

ASSOCIATION STUDIES ARTICLE

Germline variants are associated with increased primary melanoma tumor thickness at diagnosis

Ernest Mangantig¹, Stuart MacGregor^{2,†}, Mark M. Iles^{3,‡}, Richard A. Scolyer^{4,5,6,7,¶}, Anne E. Cust^{4,6,8,||}, Nicholas K. Hayward⁹, Grant W. Montgomery^{10,††}, David L. Duffy¹¹, John F. Thompson^{4,5,6}, Anjali Henders^{10,11}, Lisa Bowdler¹¹, Casey Rowe^{12,13}, Gemma Cadby¹⁴, Graham J. Mann^{4,15,16}, David C. Whiteman¹⁷, Georgina V. Long^{4,6,18,19}, Sarah V. Ward¹⁴, Kiarash Khosrotehrani^{12,13,‡‡}, Jennifer H. Barrett³ and Matthew H. Law^{2,*,§}

¹Regenerative Medicine Cluster, Advanced Medical and Dental Institute, Universiti Sains Malaysia, 13200, Pulau Pinang, Malaysia, ²Statistical Genetics, QIMR Berghofer Medical Research Institute, Brisbane, Queensland, 4006, Australia, ³Leeds Institute for Data Analytics, University of Leeds, Leeds LS2 9JT, UK, ⁴Melanoma Institute Australia, The University of Sydney, Sydney, New South Wales, 2065, Australia, ⁵Department of Tissue Oncology and Diagnostic Pathology, Royal Prince Alfred Hospital, Sydney, New South Wales, 2050, Australia, ⁶Faculty of Medicine and Health, The University of Sydney, Sydney, New South Wales, 2050, Australia, ⁷Department of Tissue Oncology and Diagnostic Pathology, New South Wales Health Pathology, Sydney, New South Wales, 2000, Australia, ⁸School of Public Health, The University of Sydney, Sydney, New South Wales, 2006, Australia, ⁹Oncogenomics, QIMR Berghofer Medical Research Institute, Brisbane, Queensland, 4006, Australia, ¹⁰Molecular Biology, The University of Queensland, Brisbane, Queensland, 4102, Australia, ¹¹Genetic Epidemiology, QIMR Berghofer Medical Research Institute, Brisbane, Queensland, 4006, Australia, ¹²Experimental Dermatology Group, Diamantina Institute, The University of Queensland, Brisbane, Queensland, 4102, Australia, ¹³Department of Dermatology, Princess Alexandra Hospital, Brisbane, Queensland, 4102, Australia, ¹⁴School of Population and Global Health, The University of Western Australia, Perth, Western Australia, 6009, Australia, ¹⁵Centre for Cancer Research, Westmead Institute for Medical Research, University of Sydney, New South Wales, 2145, Australia, ¹⁶John Curtin School of Medical Research, Australian National University, Canberra, Australian Capital Territory, 2601, Australia, ¹⁷Cancer Control, QIMR Berghofer Medical Research Institute, Brisbane, Queensland, 4006, Australia, ¹⁸Department of Medical Oncology, Mater Hospital, North Sydney, NSW, 2060, Australia and ¹⁹Department of Medical Oncology, Royal North Shore Hospital, St Leonards, New South Wales, 2065, Australia

*To whom correspondence should be addressed at: 300 Herston Road, Herston QLD 4006. Tel: +61 7 3362 0213; Fax: +61 7 3362 0101; Email: matthew.law@qimrberghofer.edu.au

[†]Stuart MacGregor, <https://orcid.org/0000-0001-6731-8142>

[‡]Mark M. Iles, <https://orcid.org/0000-0002-2603-6509>

[¶]Richard A. Scolyer, <http://orcid.org/0000-0002-8991-0013>

^{||}Anne E. Cust, <https://orcid.org/0000-0002-5331-6370>

^{††}Grant W. Montgomery, <https://orcid.org/0000-0002-4140-8139>

^{‡‡}Kiarash Khosrotehrani, <https://orcid.org/0000-0002-6406-4076>

[§]Matthew H. Law, <http://orcid.org/0000-0002-4303-8821>

Received: March 24, 2020. Revised: July 29, 2020. Accepted: October 8, 2020

© The Author(s) 2020. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

Abstract

Germline genetic variants have been identified, which predispose individuals and families to develop melanoma. Tumor thickness is the strongest predictor of outcome for clinically localized primary melanoma patients. We sought to determine whether there is a heritable genetic contribution to variation in tumor thickness. If confirmed, this will justify the search for specific genetic variants influencing tumor thickness. To address this, we estimated the proportion of variation in tumor thickness attributable to genome-wide genetic variation (variant-based heritability) using unrelated patients with measured primary cutaneous melanoma thickness. As a secondary analysis, we conducted a genome-wide association study (GWAS) of tumor thickness. The analyses utilized 10 604 individuals with primary cutaneous melanoma drawn from nine GWAS datasets from eight cohorts recruited from the general population, primary care and melanoma treatment centers. Following quality control and filtering to unrelated individuals with study phenotypes, 8125 patients were used in the primary analysis to test whether tumor thickness is heritable. An expanded set of 8505 individuals (47.6% female) were analyzed for the secondary GWAS meta-analysis. Analyses were adjusted for participant age, sex, cohort and ancestry. We found that 26.6% (SE 11.9%, $P = 0.0128$) of variation in tumor thickness is attributable to genome-wide genetic variation. While requiring replication, a chromosome 11 locus was associated ($P < 5 \times 10^{-8}$) with tumor thickness. Our work indicates that sufficiently large datasets will enable the discovery of genetic variants associated with greater tumor thickness, and this will lead to the identification of host biological processes influencing melanoma growth and invasion.

Introduction

Cutaneous melanoma (hereafter melanoma) is a potentially fatal skin cancer resulting from the uncontrolled growth of melanocytes, the pigment-producing cells of the skin. In 2019, in Australia, there were estimated to be over 38 000 new cases of melanoma (15 000 invasive) and nearly 2000 deaths, and in the United States an estimated 192 000 cases of melanoma (97 000 invasive) with over 7000 deaths (1,2).

Tumor thickness, measured from the top of the epidermal granular layer to the deepest point of tumor invasion, is the single strongest predictive factor for mortality in clinically localized primary cutaneous melanoma and widely used in disease staging and prognostication (3–5). It is not yet known whether primary melanoma thickness at the time of initial diagnosis is influenced by heritable factors. Beyond the time of presentation for biopsy/excision of the lesion, factors such as invasion potential, growth rate and immunosurveillance are likely to be relevant to tumor thickness at presentation, and these may be influenced by host germline characteristics (6–9).

Previous investigations have identified a limited number of potential loci associated with primary melanoma tumor thickness including at *MMP1*, *FGFR4*, *VDR* and loci previously associated with melanoma risk e.g. *MC1R* (10–16). No loci have been validated independently. Two genome-wide association studies (GWAS) of primary melanoma tumor thickness did not identify any genetic variants that were significant after multiple testing correction (17,18). These findings are summarized in [Supplementary Material, Table S1](#).

Using over 8000 patients with primary cutaneous melanoma from eight cohorts ([Table 1](#)), we sought to determine whether heritable genetic factors influence melanoma tumor thickness and the extent of such influence. Due to the possible impact of shared environment on tumor thickness, and the increased surveillance from diagnosis of melanoma in a close relative, our analysis used distantly related individuals to remove this potential for confounding (19–24). Confirmation that germline genetic variation influences primary melanoma tumor thickness would imply that, with a sufficiently large dataset, specific genetic variants associated with greater tumor thickness can be identified. Identifying these genetic variants, and the genes they influence, will lead to a better understanding of the host

biological processes that are important for the growth and invasion of melanoma.

Results

Contribution of genome-wide germline genotype to variation in tumor thickness

Age- and sex- adjusted residuals of natural log transformed tumor thickness were generated, and GREML-LDMS-I (23) was used to determine the contribution of genome-wide germline variants with a minor allele frequency > 0.001 to tumor thickness variation (heritability, h^2_{SNPs} , [Supplementary Material, Fig. S1](#), [Supplementary Material, Fig. S2](#), [Methods](#)).

In the combined dataset ($N = 8125$) with the first six ancestry principal components and an individual cohort membership variable fitted as covariates, tumor thickness h^2_{SNP} was 0.266 (standard error (SE) = 0.119, $P = 0.0128$). Analyzing the age- and sex- corrected residuals of tumor thickness, and fitting a dataset membership covariate, may not have completely accounted for all differences across datasets. However, repeating analysis with rank normalized tumor thickness residuals gave a similar $h^2_{\text{SNP}} = 0.248$ (SE = 0.119, $P = 0.0188$, [Supplementary Material, Fig. S3](#)), indicating it is unlikely individual study differences are driving the observed heritability. The estimate from the random-effects meta-analysis of the h^2_{SNP} for each discrete dataset calculated separately, while losing power by not using genetic relationships across sample sets, was consistent with the overall result ($h^2_{\text{SNP}} = 0.264$) but with wider confidence intervals (SE = 0.158, $P = 0.095$; [Methods](#), [Supplementary Material, Fig. S4](#)).

Acral lentiginous melanoma tends to be diagnosed later and have a greater thickness (25). Excluding 92 acral lentiginous melanoma cases (histology data were available for all but the MD Anderson Cancer Center (MDACC) cohort) did not meaningfully change the results ($h^2_{\text{SNP}} = 0.281$, SE = 0.121, $P = 0.0104$).

Genome-wide association study of tumor thickness

As the h^2_{SNP} was significantly different to zero, indicating a role for germline genetic variation in the thickness of

Table 1. Population demographics

	AMFS	CAMBRIDGE	EPIGENE	LEEDS	MDACC	MIA	Q-MEGA 610 k	Q-MEGA omni	WAMHS
Genotyped	549	492	787	2192	1982	1745	912	656	1289
Age, sex	532	491	729	1898	1523	1240	543	457	1253
¹ Tumor thickness	529	491	729	1895	1523	1177	534	451	1195
² Post-QC	528	488	728	1888	1522	1177	534	451	1189
³ Thickness < 0.8 mm	350	223	539	200	523	335	351	315	711
³ Thickness ≥ 0.8 mm and ≤ 1.0 mm	79	72	61	343	180	138	74	47	162
³ Thickness > 1.0 mm and ≤ 2.0 mm	68	109	82	724	405	354	76	59	183
³ Thickness > 2.0 mm and ≤ 4.0 mm	18	62	38	411	265	234	25	20	87
³ Thickness > 4.0 mm	13	22	8	210	149	116	8	10	46
Mean tumor thickness mm (SD)	0.861 (1.088)	1.288 (1.252)	0.791 (1.184)	2.145 (2.013)	1.932 (2.548)	1.950 (2.999)	0.821 (0.875)	0.833 (1.005)	1.104 (1.770)
⁴ Transformed tumor thickness mean (SD)	-0.479 (0.721)	-0.104 (0.838)	-0.593 (0.730)	0.476 (0.732)	0.190 (0.927)	0.255 (0.891)	-0.506 (0.739)	-0.518 (0.737)	-0.352 (0.855)
⁵ Women (%)	336 (63.6%)	257 (62.7%)	231 (31.7%)	1082 (57.3%)	634 (41.7%)	474 (40.3%)	287 (53.8%)	223 (49.5%)	496 (41.8%)
Men	192	231	497	806	888	703	247	228	693
Genomic inflation λ	0.971	0.995	0.994	1.000	0.995	1.000	0.991	0.961	0.995
⁶ IBD < 0.025	499	483	660	1864	1478	1079	504	420	1138

¹ Number of individuals with tumor thickness as a continuous measurement in mm. ² Number of individuals following quality control and cleaning (Methods; Post-QC). ³ For reference, individuals for each cohort are grouped based on T categories as defined in the eighth edition American Joint Committee on Cancer Melanoma Staging System (4). ⁴ Continuous tumor thickness measurements in millimeter were natural log transformed. SDs are reported where indicated. ⁵ Sex ratio is reported for the post-QC set with all phenotypes. ⁶ Determining the genetic contribution to melanoma thickness requires distantly related participants so only those pairs with identity-by-descent (IBD) $\pi_{\text{hat}} > 0.025$ are included; thickness distribution for the 8125 samples filtered to IBD 0.025 is reported in S3 Table. (Methods). Queensland study of Melanoma: Environment and Genetic Associations (Q-MEGA) samples GWAS were performed by the HumanHap610 (610 k) and Omni1-Quad (omni) genotyped arrays used (Methods). MIA, MDACC, WAMHS.

primary tumors, we therefore performed a linear regression GWAS of primary melanoma tumor thickness in each dataset (Methods). Following the subsequent meta-analysis of the individual genome-wide results, there was no evidence of genomic inflation ($N=8505$; genomic inflation $\lambda=1.00$; Supplementary Material, Fig. S5). Two genetic variants are in linkage-disequilibrium ($r^2_{\text{EUR}}=0.69$) at a single locus on chromosome 11 reached genome-wide significance (Fig. 1). The regional association plot is shown in Fig. 2. The lead variant was rs183471242 ($P=3.56 \times 10^{-9}$; Table 2). rs183471242 is in an intron of the gene low-density lipoprotein receptor class A domain containing 3 (LDLRAD3; Fig. 2). The distribution of natural log transformed tumor thickness by genotype is displayed in Fig. 3; the distribution of effect sizes across GWAS is shown in Supplementary Material, Fig. S6. Both rs183471242 (G/A) and rs566382949 (C/A) are rare (HRC v1.1 minor A allele 0.0098 and 0.0082, respectively), with the minor allele associated with thicker tumors (Table 2, Fig. 3). Each minor allele of rs183471242 translates to a 1.423-fold increase on the transformed residuals of tumor thickness.

As a sensitivity analysis, we performed the regression of residual and rank normalized residual tumor thickness on rs183471242 and rs566382949 in the same combined dataset used for the GREML analysis ($N=8125$), fitting the first six PCs and study covariates in the model. Both SNPs remained genome-wide significantly associated ($P < 5 \times 10^{-8}$) with tumor thickness.

While none of the previously reported genetic variants associated with primary tumor thickness reached genome-wide significance in this study, the IRF4 functional genetic variant rs12203592 was the most strongly associated (fixed $P=6.50 \times 10^{-4}$; Supplementary Material, Table S1) (10–18).

Discussion

Given inconsistency in identifying specific germline variants associated with primary melanoma tumor thickness, we first determined whether tumor thickness is heritable (that is, a proportion of phenotypic variance can be explained by additive genetic variants) (10–18). Traditionally, twin- or family-based approaches are used to measure trait heritability, and do so by assessing whether more closely related individuals tend to have more similar phenotypes. However, shared environment and behaviors can confound heritability estimates. For example, diagnosis of melanoma in a relative can lead to increased surveillance and earlier detection, which may influence thickness (19,20,27–33). As a result, tumor thickness can become inversely correlated to degree of genetic relationship, biasing estimates of heritability. An effective alternative approach is to estimate the genetic contribution to thickness using distantly related individuals, removing the potential confounding between thickness and melanoma diagnosis in a relative and/or earlier diagnosis (22). This alternative approach relies on genome-wide single nucleotide polymorphism (SNP) arrays, yielding an estimate of the variation attributable to these genetic variants; this term is referred to as h^2_{SNP} . Since arrays do not genotype all genetic variants, h^2_{SNP} represents a lower bound of trait heritability.

Using a large collection of distantly related individuals with melanoma ($N=8125$ following filtering to remove individuals such that there were no pairs where their relationship was closer than identity-by-descent $\pi_{\text{hat}}=0.025$, Methods, Table 1), we estimated the h^2_{SNP} for tumor thickness to be 0.266 (95% CI 0.033–0.500). This h^2_{SNP} result was robust to various sensitivity

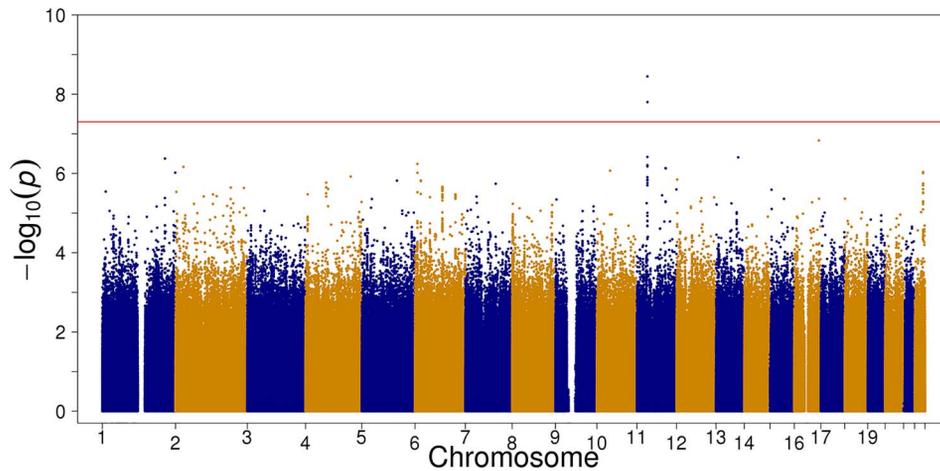


Figure 1. Manhattan plot of P -values from the meta-analysis of genome-wide association studies of tumor thickness. Negative Log_{10} of observed fixed-effects meta-analysis P -values plotted by chromosome position. Red line indicates genome-wide significance ($P = 5 \times 10^{-8}$). Sample size is reported in Table 1.

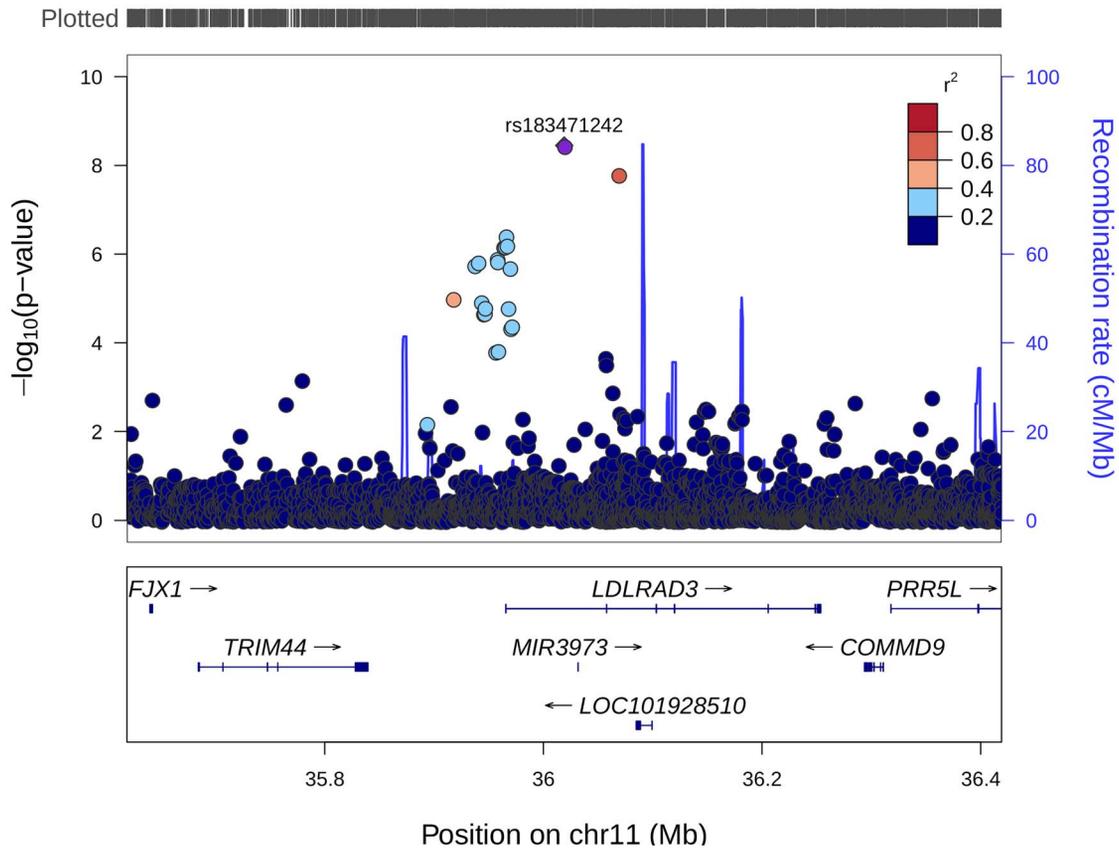


Figure 2. Regional association plot for rs183471242. Negative Log_{10} of observed fixed-effects meta-analysis P -values plotted by chromosome position. Linkage disequilibrium r^2 of plotted SNPs with the lead SNP rs183471242 is displayed. Plot generated using LocusZoom (26).

analyses (Methods, Results). As this result indicates that primary melanoma tumor thickness is heritable, it shows that with sufficiently well-powered cohorts, it will be possible to identify specific genetic variants associated with tumor thickness.

Following confirmation that tumor thickness is heritable, we used the complete dataset to perform a GWAS meta-analysis, identifying two genetic variants in a locus on chromosome 11 at genome-wide significance ($P < 5 \times 10^{-8}$). As tumor thickness

values were natural log transformed prior to analysis, the effect size for each minor allele of rs183471242 is associated with a 1.42-fold higher tumor thickness. rs183471242 is rare (minor A allele frequency 0.0098) with 190 heterozygote GA and 1 homozygous AA samples (Fig. 3). This variant appears to be rarer in non-European populations with an MAF of 0.0018 in African populations and not observed in 780 East Asian samples (34). While these genetic variants are in an intron of LDLRAD3,

Table 2. Genetic variants associated with tumor thickness following multiple testing correction

CHR	BP	rsID	EA/NEA	EA FREQ	P	BETA	Q	I ²	
rs183471242	11	36019025	A/G	0.0098	3.56×10^{-9}	0.353	0.81	0	
rs566382949	11	36068615	A/C	0.0082	1.58×10^{-8}	0.382	0.97	0	
rsID	AMFS	CAMBRIDGE	EPIGENE	LEEDS	MDACC	MIA	Q-MEGA 610 k	Q-MEGA omni	WAMHS
rs183471242	0.598	0.501	0.443	0.241	0.389	0.459	0.093	0.229	0.369
rs566382949	0.326	0.518	0.418	0.302	0.232	0.541	0.240	0.406	0.435

We report hg19 chromosome (CHR) and base pair (BP) positions for genetic variants (rsID). The effect allele (EA) and non-effect allele (NEA) are provided, as is the HRC frequency of the EA (EA FREQ). The fixed effects meta-analysis P (P), effect size (BETA) on the residuals of fitting age and sex on natural log transformed tumor thickness; the first six ancestry principal components were included as covariates in the regression (Methods). As there is no heterogeneity (Q, I^2) the random effects meta-analysis values are identical. We also report effect size estimates for individual datasets. The distribution of these measures is displayed in [Supplementary Material, Fig. S6](#). Q-MEGA samples were analyzed by HumanHap610 (610 k) and Omni1-Quad (omni) genotyped array used (Methods). Sample size is reported in [Table 1](#).

the closest gene is not always the target of associated genetic variants. However, interrogation of public gene expression and annotation resources do not reveal an obvious functional target for these genetic variants (35–38). LDLRAD3 is a member of the LDL receptor family and may play a role in activating genes involved in protein ubiquitination (39). Neither of these two genetic variants were associated ($P > 0.05$) with ease of tanning, childhood sunburns or skin and hair color in the UK Biobank (data not shown), melanoma risk, nor nevus count (40,41). While none of the previously reported genetic variants associated with tumor thickness reached a P -value $< 5 \times 10^{-8}$, the IRF4 functional genetic variant rs12203592 was the most strongly associated ($P = 6.51 \times 10^{-4}$, $I^2 = 36.4\%$, [Supplementary Table 1](#)). The direction of effect is the same as in the previous report from the Western Australian Melanoma Health Study; this dataset is included in this analysis (10). The IRF4 SNP rs12203592 has been associated with risk of melanoma (42,43). It is not clear to what extent genetic variants associated with cancer risk influence outcome (e.g. there is no overlap between genetic variants associated with risk or survival for lung cancer (44) or breast cancer (45)), and as such, it is unclear if we would expect the other known melanoma risk variants (e.g. in MC1R) to associate with tumor thickness in a larger analysis.

The thickness of a tumor at diagnosis is the outcome of its duration and rate of growth. In practice, these two factors are hard to assess post detection and unknown before diagnosis, and as a result, the relative importance of either is unclear (6,27–29,33,46). Our work shows that germline genetic variation plays a role in thickness. While the specific mechanisms are yet to be determined, host germline genetic variants may influence tumor growth rate, immunosurveillance or characteristics relating to time to discovery such as appearance of the tumor or the assiduousness of personal skin checks. While different subtypes of melanoma grow at different rates, accurate heritability estimates require very large sample sizes; thus (and to avoid repeated multiple testing), we have not performed subtype specific analyses beyond excluding the acral subtype. Other factors may influence tumor thickness. Higher body mass index (BMI) has been associated with greater thickness; however, fitting a trait that is itself genetically controlled as a covariate can bias genetic association towards false positives and we hence did not include this in our analysis (47–50).

While both our heritability estimate and the genetic variants associated with tumor thickness following correction for multiple testing require replication in a sufficiently large dataset, these findings indicate germline variants impact tumor development, a strong predictor of melanoma outcome. Discovery of

specific genetic variants will enable identification of host biological processes influencing melanoma growth and invasion.

Materials and Methods

Dataset descriptions

Overall demographics of contributing datasets are summarized in [Table 1](#). Full descriptions of each contributing dataset can be found in the [Supplementary Material, Note](#).

Quality control, cleaning and imputation of genome-wide genotype data

PLINK v1.9 and R 3.3.2 (51,52) were used for quality control and cleaning of genome-wide genotype data. Genotyped variants were filtered out if they had a minor allele frequency < 0.01 , Hardy-Weinberg equilibrium P -value $< 5 \times 10^{-4}$ in cancer-free individuals (where melanoma cases were genotyped/cleaned in concert with healthy individuals) or $< 5 \times 10^{-10}$ in those with melanoma. To remove samples with low-quality DNA, or other issues that may impact analysis (e.g. sample contamination or inbreeding) (53), individuals were excluded if they had missingness > 0.03 , heterozygosity more than three standard deviations (SDs) from the rest of the population, a mismatch between recorded sex and X chromosome determined sex or were considered non-European. European ancestry was determined by principal components analysis using 1000 Genomes European populations as a reference set (54). Individuals more than three SDs from the mean of principal component 1 or 2 were excluded. Relatedness across and within genotyped sets was measured by identity-by-descent pi-hat scores using PLINK (51). For pairs with pi-hat > 0.15 , the individual with the highest missing genotype rate was dropped.

The Michigan Imputation Server was used to impute individuals ([Table 1](#)) to the Haplotype Reference Consortium panel (HRC version 1), and genetic variants with an imputation quality score $RSQ > 0.3$, minor allele frequency > 0.001 and minor allele count > 3 were retained for analysis (55,56). While the approach we have used to determine the genetic contribution to melanoma thickness (see below) is robust to imputation quality, $RSQ > 0.3$ excludes potentially poorly imputed genetic variants (23).

Cleaning and normalization of tumor thickness

Research participant demographics are presented in [Table 1](#). Primary melanoma tumor thickness measurements were extracted from pathology reports. For participants with multiple

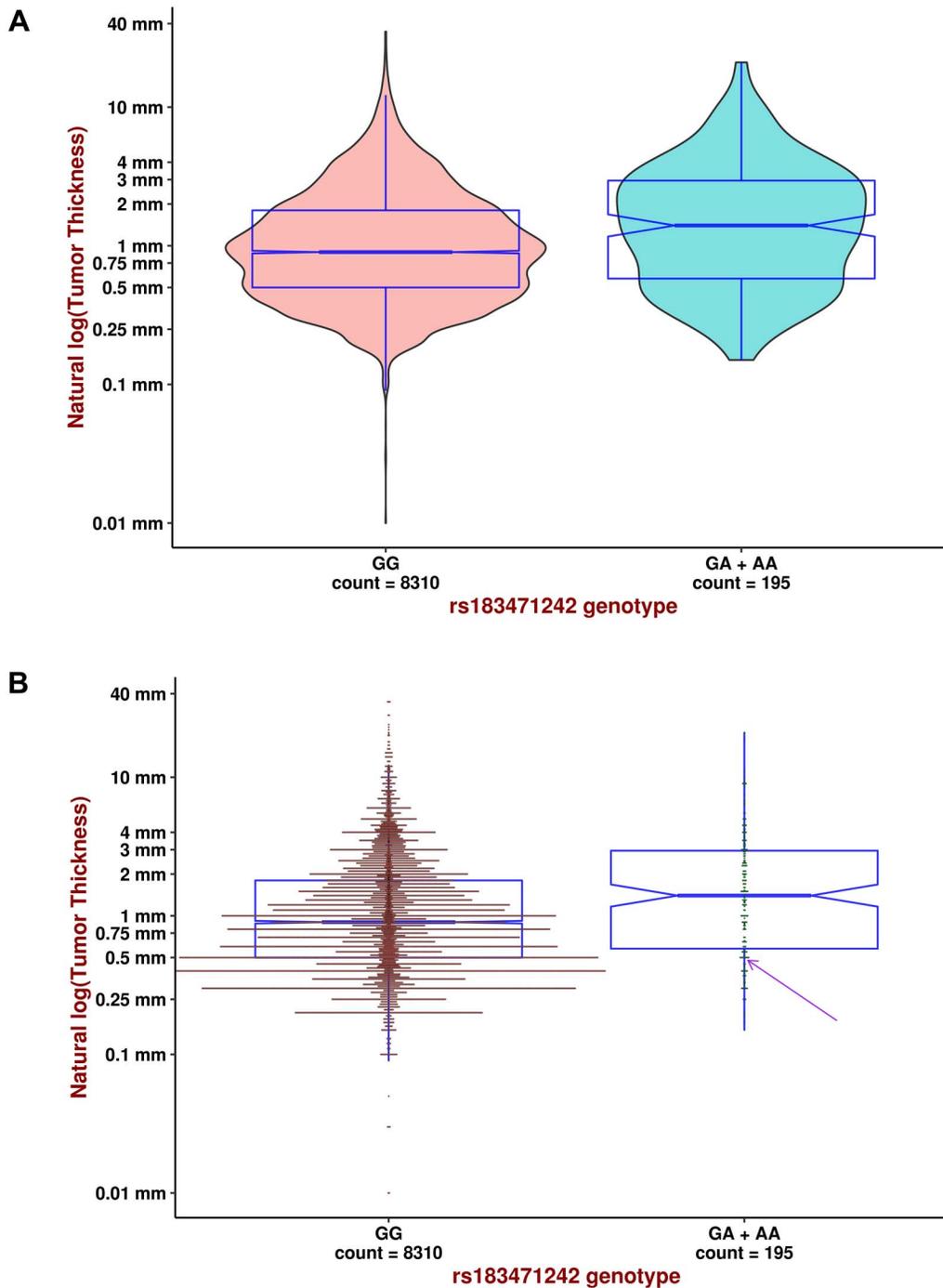


Figure 3. Distribution of natural log transformed primary cutaneous melanoma tumor thickness by rs183471242 genotypes. Data are reported for the combined meta-analysis of all studies for this genetic variant (8505 individuals). For plotting purposes, we display the natural log transform of tumor thickness by genotype rather than the residuals of tumor thickness as used in the heritability estimation and GWAS. In total, there are 8310 homozygous GG, 194 AG and a single AA genotype (HRC v1.1 minor A allele 0.0098). AG and AA genotypes have been plotted together. Distribution of tumor thickness is described using a notched whisker plot (blue) where the midpoint of the notch is the median, and the 95% confidence interval of that median is represented by the notched region. The boundaries of the boxed area extend to the first and third quartiles. The whiskers represent the $1.5 \times$ the interquartile range. The same data are displayed twice with differing secondary layers to display the distribution of tumor thickness residuals; the first is a violin plot, and the second displays the individual results. The individual with the AA genotype had a primary tumor thickness of 0.5 mm, and their position is indicated by an arrow in the second plot.

melanomas, the first primary tumor was used. The distribution of tumor thickness is reported in Table 1. As tumor thickness is not normally distributed, measurements were natural log transformed. Age- and sex-adjusted residuals of

transformed tumor thickness were used for analyses. The distribution of tumor thickness between cohorts is significantly different ($P < 2 \times 10^{-16}$ from ANOVA with age and sex fitted as covariates, see Supplementary Material, Methods). Pairwise

comparisons are reported in [Supplementary Material, Table S2](#). As a result, analyses were performed with a cohort variable fitted and as a sensitivity analysis repeated with rank transformation of tumor thickness.

Estimation of heritability for a complex trait

We sought to estimate the heritability (h^2 , defined as the proportion of phenotypic variance explained by additive genetic variants) for primary cutaneous melanoma tumor thickness. Heritability is traditionally estimated from family data, with high heritability inferred when individuals who are closely related have more similar phenotypes than those who are more distantly related. An alternative approach to family-based methods is to estimate the genetic contribution to thickness using only distantly related individuals to estimate heritability attributable to SNPs, h^2_{SNP} (22).

Early methods for estimating h^2_{SNP} such as genome-based restricted maximum likelihood as implemented in the Genome-wide Complex Trait Analysis software use only directly genotyped variants (22). However, imputation of genetic variants not present on genotyping arrays can improve discovery power and resolution for standard genome-wide association studies and heritability estimates (21,56). In aggregate, rare variants capture on average one third of the h^2_{SNP} , and their inclusion can yield more accurate h^2_{SNP} estimates (21,24).

The h^2_{SNP} was determined using an extension of the genome-based restricted maximum likelihood approach designed for imputed data, GREML-LDMS-I, as implemented in the Genome-wide Complex Trait Analysis software (21–23). GREML-LDMS was used recently to determine that ~50% of the heritability of height is accounted for by genetic variants with a minor allele frequency between 0.0001 and 0.1 (57). GREML-LDMS-I is made robust to the (unknown) underlying trait genetic architecture by dividing input genetic variants into bins based on their minor allele frequency and degree of linkage disequilibrium (23,24). This is important as incorrect modeling of the underlying genetic architecture can lead to under- or over-estimation of h^2_{SNP} (21,24).

Construction of a combined imputed dataset for array-based heritability estimates

Imputed dosage data from the Michigan Imputation Server in variant-call format were converted to best guess format (genotype dosage ≤ 0.5 as 0, 0.5–1.5 as 1 and > 1.5 as 2) and merged into a single combined dataset using PLINK v1.91.4 beta3 (21,51). PLINK binary files were converted into a genetic relationship matrix by Genome-wide Complex Trait Analysis v1.91.4 (21,22). To ensure only distantly related individuals were included, the merged dataset was filtered such that no pair had identify-by-descent π -hat > 0.025 (21,22) (Table 1; final combined sample size of 8125).

We used GREML-LDMS-I to estimate h^2_{SNP} for genetic variants binned by minor allele frequency (0.4–0.5, 0.3–0.4, 0.2–0.3, 0.1–0.2, 0.01–0.1, 0.001–0.01) and linkage-disequilibrium quartile. Linkage-disequilibrium scores were estimated for individual genetic variants rather than regions of variants, as this approach produces unbiased estimates in the presence of all possible genetic architectures (21).

Spurious genetic similarities (e.g. ancestry, cohort or batch effects) can bias h^2 estimates (21). To address this, we fitted the first six ancestry principal components and a dataset membership variable (a binary yes/no variable for membership in a

given cohort) in the GREML-LDMS-I analyses. As a sensitivity analysis, we repeated analyses following rank normalization of tumor-thickness residuals within each cohort.

While the larger merged dataset increased power by leveraging distant genetic relationships across and between individual datasets, it may have introduced bias due to subtle differences in ancestry, tumor thickness or genotyping methods within each individual dataset. As an additional sensitivity analysis, we estimated h^2_{SNP} in the individual imputed datasets (filtered to identity-by-descent π -hat < 0.025 within the individual dataset rather than across all datasets) using standard genome-based restricted maximum likelihood (22). The binned GREML-LDMS-I approach was not applied to individual datasets as they were too small. Individual dataset's h^2_{SNP} estimates were then combined using a random effects meta-analysis using the *metafor* package in R (58). *metafor* was also used to generate forest plots.

Genome-wide association study of tumor thickness

To identify specific genetic variants associated with tumor thickness, within each individual dataset, residual tumor thickness was regressed on imputed genome-wide genotype dosages with the first six ancestry principal components included as covariates. Individual dataset results were further filtered by removing variants with an extreme effect size estimate (> 2 or < -2 on the rank-normalized residuals of natural log-transformed tumor thickness; as a reference, in the Australian Melanoma Family Study (AMFS) dataset, this removes genetic variants with an effect size estimate > 8 SDs from an effect size of 0). Genome-wide association results for each individual dataset were combined by inverse variance-weighted fixed effects meta-analysis in PLINK v1.91.4 beta3 (51). The total number of genetic variants tested was 13 517 544. Genetic variants were deemed significant if their P-value was less than a genome-wide multiple testing corrected threshold of $P < 5 \times 10^{-8}$.

In a GWAS, the majority of genetic variants are not expected to be associated and their X^2 distribution should match the null and have a median of ~0.456 (59). We report the genomic inflation λ , the median meta-analysis $X^2/0.456$.

Additional analysis methods are reported in the supplementary note.

Ethics approval

Individual studies' ethical approval details are reported in the [Supplementary Material, Note](#). Overall, approval was managed by the Human Research Ethics Committee of the QIMR Berghofer Medical Research Institute.

Supplementary Material

[Supplementary Material](#) is available at HMG online.

Acknowledgements

E.M. was supported by the Malaysian Ministry of Higher Education and Universiti Sains Malaysia to study for a PhD at the University of Leeds. A.E.C. was supported by a National Health and Medical Research Council (NHMRC) of Australia Career Development Fellowship (1147843). K.K. was supported by an NHMRC Career Development Fellowship (1125290). M.M.I. was supported by Cancer Research UK (c588/a19167) and the NIH (ca083115). R.A.S. and G.V.L. are supported by NHMRC Practitioner Fellowships; R.A.S. and J.F.T. also acknowledge support from an NHMRC program grant. D.C.W., S.M. and N.K.H.

were supported by NHMRC Research Fellowships (1058522, 1155413, 1154543 and 1117663). We thank Nicholas G. Martin for assistance with access to data from the Q-MEGA cohort and with manuscript writing. This work was conducted using the UK Biobank Resource (application number 25331).

QMEGA

We are grateful for the critical support of sample processing, project management, data collection, support and research staff including (but not limited to) A. Baxter, M. de Nooyer, I. Gardner, D. Statham, B. Haddon, M.J. Wright, J. Palmer, J. Symmons, B. Castellano, L. Bowdler, S. Smith, D. Smyth, L. Wallace, M.J. Campbell, A. Caracella, M. Kvaskoff, O. Zheng, B. Chapman and H. Beeby. The Q-MEGA study was supported by the Melanoma Research Alliance, the NIH NCI (CA88363, CA83115, CA122838, CA87969, CA055075, CA100264, CA133996 and CA49449), the NHMRC (200071, 241944, 339462, 380385, 389927, 389875, 389891, 389892, 389938, 443036, 442915, 442981, 496610, 496675, 496739, 552485, 552498, APP1049894), the Cancer Councils of New South Wales, Victoria and Queensland, the Cancer Institute New South Wales, the Cooperative Research Centre for Discovery of Genes for Common Human Diseases (CRC), Cerylid Biosciences (Melbourne), the Australian Cancer Research Foundation, The Wellcome Trust (WT084766/Z/08/Z) and donations from Neville and Shirley Hawkins.

Princess Alexandra Hospital

We thank A. Greene and M. Smither for establishment of the cohort and its funding, and C. Rowe and M Malt for patient recruitment and collection, and finally the patients themselves for their donation of support and time. This cohort was funded by the University of Queensland Diamantina Institute, the Meehan Foundation, NHMRC Career Development Fellowship (1125290) and Cancer Council Queensland (1125237).

AMFS

This cohort could not have been established without the support and involvement of participants, interviewers, data collection staff and project coordinators. The AMFS was funded by the NHMRC (APP566946, APP107359, APP211172, APP402761). We also recognize the funding from Cancer Councils of Victoria, Queensland and New South Wales (project grants 77/00, 06/10, 371) and by US NIH RO1 grants (CA-83115-01A2 and 2R01CA083115-11A1).

Western Australian Melanoma Health Study (WAMHS)

In addition to thanking the patients who took part, the WAMHS recognizes the essential work undertaken by members of the research teams, the WAMHS Management Committee, the Western Australian DNA Bank, the Ark at The University of Western Australia and the Western Australian Cancer Registry. Staff and students were supported by the Scott Kirkbride Melanoma Research Centre and the Cancer Council Western Australia.

Melanoma Institute Australia

In addition to the work and support of patients, and colleagues, at Melanoma Institute Australia (MIA), this cohort was enabled by funding from MIA, the NHMRC, New South Wales Department of Health, New South Wales Health Pathology, Cancer

Institute New South Wales and infrastructure grants from Macquarie University and the Australian Cancer Research Foundation. Genotyping was supported by Worldwide Cancer Research grant 16-0101. Support from the Cameron Family, the Ainsworth Foundation is also gratefully acknowledged.

Cambridge and Leeds

Participant recruitment was undertaken with the assistance of UK National Cancer Research, research assistant and nurses, as well as clinical and scientific collaborators. We also wish to specifically recognize the hard work of P. Mack, K. Gamble, P. King and A. Downing. Funding from Cancer Research UK (UK C490/A16561, C8216/A6129, C588/A4994) and by the NIH (R01 CA83115) is gratefully recognized.

EPIGENE

We recognize the hard work and assistance of participating melanoma patients, Sullivan and Nicolaides Pathology, Queensland Medical Laboratories and IQ Pathology. Funding, which is gratefully acknowledged, was from the NHMRC (APP442960).

MDACC

Data were downloaded from dbGAP, accession number phs000187.v1.p1. For details of acknowledgements for the MDACC study, please see Amos et al. (60).

Conflict of Interest statement. R.S. reports receiving fees for professional services from Merck Sharp & Dohme, GlaxoSmithKline Australia, Bristol-Myers Squibb, Dermopedia, Novartis Pharmaceuticals Australia Pty Ltd, Myriad, NeraCare GmbH and Amgen unrelated to this work. J.F.T. has received honoraria for advisory board participation from BMS Australia and MSD Australia. He has received honoraria and travel support from GlaxoSmithKline and Provectus Inc. G.V.L. is a consultant advisor for Aduro, Amgen, Bristol-Myers Squibb, Highlight Therapeutics S.L., Mass-Array, Merck, MSD, Novartis, OncoSec Medical, Pierre Fabre, Roche, QBiotech, Skyline DX and Sandoz.

References

1. Street, W. (2019) *Cancer Facts & Figures*. American Cancer Society, Atlanta, Vol. 71.
2. Australian Institute of Health and Welfare (2019) *Cancer in Australia*. AIHW, Vol. 119.
3. Balch, C.M., Gershenwald, J.E., Soong, S.-J., Thompson, J.F., Atkins, M.B., Byrd, D.R., Buzaid, A.C., Cochran, A.J., Coit, D.G., Ding, S. et al. (2009) Final version of 2009 AJCC melanoma staging and classification. *J Clin Oncol*, **27**, 6199–6206.
4. Gershenwald, J.E. and Scolyer, R.A. (2018) Melanoma staging: American joint committee on cancer (AJCC) 8th edition and beyond. *Ann Surg Oncol*, **25**, 2105–2110.
5. Breslow, A. (1970) Thickness, cross-sectional areas and depth of invasion in the prognosis of cutaneous melanoma. *Ann Surg*, **172**, 902–908.
6. Liu, W., Dowling, J.P., Murray, W.K., McArthur, G.A., Thompson, J.F., Wolfe, R. and Kelly, J.W. (2006) Rate of growth in melanomas: characteristics and associations of rapidly growing melanomas. *Arch Dermatol*, **142**, 1551–1558.
7. Nagore, E., Martorell-Calatayud, A., Botella-Estrada, R. and Guillén, C. (2011) Growth rate as an independent prognostic factor in localized invasive cutaneous melanoma. *J Eur Acad Dermatol Venereol*, **25**, 618–620.

8. Richard, M.A., Grob, J.J., Avril, M.F., Delaunay, M., Thirion, X., Wolkenstein, P., Souteyrand, P., Dreno, B., Bonerandi, J.J., Dalac, S. et al. (1999) Melanoma and tumor thickness: challenges of early diagnosis. *Arch Dermatol*, **135**, 269–274.
9. El Sharouni, M.A., Witkamp, A.J., Sigurdsson, V., van Diest, P.J., Louwman, M.W.J. and Kukutsch, N.A. (2019) Sex matters: men with melanoma have a worse prognosis than women. *J Eur Acad Dermatol Venereol*, **33**, 2062–2067.
10. Gibbs, D.C., Ward, S.V., Orlow, I., Cadby, G., Kanetsky, P.A., Luo, L., Busam, K.J., Kricke, A., Armstrong, B.K., Cust, A.E. et al. (2017) Functional melanoma-risk variant IRF4 rs12203592 associated with Breslow thickness: a pooled international study of primary melanomas. *Br J Dermatol*, **177**, e180–e182.
11. Liu, H., Wei, Q., Gershenwald, J.E., Prieto, V.G., Lee, J.E., Duvic, M., Grimm, E.A. and Wang, L.-E. (2012) Influence of single nucleotide polymorphisms in the MMP1 promoter region on cutaneous melanoma progression. *Melanoma Res*, **22**, 169–175.
12. Streit, S., Mestel, D.S., Schmidt, M., Ullrich, A. and Berking, C. (2006) FGFR4 Arg388 allele correlates with tumour thickness and FGFR4 protein expression with survival of melanoma patients. *Br J Cancer*, **94**, 1879–1886.
13. Davies, J.R., Randerson-Moor, J., Kukalich, K., Harland, M., Kumar, R., Madhusudan, S., Nagore, E., Hansson, J., Höiom, V., Ghiorzo, P. et al. (2012) Inherited variants in the MC1R gene and survival from cutaneous melanoma: a BioGenoMEL study. *Pigment Cell Melanoma Res*, **25**, 384–394.
14. Taylor, N.J., Busam, K.J., From, L., Groben, P.A., Anton-Culver, H., Cust, A.E., Begg, C.B., Dwyer, T., Gallagher, R.P., Gruber, S.B. et al. (2015) Inherited variation at MC1R and histological characteristics of primary melanoma. *PLoS ONE*, **10**, e0119920.
15. Orlow, I., Reiner, A.S., Thomas, N.E., Roy, P., Kanetsky, P.A., Luo, L., Paine, S., Armstrong, B.K., Kricke, A., Marrett, L.D. et al. (2016) Vitamin D receptor polymorphisms and survival in patients with cutaneous melanoma: a population-based study. *Carcinogenesis*, **37**, 30–38.
16. Randerson-Moor, J.A., Taylor, J.C., Elliott, F., Chang, Y.-M., Beswick, S., Kukalich, K., Affleck, P., Leake, S., Haynes, S., Karpavicius, B. et al. (2009) Vitamin D receptor gene polymorphisms, serum 25-hydroxyvitamin D levels, and melanoma: UK case-control comparisons and a meta-analysis of published VDR data. *Eur J Cancer*, **45**, 3271–3281.
17. Vaysse, A., Fang, S., Brossard, M., Wei, Q., Chen, W.V., Mohamdi, H., Vincent-Fetita, L., Margaritte-Jeannin, P., Lavielle, N., Maubec, E. et al. (2016) A comprehensive genome-wide analysis of melanoma Breslow thickness identifies interaction between CDC42 and SCIN genetic variants. *Int J Cancer*, **139**, 2012–2020.
18. Fang, S., Vaysse, A., Brossard, M., Wang, Y., Deng, D., Liu, Q., Zhang, P., Xu, K., Li, M., Feng, R. et al. (2017) Melanoma expression genes identified through genome-wide association study of Breslow tumor thickness. *J Invest Dermatol*, **137**, 253–257.
19. Li, W.-Q., Cho, E., Wu, S., Li, S., Matthews, N.H. and Qureshi, A.A. (2019) Host characteristics and risk of incident melanoma by Breslow thickness. *Cancer Epidemiol Biomark Prev*, **28**, 217–224.
20. Fisher, N.M., Schaffer, J.V., Berwick, M. and Bolognia, J.L. (2005) Breslow depth of cutaneous melanoma: impact of factors related to surveillance of the skin, including prior skin biopsies and family history of melanoma. *J Am Acad Dermatol*, **53**, 393–406.
21. Evans, L.M., Tahmasbi, R., Vrieze, S.I., Abecasis, G.R., Das, S., Gazal, S., Bjelland, D.W., de Candia, T.R., Haplotype Reference Consortium, Goddard, M.E. et al. (2018) Comparison of methods that use whole genome data to estimate the heritability and genetic architecture of complex traits. *Nat Genet*, **50**, 737–745.
22. Yang, J., Lee, S.H., Goddard, M.E. and Visscher, P.M. (2011) GCTA: a tool for genome-wide complex trait analysis. *Am J Hum Genet*, **88**, 76–82.
23. Yang, J., Bakshi, A., Zhu, Z., Hemani, G., Vinkhuyzen, A.A.E., Lee, S.H., Robinson, M.R., Perry, J.R.B., Nolte, I.M., van Vliet-Ostaptchouk, J.V. et al. (2015) Genetic variance estimation with imputed variants finds negligible missing heritability for human height and body mass index. *Nat. Gen.*, **47**, 1114–1120.
24. Speed, D., Cai, N., UCLEB Consortium, Johnson, R. M., Nejentsev, S. and Balding, D.J. (2017) Reevaluation of SNP heritability in complex human traits. *Nat Genet*, **49**, 986–992.
25. Carrera, C., Gual, A., Díaz, A., Puig-Butillé, J.A., Nogués, S., Vilalta, A., Conill, C., Rull, R., Vilana, R., Arguis, P. et al. (2017) Prognostic role of the histological subtype of melanoma on the hands and feet in Caucasians. *Melanoma Res*, **27**, 315–320.
26. Pruim, R.J., Welch, R.P., Sanna, S., Teslovich, T.M., Chines, P.S., Gliedt, T.P., Boehnke, M., Abecasis, G.R. and Willer, C.J. (2010) LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics*, **26**, 2336–2337.
27. Swetter, S.M., Pollitt, R.A., Johnson, T.M., Brooks, D.R. and Geller, A.C. (2012) Behavioral determinants of successful early melanoma detection: role of self and physician skin examination. *Cancer*, **118**, 3725–3734.
28. Uliasz, A. and Lebwohl, M. (2007) Patient education and regular surveillance results in earlier diagnosis of second primary melanoma. *Int J Dermatol*, **46**, 575–577.
29. Swetter, S.M., Johnson, T.M., Miller, D.R., Layton, C.J., Brooks, K.R. and Geller, A.C. (2009) Melanoma in middle-aged and older men: a multi-institutional survey study of factors related to tumor thickness. *Arch Dermatol*, **145**, 397–404.
30. Rowe, C.J., Law, M.H., Palmer, J.M., MacGregor, S., Hayward, N.K. and Khosrotehrani, K. (2015) Survival outcomes in patients with multiple primary melanomas. *J Eur Acad Dermatol Venereol*, **29**, 2120–2127.
31. Youlden, D.R., Baade, P.D., Soyer, H.P., Youl, P.H., Kimlin, M.G., Aitken, J.F., Green, A.C. and Khosrotehrani, K. (2016) Ten-year survival after multiple invasive melanomas is worse than after a single melanoma: a population-based study. *J Invest Dermatol*, **136**, 2270–2276.
32. Baumert, J., Plewig, G., Volkenandt, M. and Schmid-Wendtner, M.-H. (2007) Factors associated with a high tumour thickness in patients with melanoma. *Br J Dermatol*, **156**, 938–944.
33. Haenssle, H.A., Hoffmann, S., Holzkamp, R., Samhaber, K., Lockmann, A., Fliesser, M., Emmert, S., Schön, M.P., Rosenberger, A. and Buhl, T. (2015) Melanoma thickness: the role of patients' characteristics, risk indicators and patterns of diagnosis. *J Eur Acad Dermatol Venereol*, **29**, 102–108.
34. Karczewski, K.J., Francioli, L.C., Tiao, G., Cummings, B.B., Alföldi, J., Wang, Q., Collins, R.L., Laricchia, K.M., Ganna, A., Birnbaum, D.P. et al. (2020) The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature*, **581**, 434–443.
35. Watanabe, K., Taskesen, E., van Bochoven, A. and Posthuma, D. (2017) Functional mapping and annotation of genetic associations with FUMA. *Nat Commun*, **8**, 1826.

36. GTEx Consortium, Laboratory, Data Analysis & Coordinating Center (LDACC), Analysis Working Group, Statistical Methods groups—Analysis Working Group, Enhancing GTEx (eGTEx) groups, NIH Common Fund, NIH/NCI, NIH/NHGRI, NIH/NIMH, NIH/NIDA, Biospecimen Collection Source Site—NDRI, Biospecimen Collection Source Site—NDRI et al. (2017) Genetic effects on gene expression across human tissues. *Nature*, **550**, 204–213.
37. Carvalho-Silva, D., Pierleoni, A., Pignatelli, M., Ong, C., Fumis, L., Karamanis, N., Carmona, M., Faulconbridge, A., Hercules, A., McAuley, E. et al. (2019) Open targets platform: new developments and updates two years on. *Nucleic Acids Res*, **47**, D1056–D1065.
38. Ward, L.D. and Kellis, M. (2012) HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucleic Acids Res*, **40**, D930–D934.
39. Noyes, N.C., Hampton, B., Migliorini, M. and Strickland, D.K. (2016) Regulation of itch and Nedd4 E3 ligase activity and degradation by LRAD3. *Biochemistry*, **55**, 1204–1213.
40. Duffy, D.L., Zhu, G., Li, X., Sanna, M., Iles, M.M., Jacobs, L.C., Evans, D.M., Yazar, S., Beesley, J., Law, M.H. et al. (2018) Novel pleiotropic risk loci for melanoma and nevus density implicate multiple biological pathways. *Nat Commun*, **9**, 4774.
41. Law, M.H., Bishop, D.T., Lee, J.E., Brossard, M., Martin, N.G., Moses, E.K., Song, F., Barrett, J.H., Kumar, R., Easton, D.F. et al. (2015) Genome-wide meta-analysis identifies five new susceptibility loci for cutaneous malignant melanoma. *Nat Genet*, **47**, 987–995.
42. Duffy, D.L., Iles, M.M., Glass, D., Zhu, G., Barrett, J.H., Höiö, V., Zhao, Z.Z., Sturm, R.A., Soranzo, N., Hammond, C. et al. (2010) IRF4 variants have age-specific effects on nevus count and predispose to melanoma. *Am J Hum Genet*, **87**, 6–16.
43. Kvaskoff, M., Whiteman, D.C., Zhao, Z.Z., Montgomery, G.W., Martin, N.G., Hayward, N.K. and Duffy, D.L. (2011) Polymorphisms in nevus-associated genes MTAP, PLA2G6, and IRF4 and the risk of invasive cutaneous melanoma. *Twin Res Hum Genet*, **14**, 422–432.
44. Bossé, Y. and Amos, C.I. (2018) A decade of GWAS results in lung cancer. *Cancer Epidemiol Biomark Prev*, **27**, 363–379.
45. Escala-Garcia, M., Guo, Q., Dörk, T., Canisius, S., Keeman, R., Dennis, J., Beesley, J., Lecarpentier, J., Bolla, M.K., Wang, Q. et al. (2019) Genome-wide association study of germline variants and breast cancer-specific mortality. *Br J Cancer*, **120**, 647–657.
46. Baade, P.D., English, D.R., Youl, P.H., McPherson, M., Elwood, J.M. and Aitken, J.F. (2006) The relationship between melanoma thickness and time to diagnosis in a large population-based study. *Arch Dermatol*, **142**, 1422–1427.
47. Aschard, H., Vilhjálmsson, B.J., Joshi, A.D., Price, A.L. and Kraft, P. (2015) Adjusting for heritable covariates can bias effect estimates in genome-wide association studies. *Am J Hum Genet*, **96**, 329–339.
48. Day, F.R., Loh, P.-R., Scott, R.A., Ong, K.K. and Perry, J.R.B. (2016) A robust example of collider bias in a genetic association study. *Am J Hum Genet*, **98**, 392–393.
49. Skowron, F., Bérard, F., Balme, B. and Maucort-Boulch, D. (2015) Role of obesity on the thickness of primary cutaneous melanoma. *J Eur Acad Dermatol Venereol*, **29**, 262–269.
50. Gandini, S., Montella, M., Ayala, F., Benedetto, L., Rossi, C.R., Vecchiato, A., Corradin, M.T., DE Giorgi, V., Queirolo, P., Zannetti, G. et al. (2016) Sun exposure and melanoma prognostic factors. *Oncol Lett*, **11**, 2706–2714.
51. Chang, C.C., Chow, C.C., Tellier, L.C., Vattikuti, S., Purcell, S.M. and Lee, J.J. (2015) Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience*, **4**, 7.
52. Core Team, R. (2013) *R: A language and environment for statistical computing*. R Foundation for statistical computing, Vienna.
53. Marees, A.T., de Kluiver, H., Stringer, S., Vorspan, F., Curis, E., Marie-Claire, C. and Derks, E.M. (2018) A tutorial on conducting genome-wide association studies: quality control and statistical analysis. *Int J Methods Psychiatr Res*, **27**, e1608.
54. 1000 Genomes Project Consortium, Auton, A., Brooks, L.D., Durbin, R.M., Garrison, E.P., Kang, H.M., Korbel, J.O., Marchini, J.L., McCarthy, S., McVean, G.A. et al. (2015) A global reference for human genetic variation. *Nature*, **526**, 68–74.
55. McCarthy, S., Das, S., Kretzschmar, W., Delaneau, O., Wood, A.R., Teumer, A., Kang, H.M., Fuchsberger, C., Danecek, P., Sharp, K. et al. (2016) A reference panel of 64,976 haplotypes for genotype imputation. *Nat Genet*, **48**, 1279–1283.
56. Loh, P.-R., Danecek, P., Palamara, P.F., Fuchsberger, C., A Reshef, Y., K Finucane, H., Schoenherr, S., Forer, L., McCarthy, S., Abecasis, G.R. et al. (2016) Reference-based phasing using the haplotype reference consortium panel. *Nat. Gen*, **48**, 1443–1448.
57. Wainschtein, P., Jain, D.P., Yengo, L., Zheng, Z., TOPMed Anthropometry Working Group, Trans-Omics for Precision Medicine Consortium, Adrienne Cupples, L., Shadyab, A.H., McKnight, B., Shoemaker, B.M., Mitchell, B.D. et al. (2019) Recovery of trait heritability from whole genome sequence data. *bioRxiv*, 588020.
58. Viechtbauer, W. (2010) Conducting meta-analyses in R with the metafor package. *J Stat Softw*, **36**, 1–48.
59. Marchini, J., Cardon, L.R., Phillips, M.S. and Donnelly, P. (2004) The effects of human population structure on large genetic association studies. *Nat Genet*, **36**, 512–517.
60. Amos, C.I., Wang, L.-E., Lee, J.E., Gershenwald, J.E., Chen, W.V., Fang, S., Kosoy, R., Zhang, M., Qureshi, A.A., Vattathil, S. et al. (2011) Genome-wide association study identifies novel loci predisposing to cutaneous melanoma. *Hum Mol Genet*, **20**, 5012–5023.