for genetic analyses. DNA samples were genotyped with Illumina's Omni Express chip. Association testing was performed on ~8,703,729 SNPs across the autosomal genome.

Results: Genome-wide significant association was observed at four genome regions: two were associated with wrinkling (in 1p13.3 and 21q21.2), one with mole count (in 1q32.3), and one with both wrinkling and mole count (in 5p13.2). Associated SNPs in 5p13.2 and in 1p13 are intronic within *SLC45A2* and *VAV3*, respectively, while SNPs in 1q32.3 are near the *SLC30A1* gene, and those in 21q21.2 occur in a gene desert. Analysis of SNPs in *IRF4* and *MC1R* are consistent with a role of these genes in skin ageing.

Conclusions: We replicate the association of wrinkling with variants in *SLC45A2, IRF4* and *MC1R* reported in Europeans. We identify *VAV3* and *SLC30A1* as two novel candidate genes impacting on wrinkling and mole count, respectively. We provide the first evidence that *SLC45A2* influences mole count, in addition to variants in this gene affecting melanoma risk in Europeans.

Introduction

Skin appearance in the general human population varies with age, reflecting changes in pigmentation and in various physical properties of the skin (such as cellularity, thickness, elasticity and moisture)[1–4], partly manifested as wrinkling. In addition, the number of moles has been reported to increase with age until about the third decade of life, after which it starts decreasing[5]. These changes in various skin features are influenced by a range of individual and environmental factors, including sun exposure (photo-damage), smoking, sex and ethnicity[6–13]. Two types of skin ageing have therefore been proposed[14]: intrinsic ageing (due largely to personal factors, such as genetics) and extrinsic ageing (due mainly to environmental/lifestyle factors[15,16]). Although changes to the face are combination of intrinsic and extrinsic skin ageing, the frank expression of extrinsic ageing (typically manifested with a coarse wrinkling) is readily detected after a considerable amount of time of exposure to skin damaging exogenous factors (generally decades, although photodamage occurs earlier in people of Northern European ancestry)[7,16]. Intrinsic ageing proceeds more subtly[7,14] but can be detected in relatively young facial skin, mostly as fine wrinkling. Recently, efforts to identify genetic determinants of skin ageing have employed genome-wide association studies (GWAS), usually in elderly individuals. These studies have detected significant associations of wrinkling with a handful of genomic regions[15,17–24], however, the robustness of these associations remains to be established as most associations have been observed in single studies, and often these studies have examined relatively small samples (i.e. less than ~1,000 individuals). By contrast, a relatively large number of loci influencing nevus count have been detected in several independent GWAS and their robustness further established by broad meta-analyses[20] comprising thousands of individuals. A drawback of these skin GWAS findings is that they have been essentially limited to individuals of European ancestry. Further research is needed in order to establish the role that the genetic associations identified in Europeans could have in in non-European populations and also to evaluate whether additional genetic variants, present in non-Europeans, could also influence these traits[25]. To this aim, we performed a GWAS for facial skin ageing and mole count in over 6,000 Latin Americans of mixed European, Native American and African ancestry. We studied relatively young adults so as to focus mostly on intrinsic ageing, which is likely to be under stronger genetic influence. We replicate the association of wrinkling with *IRF4, MC1R* and *SLC45A2* polymorphisms (previously reported in Europeans) and identify two novel loci associated with this trait, most notably *VAV3,* a

gene highly expressed in the skin. We also detect a previously unreported association of mole count with *SLC45A2*, a gene well known to be involved in pigmentation and in melanoma susceptibility.

Material and methods

Study subjects

We studied 6,254 individuals from 5 Latin American countries (Brazil=628; Chile=1,649; Colombia=1,633; Mexico=1,146; Peru=1,198). These individuals were recruited by the CANDELA consortium as part of a study on the genetics of physical appearance[26]. Most subjects recruited were students and staff from the universities participating in the CANDELA consortium. Individuals with any long-term medical condition or treatment were excluded. Data from these individuals have been included in previous GWAS for a range of traits, including pigmentation and facial features[27–31]. Summary information on the subjects studied here is presented in Table 1. All individuals provided written informed consent. Ethics approvals were obtained from: Universidad Nacional Autónoma de México (México), Universidad de Antioquia (Colombia), Universidad Perúana Cayetano Heredia (Perú), Universidad de Tarapacá (Chile), Universidade Federal do Rio Grande do Sul (Brazil) and University College London (UK).

Phenotyping

Due to the relatively young age of our study subjects (Table 1 and Supplementary figure S1), we adapted 12 traits from the SCINEXA protocol (*Score for INtrinsic and EXtrinsic skin Aging*) protocol[32] by including half points so as to allow for finer trait variation. SCINEXA includes four assesments of pigmentation spots (qualitative and quantitative scorings; on forehead and on cheeks), five types of wrinkles (on forehead, frown lines, crow's feet, below eyes and on upper lip) and three additional traits (cheeks laxity, telangiectasia on cheeks and prominence of the nasolabial fold – See Supplementary table S1 and figure S2). Moles were counted on a semi-quantitative scale: (0) no moles, (1) less than 5 pale moles, (2) 5 or more pale moles, (3) one dark mole, in relief, and (4) more than one dark mole, in relief. The full sample was randomized and divided between two raters that performed the scoring (Y.C.: ~65% and M.A.: ~35%). We then discarded traits with low variability and assessed the reliability of the trait scores by computing

Spearman correlations between four types of comparisons between rater scores (Supplementary table S2).

In addition to obtaining SCINEXA and mole count scores, we had available Melanin Index values for the inner arm of each individual as this study sample has been previously included in a GWAS of skin pigmentation[27].

DNA genotyping

DNA samples from participants were genotyped on the Illumina HumanOmniExpress chip including 730,525 SNPs. Genotyping quality control was as described in Adhikari *et al.* (2019)[27]. Briefly, we excluded markers with >5% missing data or minor-allele frequency <1% and individuals either with >5% missing data or failing the sex concordance check. After these checks, a total of 636,195 autosomal SNPs were retained for further analysis. Genotypes were phased using SHAPEIT2 and then genotypes at untyped SNPs were imputed from the 1,000 Genomes Phase 3 using IMPUTE2, resulting in 11,218,392 autosomal biallelic variants being available for study. Of these, we removed SNPs with poor imputation or genotyping quality scores (as in Adhikari *et al.* 2019[27]). The final dataset included genotypes for 8,703,729 autosomal biallelic SNPs.

Statistical genetics analyses

We estimated the narrow-sense heritability (h^2) and the genetic correlation for each pair of these traits using LDAK5 with default parameters[33,34]. For the genome-wide association analyses we excluded individuals with a missing trait score as well as all phenotypic and genetic outliers. Association was performed with PLINK v1.90b6.1[35] using a regression model including 10 covariates (age, sex, BMI, rater and 6 genetic PCs reflecting continental and sub-continental ancestry – as stated in previous GWAS studies of the CANDELA sample[27]). We computed association statistics only for SNPs with Minor Allele Frequency (MAF) ≥1% and with < 5% uncalled genotypes. To account for multiple testing across traits and SNPS we calculated a an adjusted significance threshold via the false-discovery rate (FDR) procedure of Benjamini-Hochberg[36]. In addition to the GWAS in the full study sample, we performed a GWAS separately in each of the five countries sampled. The resulting summary statistics were then combined in a meta-analysis using PLINK's implementation (--meta-analysis weighted-z) of METAL[37].

All additional numerical analyses (computation of phenotype principal components, phenotypic correlations, correlations between traits and covariates) were carried out with MATLAB ® 9.6.0. (R2019a). Individual continental (African, European, Native American) ancestry was previously[26] estimated on the CANDELA dataset using the software ADMIXTURE[38] with $K = 3$. The average African, European and Native American ancestry of the individuals included in the GWAS was estimated as 4.4%, 51.1% and 44.5%, respectively.

Results and discussion

Traits examined

Using a modified SCINEXA[32], we evaluated six wrinkling and four pigmentation changes related to skin ageing in over 6,000 Latin Americans (Supplementary figures S2, S3 and Table S1). To focus on the most reliably scored traits, we restricted subsequent analyses to traits having correlations >60% for 3/4 quality assessments (Supplementary table S2). This led to the four pigmentation traits being discarded. Among the wrinkling traits, two (wrinkles in the upper lip and teleangiectasia on cheeks) showed minimal variation (Supplementary figure S4) and were also excluded from further analysis. This resulted in four wrinkling traits (wrinkles under eyes, glabellar -"frown lines"-, fine lateral canthal rhytids -"crow's feet"- and fine forehead wrinkles) and mole count being retained (Supplementary figure S5). In addition, since the four wrinkling traits retained are positively correlated (Supplementary table S3), we performed a Principal Components Analysis (PCA) and retained the first PC (PC1), as an overall wrinkling score. PC1 explains ~57% of the total phenotypic variance and correlates strongly with the sum of the 4 wrinkling scores ($\rho = 0.99$).

We evaluated the correlation between skin traits and individual covariates (age, BMI, sex, constitutive pigmentation and genetic ancestry, Supplementary table S4). Wrinkling showed a moderate but highly significant correlations with age (Spearman ρ +0.29 to +0.41, p- $1x10^{-113}$ to 110^{-162}) . Wrinkling also correlated with body mass index (BMI – Spearman ρ ~0.15; p-values $1x10^{-113}$ to $1x10^{-162}$), probably reflecting the positive correlation of BMI with age in young adults (Supplementary table S5). Forehead wrinkles were more frequent in men than in women (p-value $<1x10^{-182}$). A weak but significant correlation was observed between wrinkles PC1 and constitutive skin pigmentation (i.e. Melanin Index, Spearman ρ = -0.06, p-value $3x10^{-5}$) and with genetic ancestry (European: Spearman ρ = +0.11, p-value $6x10^{-17}$; Native American: Spearman ρ

= -0.07, p-value $2x10^{-8}$; African: Spearman ρ = -0.09, p-value $8x10^{-11}$). These observations are consistent with the reported greater wrinkling with age in Europeans compared to non-Europeans[6,39]. From the chip data, a low to moderate heritability was estimated for the skin traits (h^2 = 0.15-0.45, Supplementary figure S6 and table S6), genetic correlations being on average ~10% higher than the phenotypic estimates (Supplementary figure S6 and table S3).

Genome-wide association analyses

We detected genome-wide significant association (p-values $5<10^{-8}$) for SNPs in four genomic regions (Figure 1 and Table 2). Of these, one region (5p13.2) is associated with both wrinkling (PC1) and mole count, two regions only with wrinkling (1p13.3 with PC1 and 21q21.2 with forehead wrinkles) and one region only with mole count (1q32.3). The p-values for the index SNPs (the one with the smallest p-value) in these four regions exceed the False Discovery Rate threshold accounting for the number of traits and SNPs tested (4.95×10^{-8}). We evaluated replication of these association signals across the five country sampled by conducting a GWAS in each country separately. We found allelic effects in the same direction across all countries for the four regions showing genome-wide significant association (Supplementary figure S7).

The region in 5p13.2, associated with wrinkles PC1 and mole count, comprises a cluster of SNPs within *SLC45A2* (Table 2, Figure 2). Strongest association with wrinkles PC1 (p-value 3×10^{-9}) was observed for SNP rs173662. Several SNPs in this region also approach genome-wide significant association for Crow's Feet (p-value = 3×10^{-7}, Supplementary table S7). A recent European study by Law et al. [40] reported a significant association of SNPs in *SLC45A2* with skin patterning on imprints from the back of the hand, with smallest p-value for rs185146, a SNP also included in the associated region shown in Figure 2. In our data, rs185146 exceeds the genome-wide significant threshold for wrinkles PC1 (p-value = 1×10^{-8}), and the suggestive threshold for Crow's Feet (p-value = 1×10^{-6} – Supplementary table S7).

Mole count showed strongest association with rs34466007 (p-value = 9×10^{-11}). This SNP has a low polymorphism in Northern Europeans (minor allele frequency or MAF of 1.5% in CEU – Supplementary table S8), in whom most mole count association studies have been performed so far. By contrast rs34466007 is highly polymorphic in the CANDELA sample (MAF of 50%), indicating that we have considerably higher power, than previous studies, to detect an effect of this

SNP on mole count. As variants in *SLC45A2* are associated with pigmentation and since it has been reported that mole count could be biased by skin colour[41], we repeated the mole count association analyses including skin pigmentation as a covariate. We obtained very similar association results (p-value = 4.3×10^{-10}), indicating that the association detected here is not due to pigmentation acting as a confounder.

The cluster of SNPs in *SCL45A2* for which we found association with mole count and wrinkles PC1 spans about 20kb and defines a high LD block (r^2>0.93; Figure 2). This region includes rs16891982, a SNP encoding a functional non-synonymous amino-acid substitution in SCL45A2 (F374L)[27]. We recently reported strong association of rs16891982 with skin pigmentation in the Latin American sample examined here[27]. Figure 3 contrasts three skin features of the CANDELA sample based on rs16891982 genotype: pigmentation, wrinkles PC1 and mole count. The derived G allele at rs16891982 (of high frequency in Europeans, Supplementary table S9), is associated with lower pigmentation, lower mole count and higher wrinkling.

SNPs in 1p13.3 associated with wrinkles PC1 are intronic in the VAV3 (Vav Guanine Nucleotide Exchange Factor 3) gene. VAV proteins are guanine nucleotide exchange factors (GEFs) for Rho family GTPases[42], playing an important role in the regulation of the cytoskeleton[43], and have been shown to be involved in cancers of the skin and other epithelia[44,45]. Consitent with an important role of VAV3 in skin biology, in the GTEx database (https://gtexportal.org/home/[46]), *VAV3* is observed to be maximally expressed in skin and other epithelial tissues (Figure 4). Furthermore, in the Regulome database, the index SNP in 1p13.3 (rs2504460) is annotated as an active transcription starting site (TSS)[47]. In fact, two major transcripts for *VAV3* have been identified: a Vav3 alpha (4,776 bp; NM_006113.4) and a truncated version denoted Vav3.1 (3,115 bp; NM_001079874.1[43]). Our index SNP (rs2504460) is located near the VAV3.1 transcription start site, a region rich in enhancer-like signatures and a high H3K27Ac histone mark in epidermal keratinocytes (Figure 4). This suggests that rs2504460 could play a role in regulating transcription of the Vav3.1 variant. Both VAV3 transcripts are protein-coding (Vav3 alpha, NP_006104.4, 847 aa; Vav3.1, NP_001073343.1, 287 aa); with the truncated form lacking several conserved protein domains (Figure 4). The differential expression of these two VAV3 variants has been linked to cancer progression[43].

The SNPs in 1q32.3 associated with mole count, show strongest association for rs6679498 (p-value 2.9 10^{-8}). The gene closest to rs6679498 is *SLC30A1* (~32 Kb). In the GTEx database,

this gene is highly expressed in sun-exposed skin (second among 55 tissues – Supplementary figure S8). Although, this gene has not been associated with any skin-related traits, the known role of SLC30A1 in the cellular transport of Ca^{++} and Zn^{++} could be relevant for skin biology[48,49]. In particular, Ca^{++} levels impact on cell proliferation and differentiation, and in the mobility and viability of melanoma cells, and Zn^{++} can induce apoptosis in melanoma cells in vitro[50]. SLC30A1 could therefore be involved in the emergence of skin moles, which result from melanocyte neoplasia and hyperplasia[51,52], by affecting melanocyte proliferation and viability.

The fourth region showing evidence of association in the Latin American sample examined here is in 21q21.2. A total of 18 SNPs in this region are associated with forehead wrinkles at genome-wide significance, with strongest association being observed for rs147991442 (p-value = 6.0×10^{-10}). The associated SNPs are located in a gene desert, with no noticeable candidate gene in the vicinity (LINC01689 being ~330kb from rs147991442) and no relevant association for these SNPs has been reported in other studies.

Since we detected a significant effect of sex on several skin ageing traits (Supplementary table S4), we also performed GWASes for these traits separately in men and women (Supplementary figure S9). As expected from a drop in power due to the smaller sample size (relative to the combined study), we did not detect the association signals discussed above. We detected four novel sex-specific association signals, two women- and two men-specific (Supplementary table S10). However, three of these signals are only marginally significant (and involve a single SNP), while the fourth occurs in a large intergenic region (on 12p12.3, Supplementary figure S9-3), complicating the biological interpretations of these observations.

Evaluation of *MC1R* and *IRF4* variants associated with skin ageing in Europeans

A substantial number of loci have been robustly associated with mole count[23]. However, the same is not the case for skin wrinkling (and other skin ageing features), for which only two gene regions, *MC1R* and *IRF4,* have been associated in more than one study[19,40]. These two gene regions had been previously identified as playing a role in hair colour, specifically red hair (*MC1R*)[53,54] and blond hair (*IRF4*) [27,55]. These hair colour variants are effectively West Eurasian traits, and the underlying associated *MC1R/IRF4* alleles having low to moderate frequency (<15%) only in West Eurasians[28,56,57]. Considering the partial (51.1%) European ancestry of the CANDELA sample we examined the evidence of association with wrinkles PC1 specifically for

variants in *IRF4* and *MC1R* previously implicated in skin ageing in Europeans. Specifically, we examined (Figure 5): (i) three index SNPs from a GWAS[40] (rs12203592 in *IRF4*, and rs4268748 and rs35063026 in *MC1R*) and (ii) compound genotypes comprising *MC1R* red-hair alleles of variable penetrance (Supplementary text S1)[19,40].

We find significant association (at a Bonferroni-corrected *p*-value threshold of 0.0125) of wrinkling PC 1 with rs12203592 in *IRF4* (*p*-value 2.8×10^{-4}), the European T allele leading to greater wrinkling. This SNP also has the smallest *p*-value across the entire *IRF4* region in the CANDELA data. Interestingly, we previously reported a genome-wide significant association of rs12203592 with hair graying, another integumentary ageing feature, in another GWAS of the CANDELA sample[28], suggesting a broader role of *IRF4* in ageing. Hair graying, like fine wrinkles, can occur even at a relatively early age, and is also not strongly associated with exogenous ageing via noxious environmental influences[58]. Consistent with the admixed ancestry of Latin Americans, the frequency of the T allele at rs12203592 is ~50% of that reported for Europeans (~15%). Thus, although there is lower power to detect phenotypic effects for this allele in Latin Americans relative to Europeans, this is to some extent compensated by the size of the CANDELA sample.

Of the two SNPs examined in *MC1R* we find significant association for rs35063026 (*p*-value = 2×10^{-3}, the direction of the effect being the same as previously reported). We also observe a marginally significant association with the *MC1R* red-hair compound genotypes (*p*-value = 0.012 – Figure 5b). Consistent with previous reports, the genotypes associated with greater wrinkling in the CANDELA sample are those carrying highly penetrant R alleles, suggesting that these alleles have a larger additive effect on wrinkling but that their low frequency in Latin Americans reduces the significance of the association. Overall, our findings are consistent with *MC1R* variants impacting on wrinkling in Latin Americans, although the association signal is probably weakened by their low frequency.

Discussion

Despite the marked differences in phenotyping approach and study population, here we provide the first independent replication of the effect of *SLC45A2* on wrinkling reported by Law et al. [40]. Noticeably, Law et al. analysed young cohorts (mean age 10-45), as we have done here, further supporting a role for *SLC45A2* in intrinsic skin ageing. More broadly, our observations are

consistent with the higher rate of wrinkling of Europeans, probably reflecting the greater sensitivity to UV damage of more lightly pigmented skin[6,39]. In addition, although other studies have strongly associated *SLC45A2* with pigmentation and melanoma[59,60], to our knowledge this is the first time that variants in this gene region are associated with mole count, a well-established risk factor for melanoma[61,62].

In conclusion, the effect of variants at *SLC45A2, IRF4* and *MC1R* on skin ageing is detectable in young adults of mixed Native American, European and African ancestry, which have not yet developed the classic deep-wrinkling phenotype associated with long-term skin photodamage. The high non-European ancestry of Latin Americans facilitated the observation of an effect on mole count of *SLC45A2* variants that are rare in Europeans. It will be important to follow-up the novel genetic loci detected here by examining association in other large study cohorts and by evaluating if a differential impact of the two *VAV3* variants on cytoskeleton regulation could explain the association of this gene with skin ageing[63].

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, 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 acknowledge the Centre de Calcul Intensif d'Aix-Marseille for granting access to high-performance computing resources. We thank Nick Martin, Matthew Law and David Duffy for helpful discussions.

Ethics approval or consent to participate

Ethics approvals were obtained from: Universidad Nacional Autónoma de México (México), Universidad de Antioquia (Colombia), Universidad Perúana Cayetano Heredia (Perú), Universidad de Tarapacá (Chile), Universidade Federal do Rio Grande do Sul (Brazil) and University College London (UK).

Funding: Work leading to this publication received funded from: the National Natural Science Founddation 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), 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), 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. B.B. is supported by a doctoral scholarship from Ecole Doctorale 251 Aix-Marseille Université.

Availability of data and materials

Raw genotype or phenotype data cannot be made available due to restrictions imposed by the ethics approval. Summary statistics from the GWAS analyses will be deposited at GWAS central and a link to the submission will be available before publication.

URLs

The GTEx data used in this manuscript were obtained from file GTEx_Analysis_2017-06-05_v8_RNASeQCv1.1.9_gene_tpm.gct.gz, downloaded from the GTEx Portal (https://gtexportal.org/home/datasets) on 06/18/20. We used the 2020-06-94 release of the version 1.02 of the GWAS catalog, downloaded from https://www.ebi.ac.uk/gwas/docs/file-downloads. The data showing regional enrichment in H3K27Ac histone mark in NHEK cell line (Figure 4)

was obtained from file wgEncodeBroadHistoneNhekH3k27acStdSig.bigWig, downloaded on the UCSC Genome Browser repository on 09/07/20 at https://hgdownload.soe.ucsc.edu/gbdb/hg38/bbi/wgEncodeReg/wgEncodeRegMarkH3k27ac/.

References

1. Shuster S, Black MM, McVITIE EVA. The influence of age and sex on skin thickness, skin collagen and density. *Br J Dermatol*. 1975;93(6):639-643. doi:10.1111/j.1365-2133.1975.tb05113.x

2. Jacobs LC, Liu F, Bleyen I, et al. Intrinsic and Extrinsic Risk Factors for Sagging Eyelids. *JAMA Dermatol*. 2014;150(8):836. doi:10.1001/jamadermatol.2014.27

3. Verdier-Sévrain S, Bonté F. Skin hydration: a review on its molecular mechanisms. *J Cosmet Dermatol*. 2007;6(2):75-82. doi:10.1111/j.1473-2165.2007.00300.x

4. Slominski A, Tobin DJ, Shibahara S, Wortsman J. Melanin Pigmentation in Mammalian Skin and Its Hormonal Regulation. *Physiol Rev*. 2004;84(4):1155-1228. doi:10.1152/physrev.00044.2003

5. Schäfer T, Merkl J, Klemm E, Wichmann H-E, Ring J, group K study. The Epidemiology of Nevi and Signs of Skin Aging in the Adult General Population: Results of the KORA-Survey 2000. *J Invest Dermatol*. 2006;126(7):1490-1496.

6. Alexis AF, Obioha JO. Ethnicity and Aging Skin. *J Drugs Dermatol JDD*. 2017;16(6):s77-s80.

7. Tobin DJ. Pigmentation and Photoaging. *Cutan Photoaging*. 2019;19:145.

8. Yaar M, Gilchrest BA. Photoageing: mechanism, prevention and therapy. *Br J Dermatol*. 2007;157(5):874-887. doi:10.1111/j.1365-2133.2007.08108.x

9. Bulpitt CJ. Why do some people look older than they should? *Postgrad Med J*. 2001;77(911):578-581. doi:10.1136/pmj.77.911.578

10. Giacomoni PU, Mammone T, Teri M. Gender-linked differences in human skin. *J Dermatol Sci*. 2009;55(3):144-149. doi:10.1016/j.jdermsci.2009.06.001

11. Langton AK, Alessi S, Hann M, et al. Aging in skin of color: disruption to elastic fiber organization is detrimental to skin's biomechanical function. *J Invest Dermatol*. 2019;139(4):779-788.

12. Campiche R, Trevisan S, Séroul P, et al. Appearance of aging signs in differently pigmented facial skin by a novel imaging system. *J Cosmet Dermatol*. 2019;18(2):614-627.

13. Langton AK, Hann M, Costello P, et al. Heterogeneity of fibrillin-rich microfibrils extracted from human skin of diverse ethnicity. *J Anat*. 2020;237(3):478-486.

14. Tobin DJ. Introduction to skin aging. *J Tissue Viability*. 2017;26(1):37-46. doi:10.1016/j.jtv.2016.03.002

15. Flood KS, Houston NA, Savage KT, Kimball AB. Genetic basis for skin youthfulness. *Clin Dermatol*. 2019;37(4):312-319.

16. Sachs DL, Varani J, Chubb H, et al. Atrophic and hypertrophic photoaging: Clinical, histologic, and molecular features of 2 distinct phenotypes of photoaged skin. *J Am Acad Dermatol*. 2019;81(2):480-488.

17. Le Clerc S, Taing L, Ezzedine K, et al. A Genome-Wide Association Study in Caucasian Women Points Out a Putative Role of the STXBP5L Gene in Facial Photoaging. *J Invest Dermatol*. 2013;133(4):929-935. doi:10.1038/jid.2012.458

18. Chang ALS, Atzmon G, Bergman A, et al. Identification of Genes Promoting Skin Youthfulness by Genome-Wide Association Study. *J Invest Dermatol*. 2014;134(3):651-657. doi:10.1038/jid.2013.381

19. Jacobs LC, Hamer MA, Gunn DA, et al. A Genome-Wide Association Study Identifies the Skin Color Genes IRF4 , MC1R , ASIP , and BNC2 Influencing Facial Pigmented Spots. *J Invest Dermatol*. 2015;135(7):1735-1742. doi:10.1038/jid.2015.62

20. Martin AR, Lin M, Granka JM, et al. An Unexpectedly Complex Architecture for Skin Pigmentation in Africans. *Cell*. 2017;171(6):1340-1353.e14. doi:10.1016/j.cell.2017.11.015

21. Sulem P, Gudbjartsson DF, Stacey SN, et al. Genetic determinants of hair, eye and skin pigmentation in Europeans. *Nat Genet*. 2007;39(12):1443-1452. doi:10.1038/ng.2007.13

22. Sulem P, Gudbjartsson DF, Stacey SN, et al. Two newly identified genetic determinants of pigmentation in Europeans. *Nat Genet*. 2008;40(7):835-837. doi:10.1038/ng.160

23. Duffy DL, Zhu G, Li X, et al. Novel pleiotropic risk loci for melanoma and nevus density implicate multiple biological pathways. *Nat Commun*. 2018;9(1). doi:10.1038/s41467-018-06649-5

24. Laville V, Clerc SL, Ezzedine K, et al. A genome wide association study identifies new genes potentially associated with eyelid sagging. *Exp Dermatol*. 2019;28(8):892-898. doi:10.1111/exd.13559

25. Liu Y, Gao W, Koellmann C, et al. Genome-wide scan identified genetic variants associated with skin aging in a Chinese female population. *J Dermatol Sci*. 2019;96(1):42-49. doi:10.1016/j.jdermsci.2019.08.010

26. Ruiz-Linares A, Adhikari K, Acuña-Alonzo V, et al. Admixture in Latin America: Geographic Structure, Phenotypic Diversity and Self-Perception of Ancestry Based on 7,342 Individuals. Di Rienzo A, ed. *PLoS Genet*. 2014;10(9):e1004572. doi:10.1371/journal.pgen.1004572

27. Adhikari K, Mendoza-Revilla J, Sohail A, et al. A GWAS in Latin Americans highlights the convergent evolution of lighter skin pigmentation in Eurasia. *Nat Commun*. 2019;10(1). doi:10.1038/s41467-018-08147-0

28. Adhikari K, Fontanil T, Cal S, et al. A genome-wide association scan in admixed Latin Americans identifies loci influencing facial and scalp hair features. *Nat Commun*. 2016;7:10815. doi:10.1038/ncomms10815

29. Adhikari K, Fuentes-Guajardo M, Quinto-Sánchez M, et al. A genome-wide association scan implicates DCHS2, RUNX2, GLI3, PAX1 and EDAR in human facial variation. *Nat Commun*. 2016;7:11616. doi:10.1038/ncomms11616

30. Adhikari K, Reales G, Smith AJP, et al. A genome-wide association study identifies multiple loci for variation in human ear morphology. *Nat Commun*. 2015;6(1). doi:10.1038/ncomms8500

31. Bonfante B, Faux P, Navarro N, et al. A GWAS in Latin Americans identifies novel face shape loci, implicating VPS13B and a Denisovan introgressed region in facial variation. *Sci Adv*. 2021;7(6):eabc6160. doi:10.1126/sciadv.abc6160

32. Vierkötter A, Ranft U, Krämer U, Sugiri D, Reimann V, Krutmann J. The SCINEXA: A novel, validated score to simultaneously assess and differentiate between intrinsic and extrinsic skin ageing. *J Dermatol Sci*. 2009;53(3):207-211. doi:10.1016/j.jdermsci.2008.10.001

33. Speed D, Hemani G, Johnson MR, Balding DJ. Improved Heritability Estimation from Genome-wide SNPs. *Am J Hum Genet*. 2012;91(6):1011-1021. doi:10.1016/j.ajhg.2012.10.010

34. Speed D, Cai N, Johnson MR, Nejentsev S, Balding DJ. Re-evaluation of SNP heritability in complex human traits. *Nat Genet*. 2017;49(7):986-992. doi:10.1038/ng.3865

35. Purcell S, Neale B, Todd-Brown K, et al. PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *Am J Hum Genet*. 2007;81(3):559-575. doi:10.1086/519795

36. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J R Stat Soc Ser B Methodol*. 1995;57(1):289-300. doi:10.1111/j.2517-6161.1995.tb02031.x

37. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics*. 2010;26(17):2190-2191. doi:10.1093/bioinformatics/btq340

38. Alexander DH, Novembre J, Lange K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res*. 2009;19(9):1655-1664. doi:10.1101/gr.094052.109

39. Venkatesh S, Maymone MBC, Vashi NA. Aging in skin of color. *Clin Dermatol*. 2019;37(4):351-357. doi:10.1016/j.clindermatol.2019.04.010

40. Law MH, Medland SE, Zhu G, et al. Genome-Wide Association Shows that Pigmentation Genes Play a Role in Skin Aging. *J Invest Dermatol*. 2017;137(9):1887-1894. doi:10.1016/j.jid.2017.04.026

41. Zalaudek I, Argenziano G, Mordente I, et al. Nevus Type in Dermoscopy Is Related to Skin Type in White Persons. *Arch Dermatol*. 2007;143(3). doi:10.1001/archderm.143.3.351

42. Bustelo XR. Vav family exchange factors: an integrated regulatory and functional view. *Small GTPases*. 2014;5(2):e973757. doi:10.4161/21541248.2014.973757

43. Boesch M, Reimer D, Sopper S, Wolf D, Zeimet AG. (Iso-)form Matters: Differential Implication of Vav3 Variants in Ovarian Cancer. *The Oncologist*. 2018;23(7):757-759. doi:10.1634/theoncologist.2017-0683

44. Menacho-Márquez M, García-Escudero R, Ojeda V, et al. The Rho Exchange Factors Vav2 and Vav3 Favor Skin Tumor Initiation and Promotion by Engaging Extracellular Signaling Loops. Hynes N, ed. *PLoS Biol*. 2013;11(7):e1001615. doi:10.1371/journal.pbio.1001615

45. Boesch M, Sopper S, Marth C, et al. Evaluation of Vav3.1 as prognostic marker in endometrial cancer. *J Cancer Res Clin Oncol*. 2018;144(10):2067-2076. doi:10.1007/s00432-018-2725-2

46. Lonsdale J, Thomas J, Salvatore M, et al. The Genotype-Tissue Expression (GTEx) project. *Nat Genet*. 2013;45(6):580-585. doi:10.1038/ng.2653

47. Boyle AP, Hong EL, Hariharan M, et al. Annotation of functional variation in personal genomes using RegulomeDB. *Genome Res*. 2012;22(9):1790-1797. doi:10.1101/gr.137323.112

48. Segal D, Ohana E, Besser L, Hershfinkel M, Moran A, Sekler I. A role for ZnT-1 in regulating cellular cation influx. *Biochem Biophys Res Commun*. 2004;323(4):1145-1150. doi:10.1016/j.bbrc.2004.08.211

49. Bin B-H, Hojyo S, Seo J, et al. The Role of the Slc39a Family of Zinc Transporters in Zinc Homeostasis in Skin. *Nutrients*. 2018;10(2):219. doi:10.3390/nu10020219

50. Provinciali M, Pierpaoli E, Bartozzi B, Bernardini G. Zinc Induces Apoptosis of Human Melanoma Cells, Increasing Reactive Oxygen Species, p53 and FAS Ligand. *Anticancer Res*. Published online 2015:8.

51. Capiod T. The Need for Calcium Channels in Cell Proliferation. :14.

52. Roh MR, Eliades P, Gupta S, Tsao H. Genetics of melanocytic nevi. *Pigment Cell Melanoma Res*. 2015;28(6):661-672. doi:10.1111/pcmr.12412

53. Valverde P, Healy E, Jackson I, Rees JL, Thody AJ. Variants of the melanocyte–stimulating hormone receptor gene are associated with red hair and fair skin in humans. *Nat Genet*. 1995;11(3):328-330.

54. Box NF, Wyeth JR, O'gorman LE, Martin NG, Sturm RA. Characterization of melanocyte stimulating hormone receptor variant alleles in twins with red hair. *In*. Published online 1997.

55. Han J, Kraft P, Nan H, et al. A Genome-Wide Association Study Identifies Novel Alleles Associated with Hair Color and Skin Pigmentation. Abecasis G, ed. *PLoS Genet*. 2008;4(5):e1000074. doi:10.1371/journal.pgen.1000074

56. Katsara M-A, Nothnagel M. True colors: A literature review on the spatial distribution of eye and hair pigmentation. *Forensic Sci Int Genet*. 2019;39:109-118. doi:10.1016/j.fsigen.2019.01.001

57. Coop G, Pickrell JK, Novembre J, et al. The Role of Geography in Human Adaptation. Schierup MH, ed. *PLoS Genet*. 2009;5(6):e1000500. doi:10.1371/journal.pgen.1000500

58. O'Sullivan JDB, Nicu C, Picard M, et al. The biology of human hair greying. *Biol Rev*. Published online September 23, 2020. doi:10.1111/brv.12648

59. Barrett JH, Iles MM, Harland M, et al. Genome-wide association study identifies three new melanoma susceptibility loci. *Nat Genet*. 2011;43(11):1108-1113.

60. Ransohoff KJ, Wu W, Cho HG, et al. Two-stage genome-wide association study identifies a novel susceptibility locus associated with melanoma. *Oncotarget*. 2017;8(11):17586-17592. doi:10.18632/oncotarget.15230

61. Bataille V, Bishop J, Sasieni P, et al. Risk of cutaneous melanoma in relation to the numbers, types and sites of naevi: a case-control study. *Br J Cancer*. 1996;73(12):1605-1611. doi:10.1038/bjc.1996.302

62. Gandini S, Sera F, Cattaruzza MS, et al. Meta-analysis of risk factors for cutaneous melanoma: I. Common and atypical naevi. *Eur J Cancer*. 2005;41(1):28-44. doi:10.1016/j.ejca.2004.10.015

63. Zouboulis CC, Makrantonaki E, Nikolakis G. When the skin is in the center of interest: An aging issue. *Clin Dermatol*. 2019;37(4):296-305. doi:10.1016/j.clindermatol.2019.04.004

Figure legends

Figure 1. Aggregated Manhattan plot of association p-values for the skin traits examined. Each dot represents a SNP, with **colours** highlighting different chromosomes. For each genome-wide significant hit ($-\log_{10}$ *p*-value >7.3; green line), we indicate the trait showing strongest association. The red line marks the genome-wide suggestive (1E-5) threshold.

Figure 2. Regional association *p*-values for SNPs in *SLC45A2* (in 5p13.2) with mole count (top) and wrinkles PC1 (bottom). The plots cover a window of 50 Kb overlapping *SLC45A2* (intron/exon boundaries are shown at the bottom of the figure). Index SNPs (rs34466007 for mole count and rs173662 for wrinkles PC1) are shown as diamonds. The **colour** palette represents the strength of LD between each SNP and the index SNP. The double-end arrow at the top of the figure delimits the haplotype block discussed in the text (and Supplementary table S9), with the the amino-acid-changing rs16891982 labelled in green (a dotted line indicating the exon location of this SNP).

Figure 3. Box plots of skin pigmentation (top), wrinkles PC1 (middle) and mole count (bottom) in the CANDELA sample for genotypes of SNP rs16891982 in *SLC45A2*. Red lines represent the median phenotypic value, with blue boxes covering the 25th to 75th percentiles and whiskers marking the extent of individual values. The average phenotypic value is given above each box plot.

Figure 4. Association plot for the 1p13.3 region and *VAV3* mRNA expression levels in epithelial tissues. (a) Annotations are as in Figure 2 with the purple line below the association plot showing the regional enrichment of the H3K27Ac histone mark in normal human epidermal keratinocytes (NHEK cell line, from the Encode project data, downloaded from the UCSC Genome Browser). The two *VAV3* transcripts (VAV3 alpha and VAV3.1 – following the nomenclature of Boesch et al. 2018[43]) are shown at the bottom of the figure. The vertical green dashed line highlights that our index SNP (rs2504460 – shown as a diamond) is located near the start site of VAV3.1 transcript, in the vicinity of the H3K27Ac enrichment peak. (b) Boxplots for the 5 GTEx tissues with highest levels of *VAV3* transcription (of 55 tissues examined). Gene

expression levels are given in transcripts per million (TPM), with tissues ranked by decreasing median TPM values.

Figure 5. Association testing (with wrinkles PC1) for variants in *IRF4* and *MC1R* previously associated with skin ageing features in Europeans. Panel (a) shows results for index SNPs from a European GWAS. TA: tested allele; F(TA): frequency of the tested allele. Panel (b) shows the average wrinkling score (and s.e.) for compound *MC1R* genotypes for 3 types of red-hair alleles: + (wild), r (weakly-penetrant) and R (strongly-penetrant). The trend was tested using a regression model (Supplementary text S1).

Tables

Table 1. General characteristics of the CANDELA sample.

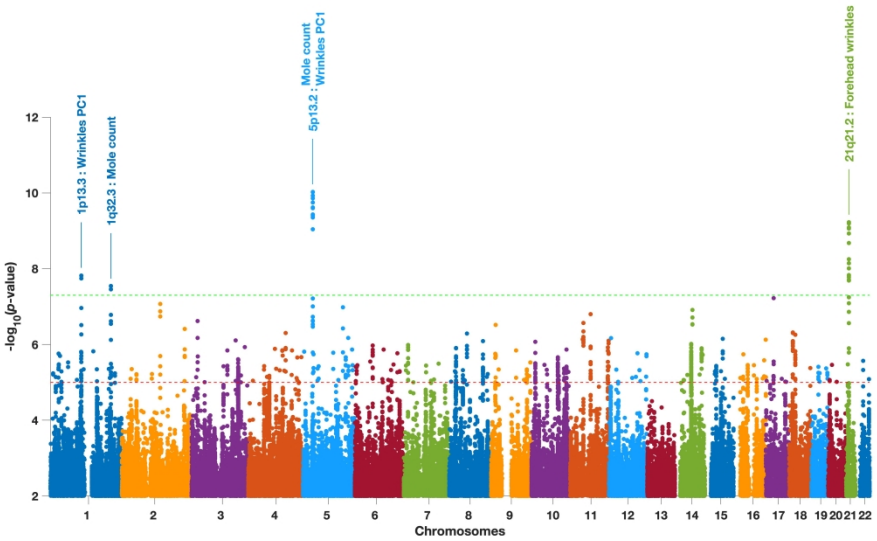
				Mean continental ancestry share [%][b]		
	N (% of females|males)	**Mean age (SD)**	**Mean Melanin Index[a]**	**African**	**European**	**Native American**
Whole study	6,254 (54|46)	24.0 (5.3)	35.1	4.4	50.8	44.8
Brazil	620 (69|31)	25.0 (5.5)	32.6	6.1	84.5	9.4
Chile	1,604 (38|62)	24.9 (5.3)	36.2	2.1	49.5	48.4
Colombia	1,647 (56|44)	24.0 (5.2)	33.8	8.2	62.9	28.9
Mexico	1,163 (60|40)	24.5 (5.4)	36.4	2.9	39.0	58.1
Peru	1,220 (58|42)	21.9 (4.7)	35.4	2.9	30.6	66.5

[a] Measured on the inner arm of each individual.

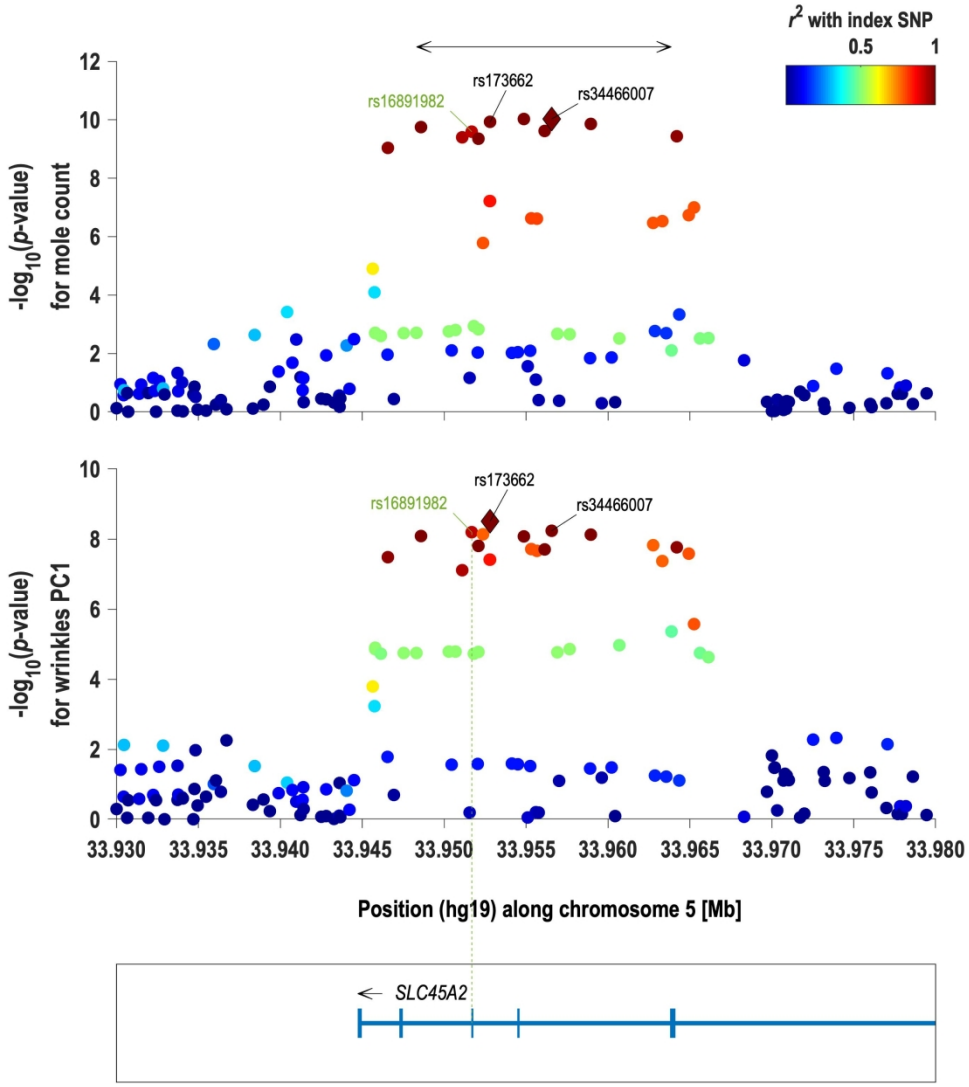
[b] Estimated using genome-wide SNP as in Ruiz-Linares et al. (2014)[26].

Table 2. Features of chromosomal regions showing genome-wide significant association the skin traits examined here (genes in bold include the index SNP).

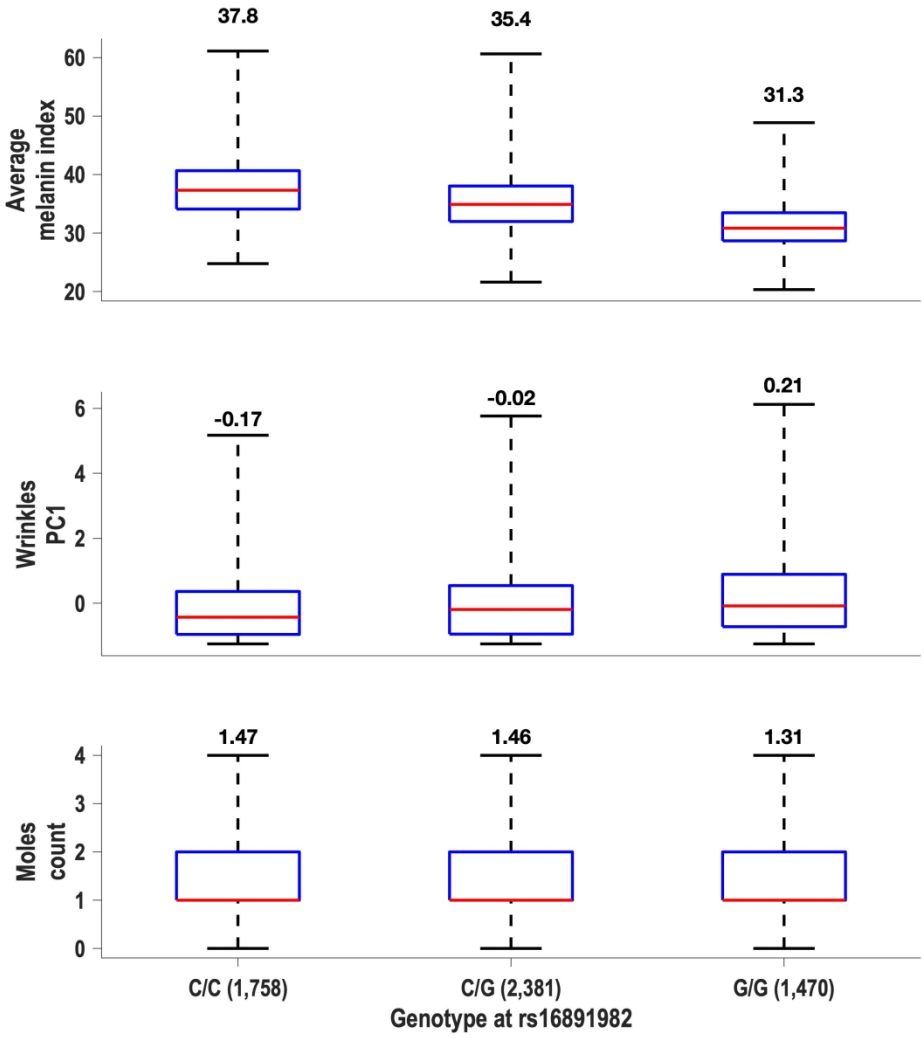
Chrom. region	Main candidate gene	Trait	Position (hg19)	Index SNP	Allele Tested	*p*-value	beta	# significant SNPs
1p13.3	***VAV3***	Wrinkles PC1	108 229 994	rs2504460	G	1×10^{-8}	0.13	2
1q32.3	*SLC30A1*	Mole count	211 712 690	rs6679498	C	2×10^{-8}	-0.11	3
5p13.2	***SLC45A2***	Mole count	33 956 560	rs34466007	G	9×10^{-11}	-0.13	11
		Wrinkles PC1	33 952 812	rs173662	C	3×10^{-9}	0.11	18
21q21.2	Intergenic	Forehead wrinkles	25 346 340	rs147991442	A	6×10^{-10}	0.20	17



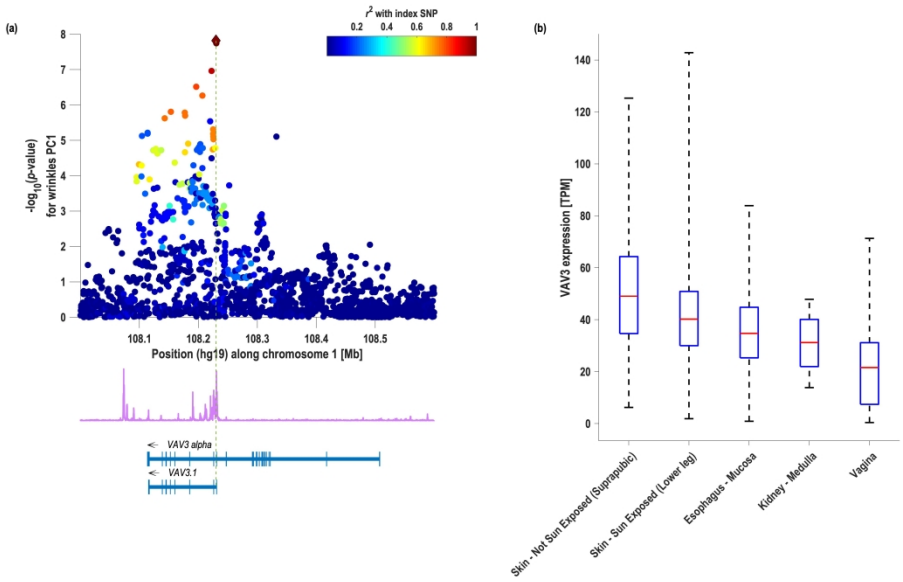
bjd_20436_f1.jpg



bjd_20436_f2.jpg



bjd_20436_f3.jpg

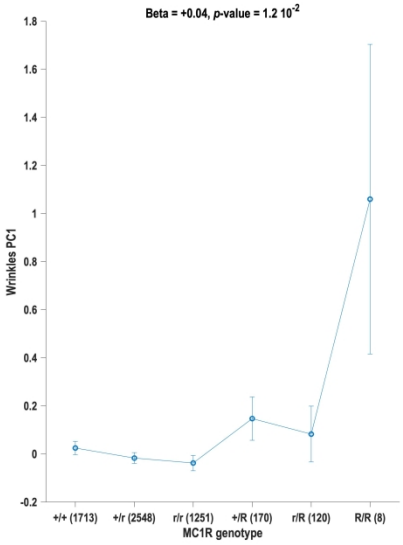


bjd_20436_f4.jpg

(a)

Gene	Variant	TA	F(TA)	*p*-value	Beta
IRF4	rs12203592	T	6.6%	$3\ 10^{-3}$	0.13
MC1R	r35063026	T	2.1%	$2\ 10^{-3}$	0.19
	rs4268748	C	11.6%	$3\ 10^{-1}$	0.03

(b)



bjd_20436_f5.jpg