

RESEARCH ARTICLE

# Integrative genetic analysis suggests that skin color modifies the genetic architecture of melanoma

Imge Hulur<sup>1#a</sup>, Andrew D. Skol<sup>2</sup>, Eric R. Gamazon<sup>2#b</sup>, Nancy J. Cox<sup>2#b</sup>, Kenan Onel<sup>3,4\*</sup>

**1** Committee on Genetics, Genomics and Systems Biology, The University of Chicago, Chicago, Illinois, United States of America, **2** Department of Medicine, The University of Chicago, Chicago, Illinois, United States of America, **3** Department of Pediatrics, Hofstra Northwell School of Medicine, Hempstead, New York, United States of America, **4** Department of Genetics and Genomics, The Feinstein Institute for Medical Research, Manhasset, New York, United States of America

<sup>#a</sup> Current address: Division of Dermatology, Department of Medicine, Albert Einstein College of Medicine, Bronx, New York, United States of America

<sup>#b</sup> Current address: Division of Genetic Medicine, Department of Medicine, Vanderbilt University, Nashville, Tennessee, United States of America

\* [konel@northwell.edu](mailto:konel@northwell.edu)



**OPEN ACCESS**

**Citation:** Hulur I, Skol AD, Gamazon ER, Cox NJ, Onel K (2017) Integrative genetic analysis suggests that skin color modifies the genetic architecture of melanoma. PLoS ONE 12(10): e0185730. <https://doi.org/10.1371/journal.pone.0185730>

**Editor:** Nathan A. Ellis, University of Arizona, UNITED STATES

**Received:** July 11, 2017

**Accepted:** September 18, 2017

**Published:** October 3, 2017

**Copyright:** © 2017 Hulur et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** Gene Environment Association Studies (GENEVA) melanoma project GWAS data are available from the Database of Genotypes and Phenotypes (dbGaP) (accession number phs000187.v1.p1). Expression quantitative trait loci data are available from the MuTHER resource (Multiple Tissue Human Expression Resource - <https://www.ebi.ac.uk/arrayexpress/experiments/E-TABM-1140>).

**Funding:** The author(s) received no specific funding for this work.

## Abstract

Melanoma is the deadliest form of skin cancer and presents a significant health care burden in many countries. In addition to ultraviolet radiation in sunlight, the main causal factor for melanoma, genetic factors also play an important role in melanoma susceptibility. Although genome-wide association studies have identified many single nucleotide polymorphisms associated with melanoma, little is known about the proportion of disease risk attributable to these loci and their distribution throughout the genome. Here, we investigated the genetic architecture of melanoma in 1,888 cases and 990 controls of European non-Hispanic ancestry. We estimated the overall narrow-sense heritability of melanoma to be 0.18 ( $P < 0.03$ ), indicating that genetics contributes significantly to the risk of sporadically-occurring melanoma. We then demonstrated that only a small proportion of this risk is attributable to known risk variants, suggesting that much remains unknown of the role of genetics in melanoma. To investigate further the genetic architecture of melanoma, we partitioned the heritability by chromosome, minor allele frequency, and functional annotations. We showed that common genetic variation contributes significantly to melanoma risk, with a risk model defined by a handful of genomic regions rather than many risk loci distributed throughout the genome. We also demonstrated that variants affecting gene expression in skin account for a significant proportion of the heritability, and are enriched among melanoma risk loci. Finally, by incorporating skin color into our analyses, we observed both a shift in significance for melanoma-associated loci and an enrichment of expression quantitative trait loci among melanoma susceptibility variants. These findings suggest that skin color may be an important modifier of melanoma risk. We speculate that incorporating skin color and other non-genetic factors into genetic studies may allow for an improved understanding of melanoma susceptibility and guide future investigations to identify melanoma risk genes.

**Competing interests:** The authors have declared that no competing interests exist.

## Introduction

With 76,380 new cases of invasive melanoma and 68,480 new cases of melanoma *in situ* expected in 2016 [1], melanoma incidence is rising more rapidly than that of any other cancer in the United States [2–5]. Although melanoma accounts for only 1% all skin cancer cases, it is responsible for the majority of skin cancer deaths [1]. In addition to its high mortality, melanoma is associated with enormous health care costs. The annual cost of productivity loss and direct medical care associated with melanoma in the United States is estimated at \$3.5 billion and \$932.5 million annually, respectively [6, 7].

Melanoma primarily affects individuals of European ancestry and is much less common among individuals of Asian, African and Hispanic ancestry [8]. The main environmental risk factor for melanoma in whites is ultraviolet radiation (UVR) from sunlight, which promotes cancer by multiple mechanisms including direct DNA damage [9, 10], immunosuppression [11, 12], and dysregulation of growth factors [13]. Another major risk factor for melanoma is light skin color, which is characterized by low levels of the UVR-absorbing pigment melanin [14]. Even in white individuals, differences in melanin are associated with differences in the amount of UVR to which melanocytes are exposed. Consequently, light-skinned whites are two to three times more likely to develop melanoma than are dark-skinned whites [15, 16].

Genetics also plays an important role in melanoma susceptibility [17]. Genome-wide association studies (GWAS) have identified several single nucleotide polymorphisms (SNPs) associated with melanoma risk, many of which influence pigmentation [18–25]. GWAS have also discovered melanoma susceptibility loci that are not associated with pigmentation; many of these are thought to mediate the response to UVR damage by regulating genes involved in DNA repair and cell cycle control [25, 26].

There have been a few studies that aim to characterize the proportion of melanoma risk attributable to genetics. The heritability of melanoma was estimated to be 55% in a study of Australian twins [27], and 58% in a study of Nordic twins [28]. More recently, a software tool called Genome-wide Complex Trait Analysis (GCTA), which allows for the estimation of heritability using genome-wide genotype data from unrelated individuals [29], was applied to quantify the contribution of genetic factors to melanoma susceptibility. In this study, heritability of melanoma was estimated to be 19% and 30% based on two different melanoma GWAS datasets from the USA and Australia, respectively [30].

In this study, we utilized an integrative approach, intersecting genome-wide genotype data from a melanoma GWAS with heritability analyses using expression quantitative trait loci (eQTL) and melanoma-associated covariates to gain insights into melanoma susceptibility, including the relative contributions of genetic factors and the underlying genetic architecture and functional mechanisms. For our analyses, we used the USA GWAS dataset mentioned in the previous paragraph, which had yielded the GCTA heritability estimate of 0.19 [30]. Here, we partitioned that heritability by chromosome, minor allele frequency (MAF) and functional annotation, including genic regions and skin *cis*-eQTLs, in an attempt to characterize the genomic distribution and allelic frequency spectrum of melanoma risk variants and to determine if any particular class of variation is driving melanoma risk. Given that pigmentation is a known melanoma risk factor, we incorporated skin color into our heritability analysis and performed a GWAS conditional on skin color. In addition to identifying risk loci that reached genome-wide significance, we also found evidence that these variants were enriched for skin *cis*-eQTLs, highlighting the importance of genetic regulation of gene expression in melanoma susceptibility.

Taken together, these results underscore the importance of incorporating the influence of non-genetic factors such as skin color into genetic analyses of complex traits such as melanoma, as they can significantly modify the underlying genetic architecture.

## Methods

### Melanoma data set

We obtained data from the Gene Environment Association Studies (GENEVA) melanoma project GWAS available through Database of Genotypes and Phenotypes (dbGaP) (accession number phs000187.v1.p1) [31]. Details of the case and control samples and genotyping have already been published [21]. Briefly, the dataset included 1,971 cases and 1,045 controls of European descent. All cases and controls were genotyped using the Illumina HumanOmni1-Quad (v1.0) array, and 1,012,904 satisfactorily genotyped SNPs were released to dbGaP. Skin color information was available for all controls and approximately half the cases ( $N = 1,050$ ). Skin color was self-reported, obtained via questionnaire in which subjects were asked to grade their skin color on a scale of 1 (very fair) to 10 (dark brown).

### SNP genotype quality control (QC)

**Standard QC.** We excluded SNPs with  $MAF < 0.02$ , missingness  $> 2\%$ , and Hardy Weinberg equilibrium (HWE) test  $P \leq 10^{-4}$ . Sex chromosomes and insertions/deletions (indels) were excluded from the analysis. We excluded samples with a genotype call rate  $\leq 95\%$ . SNP chromosomal positions were converted from hg18 to hg19 coordinates using liftOver (<http://genome.ucsc.edu/cgi-bin/hgLiftOver>). After QC, we were left with 775,275 SNPs genotyped in 1,968 cases of melanoma and 1,044 controls with a mean sample call rate of 99.93%.

**QC for heritability estimates.** We used more stringent QC criteria when filtering SNPs and individuals to estimate heritability, as recommended by Lee et al (2011) [32], in order to minimize artifactual differences between case and control allele frequencies, because not doing so can artificially inflate heritability estimates. We excluded SNPs with  $MAF < 0.01$ , missingness  $> 5\%$ , and HWE  $P < 0.05$ . We further excluded SNPs with  $P < 0.05$  for tests of differential missingness between cases and controls. We restricted our analysis to autosomal SNPs. We excluded cases with genotype call rate  $\leq 99\%$ , or an estimated genetic relatedness score (by GCTA) of  $> 0.05$ . After the stringent QC process, we were left with 1,888 cases and 990 controls, and 756,002 SNPs from which to estimate heritability.

### Principal components analysis (PCA)

We investigated the population structure of our dataset by performing a PCA-based analysis using EIGENSTRAT [33] and including the European (CEU), African (YRI) and Asian (ASN = CHB + JPT) HapMap samples (S1 Fig). 33 individuals were identified as having non-European ancestry and excluded from the association and heritability analyses. Upon the removal of population outliers, we repeated the PCA, with CEU, Tuscan (TSI) and Ashkenazi Jewish (AJ) samples included as references and the GWAS samples color-coded according to case-control status (S2 Fig). The majority of samples clustered with the CEU, while a small number clustered with the AJ or trailed towards the TSI, suggesting that our study population is mostly of Northern European ancestry. We conducted the PCA without the reference population samples and used the first two PCs from the analysis to adjust for confounding by population stratification in the association analyses (S3 Fig).

### Imputation

We imputed 1000 Genomes Phase I SNPs with IMPUTE2 v2.2.2 using the 1000 Genomes Phase I integrated variant set release (v3) [34, 35]. We converted imputation probabilities into genotypes for all SNPs with info scores  $> 0.30$  via thresholding, using a posterior genotype

probability threshold of 0.90, as implemented in GTOOL v0.7.5 (<http://www.well.ox.ac.uk/~cfreeman/software/gwas/gtool.html>). Imputed SNPs with MAF < 0.01, missingness rate > 5%, HWE  $P \leq 10^{-4}$  and all indels were excluded. Three samples with missingness rate > 5% for the imputed SNPs were removed, resulting in a final dataset of 6,311,517 SNPs in 1,952 cases and 1,025 controls.

## Genome-wide association study

We performed a melanoma GWAS using directly genotyped and imputed SNPs for 1,968 cases of melanoma and 1,044 controls. Specifically, we performed logistic regression on case-control status using the SNP genotype (additive effect), age, sex, and the first two PCs as covariates. (S3 Fig). We also performed a second GWAS in which skin color (coded as 1–10) was added as a quantitative covariate in the model.

## Heritability analysis of melanoma

We estimated the heritability of melanoma risk using GCTA v1.24 [29]. We first constructed a genetic relatedness matrix (GRM) that contains estimates of genetic relatedness for all pairs of individuals based on genotyped data after stringent QC. The additive genetic relationship between individuals  $j$  and  $k$  ( $A_{jk}$ ) is calculated by the following equation:

$$A_{jk} = \frac{1}{N} \sum_{i=1}^N \frac{(x_{ij} - 2p_i)(x_{ik} - 2p_i)}{2p_i(1 - p_i)}$$

where  $N$  is the number of SNPs,  $x_{ij}$  is the number of copies of the reference allele for the  $i^{\text{th}}$  SNP of the  $j^{\text{th}}$  individual and  $p_i$  is the frequency of the reference allele. GCTA then fits the GRM as random effect by a mixed linear model (MLM) to estimate the total additive variance explained by all SNPs ( $\sigma_g^2$ ). The MLM is represented by the following equations:

$$y = X\beta + g + \varepsilon \quad \text{with} \quad \text{var}(y) = V = A\sigma_g^2 + I\sigma_\varepsilon^2$$

where  $y$  is a vector of phenotypes,  $\beta$  is a vector of fixed effects such as age, sex or ancestry principal components (PC),  $g$  is a vector of total additive genetic effects,  $\varepsilon$  is a vector of residual effects,  $I$  is an identity matrix,  $A$  is the GRM and  $\sigma_\varepsilon^2$  is error variance.

Our strict QC measures substantially reduce the number of SNPs used for the GRM, which decreases linkage disequilibrium (LD) between the remaining SNPs and the true risk loci and leads to a downwardly biased estimate of heritability. We used the method described in Yang et al (2010) [36] to correct this bias.

PCs of the genotype data were calculated using GCTA. The top 20 PCs were included as quantitative covariates in the heritability analysis to account for any possible persistent population structure that might result in inflated heritability estimates. Case-control status was transformed to a liability scale via a probit model and fit as a mixed model that included the top 20 PCs as fixed effects and the additive genetic relatedness as a random effect. The additive genetic variance was estimated using restricted maximum likelihood (REML) in GCTA. A further adjustment to heritability is made to correct for the ascertainment bias that results due to the difference in the prevalence of melanoma in our sample (66%) relative to its population prevalence (2%).

## Heritability explained by known melanoma SNPs

We downloaded all SNPs previously identified in GWAS as associated with melanoma from the National Human Genome Research Institute (NHGRI) catalog (<http://www.ebi.ac.uk/gwas>),

using the default p-value threshold of  $1 \times 10^{-5}$  (S1 Table). Out of the 39 variants, 17 were directly genotyped in our dataset and 16 were imputed but not directly genotyped. 6 SNPs were not genotyped or imputed in our study, and there was also no LD information regarding them in either the HapMap or 1000 Genomes Project databases. Hence, these SNPs were excluded from our analysis. We generated the GRM using the remaining 33 GWAS SNPs for which we had direct or imputed genotype data. The heritability attributable to these melanoma-associated SNPs was then estimated via the REML analysis described above. In addition, we repeated the heritability analysis including all SNPs within 250 kb (up and downstream) of the 39 known melanoma-associated SNPs ( $N = 5,324$ ). In order to test the significance of the heritability estimate explained by known GWAS SNPs (or GWAS regions), we generated 1,000 permutations of the melanoma case-control status and reran GCTA on each permutation. The empirical p-value was estimated as the number of permutations in which the heritability exceeds the heritability estimated using the correct case-control assignments.

### Heritability partitioned by chromosome

We generated separate GRMs for each chromosome and fit them all simultaneously in a single joint analysis to calculate the heritability explained by each chromosome. By plotting the heritability of each chromosome by its length, we sought to see whether there was a linear relationship between chromosome length and heritability, as would be expected under a polygenic model, where thousands of risk variants are distributed evenly across the genome.

### Heritability partitioned by minor allele frequency

In order to characterize the allelic frequency spectrum of the SNPs contributing to melanoma risk and to test if there is a disproportionately high contribution to risk from a particular MAF class, we partitioned the genotyped SNPs into six non-overlapping MAF bins from 0.01–0.05, > 0.05–0.1, > 0.1–0.2, > 0.2–0.3, > 0.3–0.4, > 0.4–0.5 and calculated the GRM separately for each MAF bin. We fit them all simultaneously using GCTA to estimate the heritability attributable to each MAF bin.

### Rare variants model

We used the method described in Yang et al. (2010) [36] to calculate the heritability of melanoma under the assumption that causal variants have  $MAF \leq 0.1$ . Briefly, we estimated the GRM using only the SNPs with  $MAF \leq 0.1$  ( $G_{jk}$ ) and calibrated the prediction error by the regression of  $G_{jk}$  on the GRM estimated using all genotyped SNPs ( $A_{jk}$ ). Using the equation  $\beta = 1 - \frac{(c+1/N)}{\text{var}(A_{jk})}$ , where  $\beta$  is the regression coefficient and  $N$  is the number of SNPs used to calculate  $A_{jk}$ , we estimated  $c$ , the correction term for adjusting the GRM. We used the `—grm-adj` option in GCTA to adjust the GRM ( $A_{jk}$ ) with the estimated value of  $c$  ( $6.2 \times 10^{-6}$ ).

### Heritability partitioned by functional annotation

**Genic.** We partitioned the SNPs into genic and intergenic regions using ANNOVAR (hg19, refGene) annotations [37]. Genic variants included those located within exons, coding sites, 3'- and 5'- untranslated regions (UTRs), splice sites and within 1 kb of the transcription start or stop sites. A single GRM was created for each partition and the heritability explained by each was determined in a single joint analysis where both GRMs were fit together.

**Expression QTL.** In a second analysis, we partitioned variants into skin *cis*-eQTLs and non-eQTL SNPs based on data from the MuTHER study (Multiple Tissue Human Expression Resource— <https://www.ebi.ac.uk/arrayexpress/experiments/E-TABM-1140>) [38]. As 32% of

the SNPs that were used to estimate heritability in our study were not analyzed in the MuTHER study, we extracted SNPs that were common to both studies ( $N = 514,445$ ) and calculated the heritability attributable to this subset of SNPs only. A single GRM was created for each partition and the heritability explained by each was determined in a single joint model where both GRMs were fit together. The analysis was repeated at three different *cis*-eQTL significance thresholds (0.01, 0.001, 0.0001).

### eQTL enrichment analysis

We conducted a test to determine if melanoma-associated SNPs were enriched for eQTLs using a permutation-based analysis described previously by Gamazon et al (2013) [39]. Specifically, we performed 1,000 GWAS under the null hypothesis of no association by shuffling case-control status while keeping the genotype data fixed to preserve the LD structure. In each permuted dataset, we determined the number of skin *cis*-eQTLs (defined as either  $P < 0.01$  or  $P < 0.001$ ) among the top 1,000 and 2,000 most significant GWAS SNPs. The empirical *p*-value for eQTL enrichment among the melanoma-associated SNPs is the proportion of permutations with more skin *cis*-eQTL observed in the top melanoma-associated SNPs than in our original GWAS. We conducted the enrichment analysis twice, once with and once without using skin color as a covariate in the GWAS.

## Results

### Association analysis

Following QC and imputation, we had genotype data on 6,311,517 SNPs in 1,952 cases and 1,025 controls of European descent from the GENEVA melanoma project. Using logistic regression, we successfully replicated the findings previously reported for this dataset [21]. These findings included SNPs in the *MC1R* region on chromosome 16 which surpassed the threshold for genome-wide significance ( $P = 1.4 \times 10^{-9}$  for the index SNP rs74800773) in both our analysis and previously published analysis; rs1129038 tagging the *HERC2/OCA2* region on chromosome 15 ( $P = 3.1 \times 10^{-7}$ ) which reached genome-wide significance in previously published analysis but only suggestive levels of significance in our analysis; and rs1889680, tagging the *CDKN2A/MTAP* region on chromosome 9 which reached suggestive levels of significance ( $P = 3.0 \times 10^{-7}$ ) in both analyses (S4 Fig) [21]. The association of *HERC2/OCA2* and *CDKN2A/MTAP* loci with melanoma risk were confirmed in subsequent meta-analyses [25, 40].

### Heritability analysis of melanoma

To estimate the narrow-sense heritability ( $h^2$ ) of melanoma, we performed further QC and used the REML approach found in GCTA in 1,888 cases and 990 controls. After adjustment for imperfect LD between causal and genotyped SNPs and correction for ascertainment bias assuming a population prevalence (K) of 2% (<https://seer.cancer.gov>), we estimated the heritability of melanoma to be 0.18 ( $P = 0.03$ ), closely approximating the previously published value of 0.19, and indicating that genetics contributes significantly to melanoma etiology [30].

### Heritability explained by known melanoma variants

To determine the proportion of the heritability accounted for by known melanoma-associated variants, we used the 39 SNPs in the NHGRI catalog with association *P*-value  $< 10^{-5}$  (S1 Table). Six of the 39 SNPs were excluded from our analysis since they were not genotyped directly and there was no LD information regarding these SNPs in either HapMap or 1000 Genomes Project databases. We found that the remaining 33 SNPs explained about 4% (95%



CI = 0.02–0.06) of the variability in melanoma and 22% of the heritability explained by all common SNPs, both known and unknown. As the majority of these SNPs are likely tagging SNPs for causal variants located in their vicinity, we expanded our analysis to include SNPs located within 250 kb (upstream and downstream) of the 39 SNPs. Heritability attributable to these GWAS regions was estimated to be 0.05 (95% CI = 0.03–0.07), indicating that most of the heritability of melanoma could not be explained by already-identified variants.

### Partitioning of heritability by chromosome

It is generally assumed that common variants contribute to complex diseases under a polygenic model, in which a large number of variants distributed throughout the genome each contribute only a small amount to disease risk. A prediction of this model is that heritability should be proportional to chromosome size. Therefore, we estimated the heritability explained by each chromosome (S2 Table), and investigated the strength of the linear correlation between chromosome's physical length and heritability, using linear regression. We did not observe a linear relationship between chromosome length and heritability ( $r = 0.08$ ,  $P = 0.71$ ).

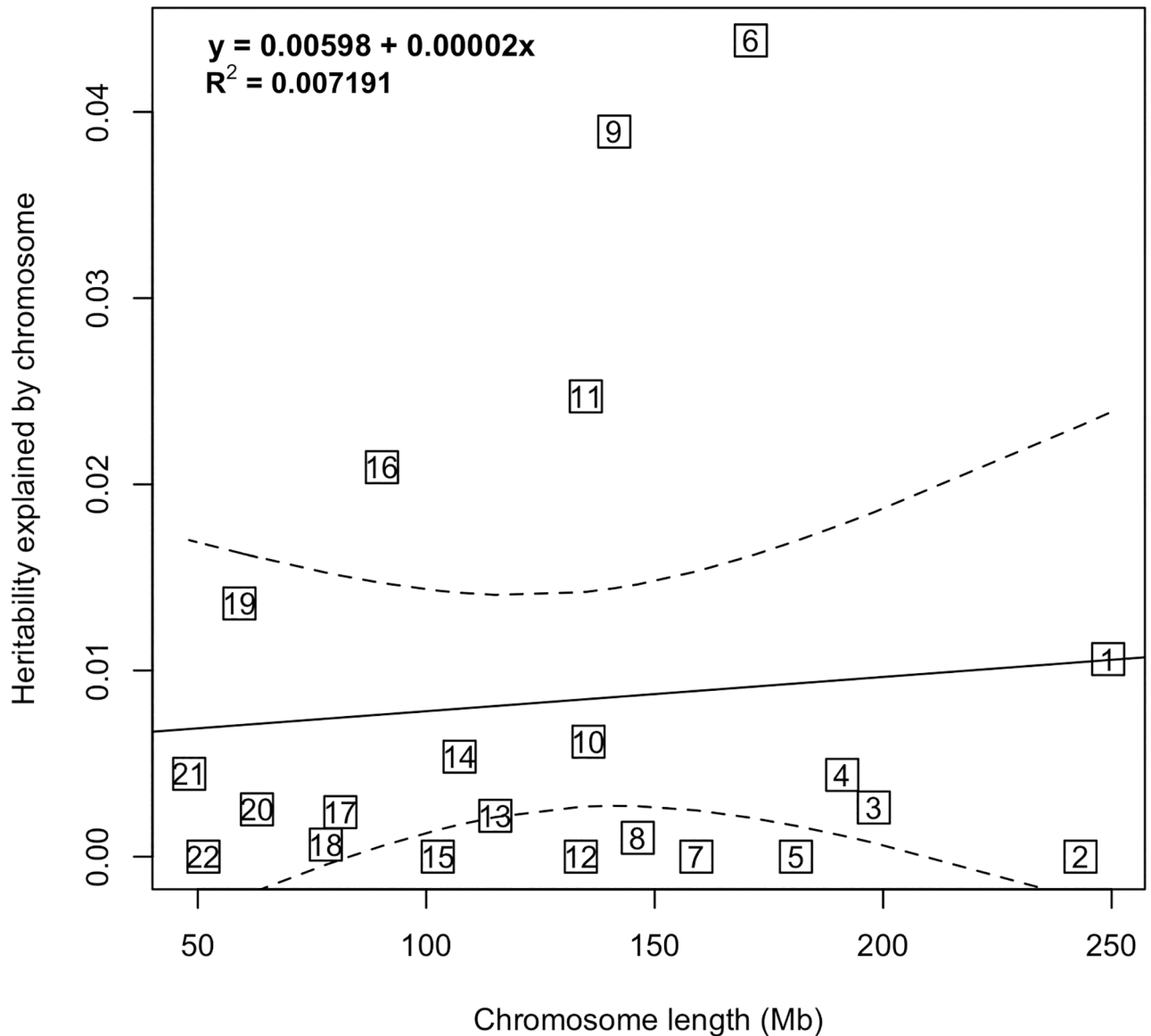
Chromosomes 6, 9, 11, 16 fell outside the 95% confidence bounds for the regression line, suggesting either that there is an excess of risk variants on these chromosomes, or that the risk variants on these chromosomes contribute disproportionately to melanoma risk (Fig 1). Chromosomes 9 and 16 harbor the known melanoma susceptibility loci, *CDKN2A* and *MC1R*, respectively. Therefore, we conditioned on the most significant SNP in each locus and re-ran the heritability analysis. Chromosome 9's heritability decreased from 0.04 to 0.02 after fitting the most significant SNP at chromosome 9 (rs2383202), suggesting the existence of other risk loci on Chromosome 9. When we removed the influence of rs4408545, the most significantly associated SNP on chromosome 16, the estimated heritability for this chromosome fell from 0.02 to 0.008, suggesting that virtually all of the heritability attributable to this chromosome was due to the *MC1R* locus.

For chromosome 6, we hypothesized that the human leukocyte antigen (HLA) region on chromosome 6, which contains essential immune system genes, may be responsible for the high heritability attributed to chromosome 6. After excluding the HLA region from the heritability analysis, the heritability estimate fell by half, from 0.04 to 0.02. For chromosome 11, although there are multiple studies pointing to the existence of melanoma tumor-suppressor genes on this chromosome [41–43], no candidate gene has been identified. Consequently, we were not able to explore the contribution of this chromosome to the heritability of melanoma heritability further.

To determine whether risk SNPs throughout the remainder of the genome fit a polygenic model, we removed HLA and conditioned on rs2383202 and rs4408545 and again found there was not a significant correlation between chromosome length and heritability ( $r = 0.18$ ,  $P = 0.42$ ). Even when we removed chromosomes 6, 9, 11, and 16 completely from the analysis, there was no correlation ( $r = -0.01$ ,  $P = 0.96$ ). Taken together, these data could mean that the contribution of common variants to melanoma susceptibility is not consistent with a polygenic model. Alternatively, it is also possible that the lack of correlation between chromosome length and heritability is due to the large error associated with each heritability estimate when performed on a per chromosome basis.

### Partitioning of heritability by MAF and rare variants model

In order to characterize the allelic frequency spectrum of the genomic variation contributing to melanoma risk and test whether rare variants play a disproportionate role in melanoma risk, we partitioned the SNPs into MAF bins and estimated the heritability to attributable to



**Fig 1. Melanoma heritability stratified by chromosome.** Heritability estimates for each chromosome are plotted against chromosome length. The solid line is the regression of heritability on chromosome length and the dashed lines represent the 95% confidence intervals for the slope of the regression line. Chromosomes 6, 9, 11 and 16 are outside the 95% confidence interval and have higher heritability estimates than that expected based on chromosome length.

<https://doi.org/10.1371/journal.pone.0185730.g001>

each bin (S5 Fig, S3 Table). There were no significant differences in heritability estimates between bins.

Additionally, to explore the contribution of rare variants to melanoma heritability, we adjusted the heritability estimate under a rare variants model using  $MAF \leq 0.01$  as proxies for causal variants. The heritability of melanoma under the rare variants model was estimated to be 0.25 (SE = 0.13,  $P = 0.03$ ), not significantly different than our initial heritability estimate.

### Partitioning of heritability by function annotation

We partitioned the heritability of melanoma by functional annotation (genic or skin *cis*-eQTLs) to determine if variants in either category explain a greater proportion of heritability



**Table 1. Heritability of melanoma partitioned based on genic annotation.**

Partition	Number of SNPs	% of Total SNPs	Heritability (SE)	% of Total Heritability
Genic	369,588	49%	0.11 (0.07)	67%
Intergenic	386,414	51%	0.06 (0.08)	33%

The number of SNPs (proportion of total SNPs), heritability (proportion of total heritability) and standard errors are listed for each partition.

<https://doi.org/10.1371/journal.pone.0185730.t001>

than would be expected based on their frequency on the GWAS array. While genic SNPs comprised 49% of all interrogated SNPs, they accounted for 67% of the total melanoma heritability ( $h^2 = 0.11$ , SE = 0.07; Table 1).

Skin *cis*-eQTLs also accounted for a greater proportion of heritability than expected. As eQTL annotation was available for only 68% of all genotyped SNPs (N = 514,445), we first calculated the heritability attributable to this subset of variants ( $h^2 = 0.11$ , SE = 0.08). When using eQTL p-value thresholds of 0.01, 0.001, and 0.0001, approximately 23%, 10%, and 6% of all genotyped SNPs were considered *cis*-eQTLs (S4 Table). When estimating heritability using only these eQTLs, we found that they accounted for 54% ( $h^2 = 0.06$ , SE = 0.05), 31% ( $h^2 = 0.03$ , se = 0.03) and 26% ( $h^2 = 0.03$ , SE = 0.02), respectively, of the total heritability attributable to variants for which eQTL annotation was available.

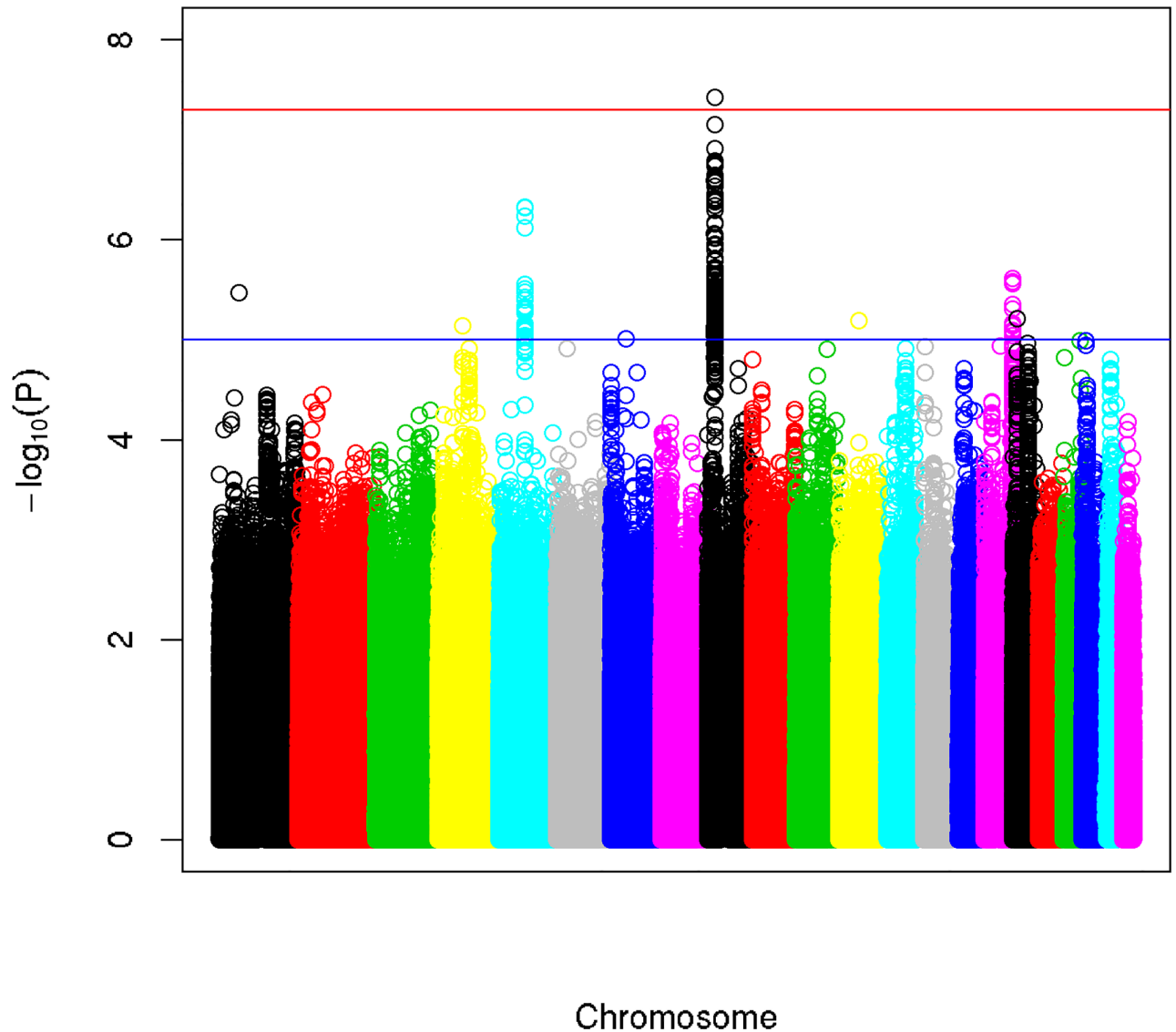
### The effect of skin color on melanoma heritability

As skin color is a modifier of biologically effective UVR exposure and melanoma susceptibility, we hypothesized that by accounting for skin color in the GWAS, we could reduce non-genetic heterogeneity and improve power to detect associations. We investigated this by performing the GWAS separately in light-skinned (skin color 1–3; 613 cases and 442 controls) and dark-skinned (skin color 4–10; 402 cases and 548 controls) whites but detected no chromosomal regions that reached genome-wide significance in either subset, likely because skin color was unavailable for significant number of cases, thereby reducing significantly the power of each skin-color-specific GWAS (data not shown).

Next, we repeated the GWAS in all cases and controls using skin color as a covariate. In doing so, the melanoma risk locus on chromosome 9 upstream of the *MTAP/CDKN2A* region reached genome-wide significance (without skin color as a covariate: per-allele odds ratio (OR) = 1.32,  $P = 3.0 \times 10^{-7}$ ; with skin color as a covariate: per-allele OR = 1.42,  $P = 3.8 \times 10^{-8}$ ) (Fig 2). In contrast, the *MC1R* (chromosome 16) and *HERC2/OCA2* (chromosome 15) loci, both of which affect pigmentation, lose their significance, implying that they influence melanoma risk through their effects on pigmentation phenotypes.

### eQTL enrichment analysis

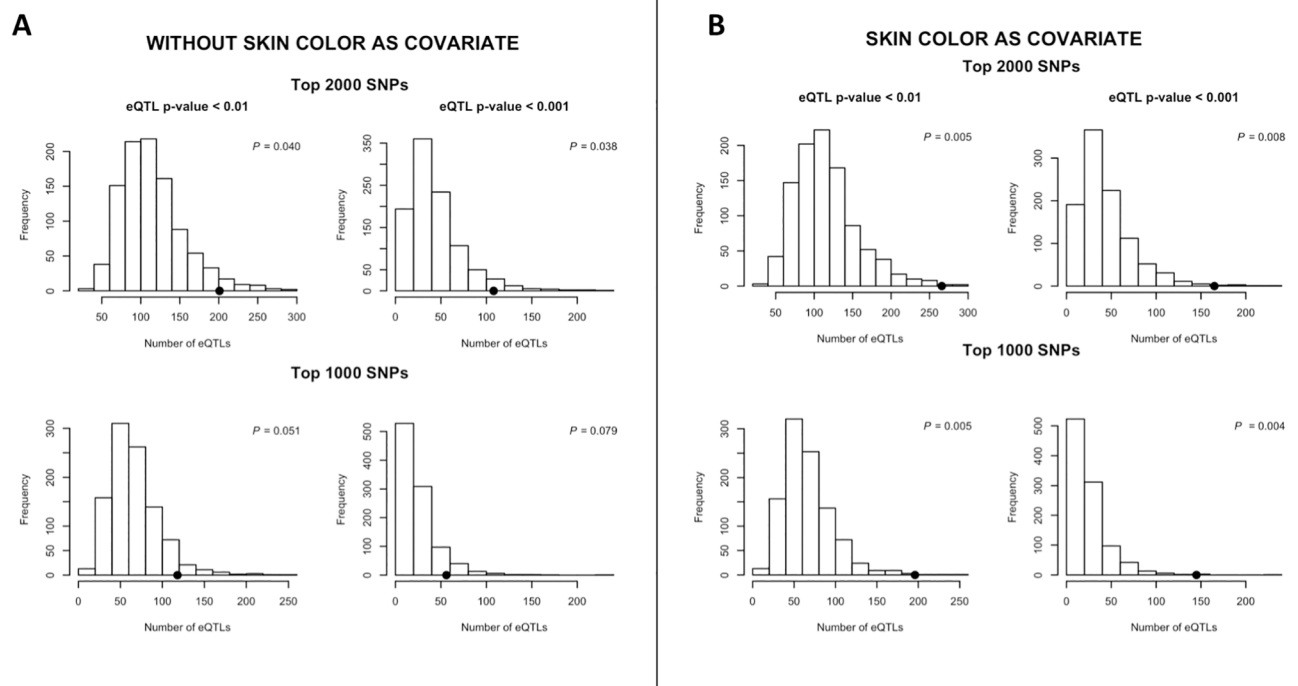
Our results suggest both that functional SNPs are enriched among melanoma-associated variants, and that skin color may be an important determinant of the genetic architecture of melanoma susceptibility. If so, then we hypothesized that incorporating skin color as a covariate in our analysis should improve the power of our study, resulting in an enrichment of functional SNPs among the set of variants most associated with melanoma risk. To test this, we investigated the consequence of conditioning on skin color on the association between melanoma risk and skin *cis*-eQTLs. By permutation analysis, we found a statistically significant enrichment of skin *cis*-eQTLs among the 2,000 most significant SNPs when not conditioning on skin color at two different p-value thresholds for *cis*-eQTLs ( $P_{.01} = 0.040$  and  $P_{.001} = 0.038$ ) (Fig 3A). When skin color was included as a covariate, however, enrichment of *cis*-eQTLs among



**Fig 2. The *CDKN2A* region surpasses the genome-wide significance threshold whereas the *MC1R* and *HERC2/OCA2* loci do not upon conditioning on skin color in the melanoma etiology GWAS.** Manhattan plot of the p-values for the association between imputed SNPs and melanoma using skin color as a covariate. The x-axis shows the chromosomal positions whereas the y-axis shows the  $-\log_{10}$  p-values of the SNPs. The p-values were obtained by logistic regression analysis including age, sex, skin color and the first two PCs from the PCA of GWAS (S3 Fig) as covariates. The red horizontal line is the widely used genome-wide significance threshold ( $p = 5 \times 10^{-8}$ ) that was estimated by correcting independent common variants, which is roughly 1,000,000. The blue horizontal line is the suggestive significance threshold ( $p = 1 \times 10^{-5}$ ). The *CDKN2A* region on chromosome 9 (black) reaches genome-wide significance upon the inclusion of skin color as a covariate in the analysis while the pigmentation loci *MC1R* (chromosome 16) and *HERC2/OCA2* (chromosome 15) are no longer significant.

<https://doi.org/10.1371/journal.pone.0185730.g002>

the top melanoma-associations became strikingly more significant ( $P_{.01} = 0.005$ ,  $P_{.001} = 0.008$ ) (Fig 3B). Results were similar when we looked only at the top 1,000 SNPs ( $P_{.01} = 0.051$  and  $P_{.001} = 0.079$ ; conditional analysis:  $P_{.01} = 0.005$ ,  $P_{.001} = 0.004$ ). Thus, our data demonstrate that by accounting for skin color, we improve significantly the power of our study to detect an enrichment of functional SNPs among those SNPs most associated with melanoma.



**Fig 3. eQTL enrichment in the melanoma GWAS dataset becomes more significant upon conditioning on skin color.** The distributions of the number of skin cis-eQTLs (at two different p-value thresholds:  $p < 0.01$  and  $p < 0.001$ ) in the permuted 1,000 GWAS datasets, generated by shuffling case-control status each time, are shown in the histograms. In each permuted dataset, the number of eQTLs are determined among the top 2,000 (top panel) or top 1,000 (bottom panel) associations. The solid black circle is the actual eQTL count among the melanoma-associated SNPs. The p-values correspond to empirical p-values, calculated as the proportion of permuted GWAS datasets in which the eQTL count exceeds the actual observed count among the top associations. **Panel (A)** displays results of the analysis without conditioning on skin color. Enrichment of eQTLs among the top 1,000 melanoma associated SNPs is statistically significant ( $P < 0.05$ ), while enrichment among the top 2,000 SNPs fails to reach statistical significance but is highly suggestive. **Panel (B)** displays results of the analysis in which skin color is included as a covariate in the analysis. Upon conditioning on skin color, enrichment of eQTLs among the top melanoma-associated SNPs becomes more pronounced and is statistically significant at all thresholds tested.

<https://doi.org/10.1371/journal.pone.0185730.g003>

## Discussion

This goal of this study was to investigate the genetic landscape of melanoma susceptibility and dissect the contributions of genetic and environmental variation by integrating GWAS data with pigmentation and expression quantitative trait loci data. We used the GCTA software tool for our analyses, which has previously been applied to study the heritability of other complex diseases including schizophrenia [44], Tourette Syndrome, obsessive-compulsive disorder [45] and multiple myeloma [46]. Although the overall heritability of the melanoma dataset we investigated had previously been estimated to be 0.19 using GCTA [30], little is known of how the genetic contribution to risk is distributed throughout the genome. Thus, by bridging genetics and biology, our study is the first to yield insights into the component parts of the genetic architecture of melanoma risk.

We found that rare variants were not disproportionately represented among melanoma-associated variants, suggesting that post-GWAS studies relying on common variation will continue to be informative for melanoma susceptibility. We then demonstrated that a large proportion of the heritability of melanoma remains unknown. The heritability explained by known GWAS SNPs for melanoma risk ( $h^2 = 0.04$ ) or the regions they reside in (including SNPs within 250 kb of the GWAS SNPs;  $h^2 = 0.05$ ) was only a small fraction of the heritability

explained by all common SNPs ( $h^2 = 0.18$ ), suggesting that more common melanoma risk loci are yet to be discovered.

Having shown the significant contribution of common variants to the heritability of melanoma, we went on to examine their genomic distribution by partitioning the heritability by chromosome. We found that melanoma susceptibility is consistent with an oligogenic model, in which a few genomic regions having relatively large effects, rather than many risk loci of very small effect sizes distributed uniformly throughout the genome, as is characteristic of highly polygenic traits [47].

We found that the high heritabilities of chromosomes 9 and 16, which explain a significantly greater proportion of the variance than expected based on their size, were mostly attributable to two regions of known association: the *CDKN2A/MTAP* region on chromosome 9 and *MC1R* region on chromosome 16. Chromosome 6 explained the largest amount of heritability, and its contribution was mostly attributable to the HLA locus which contains a large number of immune-related genes [48]. It is possible that certain HLA haplotypes are associated with melanoma risk by virtue of allowing early stage melanoma to evade the immune system. In fact, loss of class I expression has been frequently documented as an immune escape mechanism in melanomas which would go along with this hypothesis [49]. Although there are currently no known GWAS variants associated with melanoma in the HLA region, it is possible that previous GWAS were underpowered to detect these associations. One possible explanation is that there are many variants with small effect sizes within the HLA locus, such that although the overall contribution of the HLA locus to melanoma heritability is high, larger studies with higher power are needed to identify individual variants within the locus. Our finding that much of melanoma heritability is explained by HLA may be of particular interest in light of the fact that the use of immunotherapy to treat cancer was pioneered in melanoma. By stratifying patients into immune responders vs. non-responders, it may be possible to identify biomarkers in HLA that can be used clinically to aid in treatment decision-making for melanoma and other cancer patients.

Rare or low-frequency variants that are poorly captured by genotyping arrays are another potential source of missing heritability [50]. We investigated the hypothesis that melanoma heritability is mostly attributable to rare variants by partitioning the heritability by MAF as well as estimating the heritability under the assumption that causal variants have  $MAF \leq 0.01$ . However, we found no significant differences between the heritabilities captured by SNPs in different MAF bins and no substantial increase in the heritability estimate under the rare variants model. Taken together, our data provides no support for the hypothesis that rare or low-frequency variants contribute substantially to missing heritability.

The majority of variants associated with complex diseases are in non-coding regions and are thought to act by influencing gene expression [51–53]. Hence, variants identified by GWAS as reproducibly associated with a number of complex traits are enriched for eQTLs [51, 54–57]. Consistent with this observation, we detected an enrichment of skin *cis*-eQTLs among top associations in melanoma, providing strong evidence that the genetic regulation of gene expression in the skin is important in melanoma susceptibility. This has important implications for GWAS, as gene expression in the biologically relevant tissue is a potentially powerful intermediate phenotype that can help the functional characterization of GWAS variants.

Finally, by incorporating skin color, a known risk factor for melanoma, in our analysis of a European-only population, we significantly potentiated the enrichment of eQTLs among melanoma susceptibility loci. We also uncovered a significant association in the *CDKN2A/MTAP* region, previously reported in two separate GWAS of melanoma in Caucasians [19, 23]. This region contains the *CDKN2A* gene, which has a well-characterized role in cell cycle regulation [58, 59] and is responsible for a substantial proportion of familial melanomas [60, 61].

Furthermore, we found that the *MC1R* and *HERC2/OCA2* loci were no longer significant when conditioning on skin color. Several studies have previously shown the presence of variants within the *MC1R* locus that may have effects on melanoma risk independent of their effect on pigmentation [62, 63]. Although none of the *MC1R* variants remained significant after conditioning on skin color in our study, it is possible that our study was underpowered to detect the pigmentation-independent effects.

Our finding that the *CDKN2A/MTAP* region becomes significant after accounting for skin color in the model suggests that the genetics underlying melanoma risk may differ between light- and dark-skinned Caucasian populations such that the inclusion of skin color in the model allows the discovery of loci which might affect melanoma risk differently in the two subsets. Genetic studies comparing melanoma susceptibility in Northern European and darker-skinned Greek populations provide support for this hypothesis. One study reported that 15% of all melanoma cases in Greece have germline mutations in *CDKN2A* [64]. This was much higher than the prevalence of mutations in this gene in an Australian population of Northern European ancestry (0.2%) [65], suggesting that the genetic architecture of melanoma susceptibility may differ between light- and dark-skinned Caucasian populations. The hypothesis that differences in skin color may represent underlying population structure with different genetic architectures that is unaccounted for by solely pigmentation phenotypes, would also explain why the effect of the variation in pigmentation genes does not similarly affect the penetrance of the *CDKN2A/MTAP* region.

A limitation our study is that the skin color data was self-reported based on a questionnaire. However, when we ran the GWAS using self-reported skin color as the phenotype, we found that the most significant loci were rs4268748 on chromosome 16 near *MC1R* ( $P = 1.76 \times 10^{-14}$ ), rs12913832 on chromosome 15 near *OCA2* ( $P = 6.08 \times 10^{-11}$ ) and rs12203592 on chromosome 6 near *IRF4* ( $P = 5.79 \times 10^{-9}$ ), which have all been previously shown to be associated with skin color in Europeans [66, 67]. This suggests that self-reported skin color is in fact a reliable estimator and that the potential bias introduced into the study by using this subjective measure, should be minimal. A second limitation of our study is that skin color was only available for approximately half of the cases and all of the controls. This further reduced our sample sizes and thus power, when we stratified the dataset into subsets based on skin color, running the analyses separately in each subset. Another limitation of our study is that although skin color is an important modifier of melanoma risk, there are other factors such as number of nevi [68] and sun exposure [69, 70] that have also shown to be important modifiers of melanoma risk. Inclusion of these parameters in the model would allow us to investigate the full spectrum of environmental contributions to melanoma risk, on the other hand, this would add additional layers of subjectivity and complexity making the data harder to interpret.

Our conclusion that there may be differences in genetic susceptibility to melanoma risk between light- and dark-skinned individuals only applies to Caucasian populations as melanoma in individuals of African, Hispanic and Asian ancestry may have a different, perhaps not UV-related, etiology in these populations, often appearing in sun-protected atypical locations such as acral, subungual and mucosal skin [71, 72]. Although for melanoma, there appears to be an established relationship between skin color and melanoma risk, for most other complex diseases, the effect of environment on phenotype is too complex to be accounted for. This is further compounded by the fact that it is hard to objectively quantify most environmental exposures, further limiting the applicability of our approach to the study of other complex traits.

The next step in advancing our understanding of melanoma risk and heritability would be to perform the GWAS and heritability analyses separately in light- and dark-skinned individuals. We attempted these analyses with our dataset, however our sample size was too small to

reveal any significant hits or yield reliable heritability estimates in the light- and dark-skinned subsets. These analyses should be attempted in a larger cohort to see if the heritability of melanoma differs between light- and dark-skinned Caucasians and to identify variants that are specific to each skin color group.

Taken together, these results have important implications for future studies of melanoma susceptibility, as they suggest that skin color should be accounted for as a covariate and that the genetic architecture of melanoma risk differs between light-skinned and dark-skinned whites. More broadly, our strategy of integrating known non-genetic risk factors such as skin color into heritability analyses may have important implications for genomic investigations into the genetic architecture of other complex diseases as well. We speculate that this approach may facilitate the identification of patient subsets in whom the biology of risk differs significantly, thereby providing a more nuanced understanding of the role of genetics in human disease.

## Supporting information

**S1 Fig. PCA plot of the melanoma GWAS dataset.** PCA plot of the melanoma GWAS samples along with HapMap reference populations of Northern European (CEU), African (YRI) and Asian (ASN = CHB + JPT) ancestry. The samples used in PCA are color-coded as indicated in the legend on the right. While most GWAS samples cluster tightly with the HapMap CEU population, a small number tend toward YRI or ASN, indicating small amounts of non-European ancestry. The individuals who fall outside the cluster boundaries, as indicated by the black circle, were defined population outliers.

(TIF)

**S2 Fig. PCA plot of the melanoma GWAS samples showing within Europe ancestry.** PCA plot of the melanoma GWAS samples along with reference populations of Northern European (CEU), Ashkenazi Jewish (AJ) and Tuscan (TSI) ancestry. GWAS are color-coded according to case-control status (A), and skin color (B), as indicated in the legend for each panel.

(TIF)

**S3 Fig. PCA plot of the melanoma GWAS samples after removal of population outliers.** PCA plot of the melanoma GWAS samples after the removal of population outliers with non-European ancestry (individuals that are outside the black circle in S1 Fig). Cases and controls are color-coded as indicated in the legend on the right. There are no significant differences in PC1 or PC2 between cases and controls. The first two PCs from the plot are used as covariates in the association analyses of melanoma to correct for population structure.

(TIF)

**S4 Fig. Manhattan plot of the melanoma GWAS.** Manhattan plot of the p-values for the association between imputed SNPs and melanoma. The x-axis shows the chromosomal positions whereas the y-axis shows the  $-\log_{10}$  p-values of the SNPs. The p-values were obtained by logistic regression analysis including age, sex and the first two PCs from the PCA of GWAS as covariates (S3 Fig). The red horizontal line is the widely used genome-wide significance threshold ( $p = 5 \times 10^{-8}$ ) that was estimated by correcting independent common variants, which is roughly 1,000,000. The blue line is the suggestive significance threshold ( $p = 1 \times 10^{-5}$ ). The *MC1R* region on chromosome 16 (magenta) is significantly associated with melanoma risk, whereas the *CDKN2A* region on chromosome 9 (black) and the *HERC2/OCA2* region on (dark blue) chromosome 15 reach suggestive significance.

(TIF)



**S5 Fig. Melanoma heritability partitioned by minor allele frequency.** The x-axis represents the MAF bins while the y-axis represents the heritability attributed to SNPs in the corresponding MAF bins. The standard errors of the heritability estimates are represented by the error bars.

(TIF)

**S1 Table. Melanoma susceptibility SNPs from the NHGRI GWAS catalog.** Chromosomal regions, reported genes, risk allele frequencies, odds ratios (OR) or beta-coefficients and study references are listed for all melanoma-associated SNPs downloaded from the NHGRI GWAS Catalog.

(DOCX)

**S2 Table. Heritability of melanoma partitioned by chromosome.** Heritability estimates and standard errors (SE) are listed for each chromosome.

(DOCX)

**S3 Table. Heritability of melanoma partitioned by minor allele frequency (MAF).** The number of SNPs (proportion of total SNPs), heritability (proportion of total heritability) and standard errors (SE) are listed for SNPs in each MAF bin.

(DOCX)

**S4 Table. Heritability of melanoma partitioned based on skin *cis*-eQTL annotation at three different p-value thresholds.** The number of SNPs (proportion of total SNPs), heritability (proportion of total heritability) and standard errors (SE) are listed for each partition.

(DOCX)

## Author Contributions

**Conceptualization:** Imge Hulur, Andrew D. Skol, Eric R. Gamazon, Kenan Onel.

**Formal analysis:** Imge Hulur, Andrew D. Skol, Eric R. Gamazon, Kenan Onel.

**Investigation:** Imge Hulur, Andrew D. Skol, Eric R. Gamazon, Kenan Onel.

**Methodology:** Imge Hulur, Andrew D. Skol, Eric R. Gamazon, Kenan Onel.

**Resources:** Andrew D. Skol, Kenan Onel.

**Supervision:** Andrew D. Skol, Eric R. Gamazon, Nancy J. Cox, Kenan Onel.

**Writing – original draft:** Imge Hulur, Kenan Onel.

**Writing – review & editing:** Imge Hulur, Andrew D. Skol, Eric R. Gamazon, Nancy J. Cox, Kenan Onel.

## References

1. American Cancer Society. Cancer Facts & Figures 2016. Atlanta: American Cancer Society; 2016.
2. Geller AC, Miller DR, Annas GD, Demierre MF, Gilchrist BA, Koh HK. Melanoma incidence and mortality among US whites, 1969–1999. *JAMA: the journal of the American Medical Association*. 2002; 288(14):1719–20. PMID: [12365954](https://pubmed.ncbi.nlm.nih.gov/12365954/).
3. Geller AC, Swetter SM, Brooks K, Demierre MF, Yaroch AL. Screening, early detection, and trends for melanoma: current status (2000–2006) and future directions. *Journal of the American Academy of Dermatology*. 2007; 57(4):555–72; quiz 73–6. <https://doi.org/10.1016/j.jaad.2007.06.032> PMID: [17870429](https://pubmed.ncbi.nlm.nih.gov/17870429/).
4. Hall HI, Miller DR, Rogers JD, Bewerse B. Update on the incidence and mortality from melanoma in the United States. *Journal of the American Academy of Dermatology*. 1999; 40(1):35–42. PMID: [9922010](https://pubmed.ncbi.nlm.nih.gov/9922010/).

5. Linos E, Swetter SM, Cockburn MG, Colditz GA, Clarke CA. Increasing burden of melanoma in the United States. *The Journal of investigative dermatology*. 2009; 129(7):1666–74. <https://doi.org/10.1038/jid.2008.423> PMID: 19131946
6. Ekwueme DU, Guy GP Jr., Li C, Rim SH, Parelkar P, Chen SC. The health burden and economic costs of cutaneous melanoma mortality by race/ethnicity-United States, 2000 to 2006. *Journal of the American Academy of Dermatology*. 2011; 65(5 Suppl 1):S133–43. PMID: 22018062.
7. Guy GP Jr., Ekwueme DU, Tangka FK, Richardson LC. Melanoma treatment costs: a systematic review of the literature, 1990–2011. *American journal of preventive medicine*. 2012; 43(5):537–45. <https://doi.org/10.1016/j.amepre.2012.07.031> PMID: 23079178.
8. Cress RD, Holly EA. Incidence of cutaneous melanoma among non-Hispanic whites, Hispanics, Asians, and blacks: an analysis of California cancer registry data, 1988–93. *Cancer causes & control: CCC*. 1997; 8(2):246–52. PMID: 9134249.
9. Cadet J, Anselmino C, Douki T, Voituriez L. Photochemistry of nucleic acids in cells. *Journal of photochemistry and photobiology B, Biology*. 1992; 15(4):277–98. PMID: 1432396.
10. Tornaletti S, Pfeifer GP. UV damage and repair mechanisms in mammalian cells. *BioEssays: news and reviews in molecular, cellular and developmental biology*. 1996; 18(3):221–8. <https://doi.org/10.1002/bies.950180309> PMID: 8867736.
11. Noonan FP, Kripke ML, Pedersen GM, Greene MI. Suppression of contact hypersensitivity in mice by ultraviolet irradiation is associated with defective antigen presentation. *Immunology*. 1981; 43(3):527–33. PMID: 7251064
12. Molendijk A, van Gurp RJ, Donselaar IG, Benner R. Suppression of delayed-type hypersensitivity to histocompatibility antigens by ultraviolet radiation. *Immunology*. 1987; 62(2):299–305. PMID: 2960612
13. Berking C, Takemoto R, Satyamoorthy K, Shirakawa T, Eskandarpour M, Hansson J, et al. Induction of melanoma phenotypes in human skin by growth factors and ultraviolet B. *Cancer research*. 2004; 64(3):807–11. PMID: 14871803.
14. D’Orazio J, Jarrett S, Amaro-Ortiz A, Scott T. UV Radiation and the Skin. *International journal of molecular sciences*. 2013; 14(6):12222–48. <https://doi.org/10.3390/ijms140612222> PMID: 23749111
15. Aitken JF, Duffy DL, Green A, Youl P, MacLennan R, Martin NG. Heterogeneity of melanoma risk in families of melanoma patients. *American journal of epidemiology*. 1994; 140(11):961–73. PMID: 7985658.
16. Bliss JM, Ford D, Swerdlow AJ, Armstrong BK, Cristofolini M, Elwood JM, et al. Risk of cutaneous melanoma associated with pigmentation characteristics and freckling: systematic overview of 10 case-control studies. The International Melanoma Analysis Group (IMAGE). *International journal of cancer Journal international du cancer*. 1995; 62(4):367–76. PMID: 7635560.
17. Jhappan C, Noonan FP, Merlino G. Ultraviolet radiation and cutaneous malignant melanoma. *Oncogene*. 2003; 22(20):3099–112. <https://doi.org/10.1038/sj.onc.1206450> PMID: 12789287.
18. Brown KM, Macgregor S, Montgomery GW, Craig DW, Zhao ZZ, Iyadurai K, et al. Common sequence variants on 20q11.22 confer melanoma susceptibility. *Nature genetics*. 2008; 40(7):838–40. <https://doi.org/10.1038/ng.163> PMID: 18488026
19. Bishop DT, Demenais F, Iles MM, Harland M, Taylor JC, Corda E, et al. Genome-wide association study identifies three loci associated with melanoma risk. *Nature genetics*. 2009; 41(8):920–5. <https://doi.org/10.1038/ng.411> PMID: 19578364
20. Gerstenblith MR, Shi J, Landi MT. Genome-wide association studies of pigmentation and skin cancer: a review and meta-analysis. *Pigment cell & melanoma research*. 2010; 23(5):587–606. <https://doi.org/10.1111/j.1755-148X.2010.00730.x> PMID: 20546537
21. Amos CI, Wang LE, Lee JE, Gershenwald JE, Chen WV, Fang S, et al. Genome-wide association study identifies novel loci predisposing to cutaneous melanoma. *Human molecular genetics*. 2011; 20(24):5012–23. <https://doi.org/10.1093/hmg/ddr415> PMID: 21926416
22. Barrett JH, Iles MM, Harland M, Taylor JC, Aitken JF, Andresen PA, et al. Genome-wide association study identifies three new melanoma susceptibility loci. *Nature genetics*. 2011; 43(11):1108–13. <https://doi.org/10.1038/ng.959> PMID: 21983787
23. Macgregor S, Montgomery GW, Liu JZ, Zhao ZZ, Henders AK, Stark M, et al. Genome-wide association study identifies a new melanoma susceptibility locus at 1q21.3. *Nature genetics*. 2011; 43(11):1114–8. <https://doi.org/10.1038/ng.958> PMID: 21983785
24. Song F, Amos CI, Lee JE, Lian CG, Fang S, Liu H, et al. Identification of a melanoma susceptibility locus and somatic mutation in TET2. *Carcinogenesis*. 2014; 35(9):2097–101. <https://doi.org/10.1093/carcin/bgu140> PMID: 24980573

25. Law MH, Bishop DT, Lee JE, Brossard M, Martin NG, Moses EK, et al. Genome-wide meta-analysis identifies five new susceptibility loci for cutaneous malignant melanoma. *Nature genetics*. 2015; 47(9):987–95. <https://doi.org/10.1038/ng.3373> PMID: 26237428
26. Law MH, Macgregor S, Hayward NK. Melanoma genetics: recent findings take us beyond well-traveled pathways. *The Journal of investigative dermatology*. 2012; 132(7):1763–74. <https://doi.org/10.1038/jid.2012.75> PMID: 22475760.
27. Shekar SN, Duffy DL, Youl P, Baxter AJ, Kvaskoff M, Whiteman DC, et al. A population-based study of Australian twins with melanoma suggests a strong genetic contribution to liability. *The Journal of investigative dermatology*. 2009; 129(9):2211–9. <https://doi.org/10.1038/jid.2009.48> PMID: 19357710
28. Mucci LA, Hjelmberg JB, Harris JR, Czene K, Havelick DJ, Scheike T, et al. Familial Risk and Heritability of Cancer Among Twins in Nordic Countries. *Jama*. 2016; 315(1):68–76. <https://doi.org/10.1001/jama.2015.17703> PMID: 26746459.
29. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. *American journal of human genetics*. 2011; 88(1):76–82. <https://doi.org/10.1016/j.ajhg.2010.11.011> PMID: 21167468
30. Lu Y, Ek WE, Whiteman D, Vaughan TL, Spurdle AB, Easton DF, et al. Most common 'sporadic' cancers have a significant germline genetic component. *Human molecular genetics*. 2014; 23(22):6112–8. <https://doi.org/10.1093/hmg/ddu312> PMID: 24943595
31. Mailman MD, Feolo M, Jin Y, Kimura M, Tryka K, Bagoutdinov R, et al. The NCBI dbGaP database of genotypes and phenotypes. *Nature genetics*. 2007; 39(10):1181–6. <https://doi.org/10.1038/ng1007-1181> PMID: 17898773
32. Lee SH, Wray NR, Goddard ME, Visscher PM. Estimating missing heritability for disease from genome-wide association studies. *American journal of human genetics*. 2011; 88(3):294–305. <https://doi.org/10.1016/j.ajhg.2011.02.002> PMID: 21376301
33. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nature genetics*. 2006; 38(8):904–9. <https://doi.org/10.1038/ng1847> PMID: 16862161.
34. Howie BN, Donnelly P, Marchini J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet*. 2009; 5(6):e1000529. Epub 2009/06/23. <https://doi.org/10.1371/journal.pgen.1000529> PMID: 19543373
35. Marchini J, Howie B, Myers S, McVean G, Donnelly P. A new multipoint method for genome-wide association studies by imputation of genotypes. *Nature genetics*. 2007; 39(7):906–13. Epub 2007/06/19. <https://doi.org/10.1038/ng2088> PMID: 17572673.
36. Yang J, Benyamin B, McEvoy BP, Gordon S, Henders AK, Nyholt DR, et al. Common SNPs explain a large proportion of the heritability for human height. *Nature genetics*. 2010; 42(7):565–9. <https://doi.org/10.1038/ng.608> PMID: 20562875
37. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic acids research*. 2010; 38(16):e164. Epub 2010/07/06. <https://doi.org/10.1093/nar/gkq603> PMID: 20601685
38. Nica AC, Parts L, Glass D, Nisbet J, Barrett A, Sekowska M, et al. The architecture of gene regulatory variation across multiple human tissues: the MuTHER study. *PLoS genetics*. 2011; 7(2):e1002003. <https://doi.org/10.1371/journal.pgen.1002003> PMID: 21304890
39. Gamazon ER, Lamba JK, Pounds S, Stark AL, Wheeler HE, Cao X, et al. Comprehensive genetic analysis of cytarabine sensitivity in a cell-based model identifies polymorphisms associated with outcome in AML patients. *Blood*. 2013; 121(21):4366–76. <https://doi.org/10.1182/blood-2012-10-464149> PMID: 23538338
40. Chatzinasiou F, Lill CM, Kypreou K, Stefanaki I, Nicolaou V, Spyrou G, et al. Comprehensive field synopsis and systematic meta-analyses of genetic association studies in cutaneous melanoma. *Journal of the National Cancer Institute*. 2011; 103(16):1227–35. <https://doi.org/10.1093/jnci/djr219> PMID: 21693730
41. Robertson G, Coleman A, Lugo TG. A malignant melanoma tumor suppressor on human chromosome 11. *Cancer research*. 1996; 56(19):4487–92. PMID: 8813145.
42. Robertson GP, Goldberg EK, Lugo TG, Fountain JW. Functional localization of a melanoma tumor suppressor gene to a small (< or = 2 Mb) region on 11q23. *Oncogene*. 1999; 18(20):3173–80. <https://doi.org/10.1038/sj.onc.1202664> PMID: 10340390.
43. Goldberg EK, Glendening JM, Karanjawala Z, Sridhar A, Walker GJ, Hayward NK, et al. Localization of multiple melanoma tumor-suppressor genes on chromosome 11 by use of homozygosity mapping-of-deletions analysis. *American journal of human genetics*. 2000; 67(2):417–31. <https://doi.org/10.1086/302999> PMID: 10877980

44. Lee SH, DeCandia TR, Ripke S, Yang J, Schizophrenia Psychiatric Genome-Wide Association Study C, International Schizophrenia C, et al. Estimating the proportion of variation in susceptibility to schizophrenia captured by common SNPs. *Nature genetics*. 2012; 44(3):247–50. <https://doi.org/10.1038/ng.1108> PMID: 22344220
45. Davis LK, Yu D, Keenan CL, Gamazon ER, Konkashbaev AI, Derks EM, et al. Partitioning the heritability of Tourette syndrome and obsessive compulsive disorder reveals differences in genetic architecture. *PLoS genetics*. 2013; 9(10):e1003864. <https://doi.org/10.1371/journal.pgen.1003864> PMID: 24204291
46. Mitchell JS, Johnson DC, Litchfield K, Broderick P, Weinhold N, Davies FE, et al. Implementation of genome-wide complex trait analysis to quantify the heritability in multiple myeloma. *Scientific reports*. 2015; 5:12473. <https://doi.org/10.1038/srep12473> PMID: 26208354
47. Visscher PM, Hill WG, Wray NR. Heritability in the genomics era—concepts and misconceptions. *Nature reviews Genetics*. 2008; 9(4):255–66. <https://doi.org/10.1038/nrg2322> PMID: 18319743.
48. Shiina T, Hosomichi K, Inoko H, Kulski JK. The HLA genomic loci map: expression, interaction, diversity and disease. *Journal of human genetics*. 2009; 54(1):15–39. <https://doi.org/10.1038/jhg.2008.5> PMID: 19158813.
49. Paco L, Garcia-Lora AM, Casares C, Cabrera C, Algarra I, Collado A, et al. Total loss of HLA class I expression on a melanoma cell line after growth in nude mice in absence of autologous antitumor immune response. *International journal of cancer*. 2007; 121(9):2023–30. <https://doi.org/10.1002/ijc.22925> PMID: 17621628.
50. Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorf LA, Hunter DJ, et al. Finding the missing heritability of complex diseases. *Nature*. 2009; 461(7265):747–53. <https://doi.org/10.1038/nature08494> PMID: 19812666
51. Nicolae DL, Gamazon E, Zhang W, Duan S, Dolan ME, Cox NJ. Trait-associated SNPs are more likely to be eQTLs: annotation to enhance discovery from GWAS. *PLoS Genet*. 2010; 6(4):e1000888. <https://doi.org/10.1371/journal.pgen.1000888> PMID: 20369019
52. Ward LD, Kellis M. Interpreting noncoding genetic variation in complex traits and human disease. *Nat Biotechnol*. 2012; 30(11):1095–106. Epub 2012/11/10. <https://doi.org/10.1038/nbt.2422> PMID: 23138309
53. Cookson W, Liang L, Abecasis G, Moffatt M, Lathrop M. Mapping complex disease traits with global gene expression. *Nature reviews Genetics*. 2009; 10(3):184–94. <https://doi.org/10.1038/nrg2537> PMID: 19223927.
54. Davis LK, Gamazon ER, Kistner-Griffin E, Badner JA, Liu C, Cook EH, et al. Loci nominally associated with autism from genome-wide analysis show enrichment of brain expression quantitative trait loci but not lymphoblastoid cell line expression quantitative trait loci. *Molecular autism*. 2012; 3(1):3. <https://doi.org/10.1186/2040-2392-3-3> PMID: 22591576
55. Gamazon ER, Badner JA, Cheng L, Zhang C, Zhang D, Cox NJ, et al. Enrichment of cis-regulatory gene expression SNPs and methylation quantitative trait loci among bipolar disorder susceptibility variants. *Molecular psychiatry*. 2013; 18(3):340–6. Epub 2012/01/04. <https://doi.org/10.1038/mp.2011.174> PMID: 22212596
56. Gamazon ER, Huang RS, Cox NJ, Dolan ME. Chemotherapeutic drug susceptibility associated SNPs are enriched in expression quantitative trait loci. *Proceedings of the National Academy of Sciences of the United States of America*. 2010; 107(20):9287–92. <https://doi.org/10.1073/pnas.1001827107> PMID: 20442332
57. Wheeler HE, Gamazon ER, Wing C, Njiaju UO, Njoku C, Baldwin RM, et al. Integration of cell line and clinical trial genome-wide analyses supports a polygenic architecture of Paclitaxel-induced sensory peripheral neuropathy. *Clinical cancer research: an official journal of the American Association for Cancer Research*. 2013; 19(2):491–9. <https://doi.org/10.1158/1078-0432.CCR-12-2618> PMID: 23204130
58. Goldstein AM, Tucker MA. Genetic epidemiology of cutaneous melanoma: a global perspective. *Archives of dermatology*. 2001; 137(11):1493–6. PMID: 11708953.
59. Kamb A, Gruis NA, Weaver-Feldhaus J, Liu Q, Harshman K, Tavtigian SV, et al. A cell cycle regulator potentially involved in genesis of many tumor types. *Science*. 1994; 264(5157):436–40. PMID: 8153634.
60. Goldstein AM, Chan M, Harland M, Hayward NK, Demenais F, Bishop DT, et al. Features associated with germline CDKN2A mutations: a GenoMEL study of melanoma-prone families from three continents. *Journal of medical genetics*. 2007; 44(2):99–106. <https://doi.org/10.1136/jmg.2006.043802> PMID: 16905682
61. Goldstein AM, Chan M, Harland M, Gillanders EM, Hayward NK, Avril MF, et al. High-risk melanoma susceptibility genes and pancreatic cancer, neural system tumors, and uveal melanoma across GenoMEL. *Cancer research*. 2006; 66(20):9818–28. <https://doi.org/10.1158/0008-5472.CAN-06-0494> PMID: 17047042.

62. Kennedy C, ter Huurne J, Berkhout M, Gruis N, Bastiaens M, Bergman W, et al. Melanocortin 1 receptor (MC1R) gene variants are associated with an increased risk for cutaneous melanoma which is largely independent of skin type and hair color. *The Journal of investigative dermatology*. 2001; 117(2):294–300. <https://doi.org/10.1046/j.0022-202x.2001.01421.x> PMID: 11511307.
63. Matichard E, Verpillat P, Meziani R, Gerard B, Descamps V, Legroux E, et al. Melanocortin 1 receptor (MC1R) gene variants may increase the risk of melanoma in France independently of clinical risk factors and UV exposure. *Journal of medical genetics*. 2004; 41(2):e13. <https://doi.org/10.1136/jmg.2003.011536> PMID: 14757863
64. Stratigos AJ, Yang G, Dimisianos R, Nicolaou V, Stefanaki I, Katsambas AD, et al. Germline CDKN2A mutations among Greek patients with early-onset and multiple primary cutaneous melanoma. *The Journal of investigative dermatology*. 2006; 126(2):399–401. <https://doi.org/10.1038/sj.jid.5700078> PMID: 16374456.
65. Aitken J, Welch J, Duffy D, Milligan A, Green A, Martin N, et al. CDKN2A variants in a population-based sample of Queensland families with melanoma. *Journal of the National Cancer Institute*. 1999; 91(5):446–52. PMID: 10070944.
66. 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. <https://doi.org/10.1371/journal.pgen.1000074> PMID: 18483556
67. 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. <https://doi.org/10.1038/ng.160> PMID: 18488028.
68. Gandini S, Sera F, Cattaruzza MS, Pasquini P, Zanetti R, Masini C, et al. Meta-analysis of risk factors for cutaneous melanoma: III. Family history, actinic damage and phenotypic factors. *European journal of cancer*. 2005; 41(14):2040–59. <https://doi.org/10.1016/j.ejca.2005.03.034> PMID: 16125929.
69. Chang YM, Barrett JH, Bishop DT, Armstrong BK, Bataille V, Bergman W, et al. Sun exposure and melanoma risk at different latitudes: a pooled analysis of 5700 cases and 7216 controls. *International journal of epidemiology*. 2009; 38(3):814–30. <https://doi.org/10.1093/ije/dyp166> PMID: 19359257
70. Newton-Bishop JA, Chang YM, Elliott F, Chan M, Leake S, Karpavicius B, et al. Relationship between sun exposure and melanoma risk for tumours in different body sites in a large case-control study in a temperate climate. *European journal of cancer*. 2011; 47(5):732–41. <https://doi.org/10.1016/j.ejca.2010.10.008> PMID: 21084183
71. Cormier JN, Xing Y, Ding M, Lee JE, Mansfield PF, Gershenwald JE, et al. Ethnic differences among patients with cutaneous melanoma. *Archives of internal medicine*. 2006; 166(17):1907–14. <https://doi.org/10.1001/archinte.166.17.1907> PMID: 17000949.
72. Rouhani P, Hu S, Kirsner RS. Melanoma in Hispanic and black Americans. *Cancer control: journal of the Moffitt Cancer Center*. 2008; 15(3):248–53. PMID: 18596677.