



Genome-wide association meta-analyses combining multiple risk phenotypes provide insights into the genetic architecture of cutaneous melanoma susceptibility

Most genetic susceptibility to cutaneous melanoma remains to be discovered. Meta-analysis genome-wide association study (GWAS) of 36,760 cases of melanoma (67% newly genotyped) and 375,188 controls identified 54 significant ($P < 5 \times 10^{-8}$) loci with 68 independent single nucleotide polymorphisms. Analysis of risk estimates across geographical regions and host factors suggests the acral melanoma subtype is uniquely unrelated to pigmentation. Combining this meta-analysis with GWAS of nevus count and hair color, and transcriptome association approaches, uncovered 31 potential secondary loci for a total of 85 cutaneous melanoma susceptibility loci. These findings provide insights into cutaneous melanoma genetic architecture, reinforcing the importance of neovogenesis, pigmentation and telomere maintenance, together with identifying potential new pathways for cutaneous melanoma pathogenesis.

Cutaneous melanoma is a deadly malignancy with increasing incidence and burden in fair-skinned populations worldwide¹. Increased risk for cutaneous melanoma is caused by high exposure to ultraviolet radiation² as well as host factors, including family history^{3,4}, pigimentary phenotypes⁵, number of melanocytic nevi^{6,7}, longer telomeres^{8,9} and immunosuppression¹⁰.

Identified melanoma genetic risk variants include rare, highly penetrant mutations in genes such as *CDKN2A*^{11,12} and *POT1* (refs. ^{13,14}), as well as more common variants (for example, lower-penetrance variants in *MC1R*)^{15,16}. GWAS of cutaneous melanoma susceptibility in populations of European ancestry have identified 21 genetic loci reaching genome-wide significance ($P < 5 \times 10^{-8}$) (refs. ^{17–24}). Additional approaches, including family-based analyses of cutaneous melanoma^{25,26}, combining cutaneous melanoma and nevus count GWAS²⁷ and transcriptome-wide association studies (TWAS)²⁸ have identified further loci that, despite not containing single nucleotide polymorphisms (SNPs) reaching $P < 5 \times 10^{-8}$ in a cutaneous melanoma-only GWAS, most likely influence melanoma risk.

This meta-analysis of cutaneous melanoma susceptibility is powered to identify cutaneous melanoma susceptibility variants, and it provides enhanced distinction of independent variants in known cutaneous melanoma susceptibility regions. We report here 68 independent cutaneous melanoma-associated variants across 54 loci that confirm the importance of key functional pathways and highlight previously unknown cutaneous melanoma etiologic routes (Tables 1 and 2). Stratified analyses showed a lack of involvement of the pigmentation pathway for acral melanoma, in line with observational data²⁹. The combined analysis of cutaneous melanoma, nevus and hair color GWAS data, and use of expression data through TWAS, identified 31 secondary, potential loci.

Results

Study overview. We performed a GWAS meta-analysis of cutaneous melanoma susceptibility with 30,134 clinically confirmed cases of cutaneous melanoma (Methods), 6,626 cases of self-reported cutaneous melanoma and 375,188 cutaneous melanoma-free controls from the United Kingdom, the United States, Australia and

Northern and Western Europe, as well as the Mediterranean—a highly sun-exposed population often under-represented in cutaneous melanoma studies (Supplementary Table 1). Of these, 24,756 cases (67%) and 358,734 controls (96%) had not been included in any previous melanoma GWAS.

Separately, we performed total (clinically confirmed cases + self-reported cases from 23andMe and a subset of UK Biobank cases with only self-reported cutaneous melanoma status) and confirmed-only cutaneous melanoma meta-analyses to determine the power gained by including cases of self-reported cutaneous melanoma. Risk loci were deemed genome-wide significant when variants had fixed-effects meta-analysis $P < 5 \times 10^{-8}$ (P_{meta}); where variants exhibited notable heterogeneity ($P > 31\%$)³⁰ random-effects P values ($P_{\text{meta,r}}$) were also required to be $< 5 \times 10^{-8}$ (Methods). Quantile–quantile (Q–Q) plots (Supplementary Fig. 1) and linkage disequilibrium score regression³¹ (LDSC) (Methods) intercepts showed minimal inflation for individual studies (mostly < 1.04 ; Supplementary Table 1), indicating adequate control of population stratification.

Before including the self-report GWAS data, we used LDSC³¹ to verify their genetic correlation (R_g) with the confirmed-only GWAS meta-analysis (Supplementary Note and Supplementary Table 2). Based on the high R_g and similarity in SNP heritability (h^2) estimates for self-reported and clinically confirmed cases of cutaneous melanoma (Supplementary Note), we performed an overall total cutaneous melanoma meta-analysis ($h^2_{\text{total}} = 0.085$, 95% confidence interval (CI) = 0.05–0.12). The genomic inflation (λ) and LDSC intercept for the total cutaneous melanoma meta-analysis indicated that most inflation is due to polygenic signal ($\lambda = 1.165$, intercept = 1.054, ratio = 0.17; Supplementary Table 2). A similar h^2_{total} (12%) was estimated using genetic effect-size distribution inference from summary-level data (GENESIS) (Methods)³².

Conditional- and joint-analyses of the total cutaneous melanoma meta-analysis summary statistics using GCTA³³ identified a total of 54 loci meeting our requirements for genome-wide significance (Methods; Fig. 1 and Extended Data Figs. 1 and 2). Results for loci previously reported by cutaneous melanoma GWAS

Table 1 | Loci previously identified in cutaneous melanoma susceptibility GWAS

CHR:BP	rsID	Refs.	Genes	EA/ NEA	Freq	P_{meta}	OR	Nevi	Hair	Tan	Telo
1:150,938,571	rs8444	21	Multiple	G/A	0.645	3.89×10^{-14}	1.08	-	-	Y	-
1:226,603,635	rs2695237	21, 27, 88	PARP1	T/C	0.628	1.53×10^{-18}	1.10	Y	-	-	-
2:38,298,139	rs1800440	23, 27	CYP1B1	T/C	0.824	6.97×10^{-15}	1.10	Y	-	Y	-
2:202,143,928	rs10931936 ^a	20	CASP8	T/C	0.281	2.17×10^{-8}	1.08	-	-	-	-
5:1,323,212	rs13178866 ^a	20, 89, 90	TERT	C/T	0.554	2.59×10^{-18}	0.87	-	Y	-	Y
5:33,951,693	rs16891982 ^a	20, 34, 90	SLC45A2	C/G	0.122	1.96×10^{-28}	0.51	-	Y	Y	-
6:21,163,919	rs6914598	23	CDKAL1	T/C	0.683	1.18×10^{-18}	0.91	-	-	Y	-
7:17,134,708	rs117132860 ^b	23, 43	AGR3	G/A	0.981	3.83×10^{-21}	0.71	Y	-	Y	-
9:21,803,880	rs871024 ^a	18, 27	MTAP, CDKN2A	C/A	0.477	2.72×10^{-23}	1.18	Y	Y	-	-
9:109,054,417	rs10739220	23, 27	TMEM38B	C/T	0.260	1.34×10^{-18}	1.10	Y	Y	-	-
10:105,694,301	rs7902587	23, 27	OBFC1	C/T	0.904	2.68×10^{-23}	0.86	Y	-	-	Y
11:69,380,898	rs4354713	20, 23	CCND1	A/G	0.356	8.50×10^{-21}	1.10	-	Y	-	-
11:89,017,961	rs1126809 ^a	18	TYR	G/A	0.757	4.78×10^{-37}	0.83	-	Y	Y	-
11:108,175,462	rs1801516	20	ATM	G/A	0.856	2.22×10^{-21}	1.14	Y	-	-	-
15:28,365,618	rs12913832 ^a	19, 23	OCA2	A/G	0.335	4.85×10^{-12}	0.88	-	Y	Y	-
16:89,986,117	rs1805007 ^a	18	MC1R	C/T	0.937	5.86×10^{-52}	0.57	Y	Y	Y	-
20:32,665,748	rs6059655 ^a	17, 18	ASIP	A/G	0.061	2.52×10^{-42}	1.45	-	Y	Y	-
21:42,743,496	rs408825	20	MX2	C/T	0.413	1.03×10^{-32}	0.89	-	-	Y	-
22:38,545,942	rs132941	18, 27, 35	MAFF	T/C	0.549	8.80×10^{-23}	1.10	Y	-	Y	-

CHR:BP, hg19 positional information. rsID, dbSNP142 rs number. Refs., related publications. Genes, prioritizing the functional target if known, followed by melanocyte or skin tissue TWAS data, or finally the closest protein coding gene; 'multiple' indicates three or more genes (Supplementary Table 3). The effect allele (EA) and noneffect allele (NEA) are listed, as are the effect allele frequencies (Freq) in the HRC reference panel⁷¹. Total fixed-effects inverse-variance weighted meta-analysis of logistic regression two-sided P value (P_{meta}) and odds ratio (OR) are from an additive model and are reported per-allele with respect to the EA. Reported results are for the total meta-analysis (36,760 cases of melanoma and 375,188 controls; for full details of analysis and covariates included see Methods). We also indicate (yes, Y; no, -) whether this locus is associated with other traits: Nevi, pleiotropically associated with cutaneous melanoma and nevus count (Methods; Supplementary Table 9). Hair, pleiotropically associated with cutaneous melanoma and hair color (Methods; Supplementary Table 10). Tanning response (Tan) and Telomere length (Telo) indicate that the lead SNP is associated with these traits when corrected for multiple testing (Methods; Supplementary Table 5). ^aVariant meta-analysis results are heterogeneous ($I^2 > 31\%$) and random-effects estimates are presented. ^bWhile this locus overlaps the previously reported *IRF4* or *AGR3* locus, the lead variants are independent.

reaching significance in the total meta-analysis are presented in Table 1. Results for loci not previously reported by a cutaneous melanoma GWAS are summarized in Table 2. In addition to the 54 lead variants, 14 independent variants with linkage disequilibrium (LD) $r^2_{EUR} < 0.05$ with lead variants at or near six loci (*TERT*, *AGR3*, *CDKN2A*, *OCA2*, *MC1R* and *TP53*) were identified (Supplementary Table 3). Individual regional association plots for the association signals have been provided in Supplementary Data 1. Conditional- and joint-analyses of summary data identified a further nine variants at or near *SLC45A2*, *IRF4*, *AGR3*, *CCND1*, *GPRC5A*, *FTO* and *MC1R*; however, these additional variants were not carried forward, having $P_{meta} > 5 \times 10^{-8}$ in the single variant analysis or excess heterogeneity ($I^2 > 31\%$) and $P_{meta,r} < 5 \times 10^{-8}$ (Supplementary Table 4). In addition, we used GENESIS (Methods), which enables a reformulation of the variance explained by associated SNPs to estimate a theoretical optimal area under the curve (AUC), rather than formally testing this using a training and prediction set³² to estimate the potential AUC. The estimated AUC was 66.8%, compared to ~64% in the 2015 cutaneous melanoma meta-analysis²³. This estimate does not include any host factors and would require benchmarking in a prospective study for validation.

Previous cutaneous melanoma GWAS have identified 21 genome-wide-significant loci^{17–24}. Family-based methods or the combination of cutaneous melanoma with nevus count have identified a further 12 loci including *IRF4*, *MITF*, *HDAC4* and *GPRC5A*^{25–27}. The lead SNPs from many of these loci are associated with pigmentation, tanning response, nevus count and telomere maintenance (Supplementary Table 5). Other SNPs are proximal to DNA repair genes. Some loci are associated with more than one trait (Tables 1

and 2). Our analysis confirms 19 of the 21 loci previously reaching genome-wide significance (Table 1 and Supplementary Note). The total cutaneous melanoma meta-analysis also confirms the previously reported *IRF4* and *MITF* associations^{25–27,34,35}, as well as six regions previously identified only by combining nevus count and cutaneous melanoma GWAS data²⁷ (Table 2 and Supplementary Note). These results demonstrate the ability of cross-trait GWAS to identify disease loci. The remaining 27 loci have not previously been reported as cutaneous melanoma susceptibility loci (Table 2 and Supplementary Table 3). The results for the pathologically confirmed-only cases of cutaneous melanoma ($n=30,134$; Supplementary Table 1) are reported in the Supplementary Note, Extended Data Figs. 2 and 3 and Supplementary Table 6. Our full meta-analysis identified 11 loci not found in the confirmed-only GWAS meta-analysis, demonstrating the advantage of including the 6,626 self-reported cases of cutaneous melanoma and over 290,000 controls (Supplementary Table 1). Results for SNPs with a fixed or random $P < 5 \times 10^{-7}$, from the total meta-analysis are reported in Supplementary Table 7.

Melanoma associations by sex, age at diagnosis and subtype. We performed separate GWAS by sex, age at cutaneous melanoma diagnosis (≤ 40 , 40–60 and ≥ 60 yr) and major cutaneous melanoma subtypes (superficial spreading, lentigo maligna, nodular melanoma and acral lentiginous) to identify variants associated with select subgroups (Supplementary Table 8). Our analysis identified no additional variants after adjustment for multiple testing ($5 \times 10^{-8}/9$ tests), suggesting that if such variants exist they are undetectable at our current sample size.

Table 2 | New loci not previously identified in cutaneous melanoma GWAS

CHR:BP	rsID	Genes	EA/NEA	Freq	P_{meta}	OR	Nevi	Hair	Tan	Telo
1:63,727,542	rs670318	FOXD3	T/C	0.047	1.21×10^{-8}	0.86	-	-	Y	-
1:154,994,978	rs76798800	ZBTB7B, ADAM15, GBA	G/T	0.753	3.86×10^{-15}	0.92	Y	-	Y	-
1:205,181,062	rs2369633	DSTYK	T/C	0.083	1.24×10^{-8}	1.10	-	e	Y	-
2:25,778,637	rs12473635	DTNB	T/C	0.776	5.17×10^{-9}	0.93	Y	-	-	-
3:70,014,091	rs149617956 ^a	MITF	G/A	0.998	9.00×10^{-14}	0.39	-	Y	Y	-
3:169,493,283	rs3950296 ^b	TERC	C/G	0.747	4.47×10^{-11}	1.08	Y	-	-	Y
5:90,262,612	rs12523094 ^c	GPR98	T/C	0.567	1.74×10^{-6} c	1.07	-	Y	Y	-
5:149,211,868	rs32578 ^{b,d}	PPARGC1B	G/A	0.658	6.58×10^{-17}	1.09	Y	-	Y	-
6:1,145,265	rs12215602 ^a	IRF4	G/A	0.721	7.91×10^{-9}	0.94	Y	-	Y	-
6:22,719,379	rs72834823	HDGFL1	T/A	0.819	1.04×10^{-12}	1.10	Y	-	Y	-
6:32,748,953	rs28986343	HLA-DQB2	C/T	0.952	1.61×10^{-8}	1.15	-	-	-	-
6:91,005,743	rs6908626	BACH2	G/T	0.844	3.92×10^{-9}	1.09	-	-	-	-
7:22,115,454	rs12539524	RAPGEF5	C/T	0.846	1.65×10^{-8}	0.93	-	-	-	-
7:124,396,645	rs4731207	POT1	G/A	0.540	2.22×10^{-15}	0.93	Y	-	-	Y
7:130,738,666	rs7778378	MKLN1	C/T	0.248	8.93×10^{-9}	0.93	Y	Y	-	-
8:21,951,009	rs6994183	FAM160B2	A/T	0.866	4.84×10^{-9}	0.92	-	-	-	-
8:72,864,240	rs13263376 ^c	RP11-383H13.1, MSC	G/A	0.364	2.28×10^{-8} c	0.93	Y	-	Y	-
9:12,587,153	rs10960710	TYRP1	G/T	0.393	3.08×10^{-12}	0.93	-	Y	Y	-
9:110,711,586	rs1339759 ^b	KLF4	C/G	0.666	5.61×10^{-19}	1.10	Y	-	-	-
9:134,457,580	rs3780269	RAPGEF1	G/A	0.691	1.92×10^{-8}	0.94	Y	-	-	-
11:16,041,305	rs7941496	SOX6	G/T	0.516	1.40×10^{-9}	1.06	Y	-	Y	-
11:120,195,702	rs12290699	TMEM136	T/C	0.745	2.20×10^{-8}	0.94	-	-	-	-
12:13,070,752	rs1056927 ^{b,c}	Multiple	A/G	0.561	2.74×10^{-9} b	0.93	Y	-	-	-
12:17,275,460	rs4237963	LMO3	T/A	0.207	1.27×10^{-9}	0.93	-	-	-	-
12:96,378,807	rs10859996	HAL, RP11-256L6.3	C/T	0.635	2.09×10^{-10}	1.07	-	-	-	-
12:116,580,291	rs113469387	MED13L	G/A	0.907	8.76×10^{-10}	0.91	-	Y	Y	-
13:113,535,949	rs1278768	MCF2L	G/C	0.488	6.33×10^{-12}	0.94	-	-	Y	-
15:33,277,710	rs117648907 ^b	FNMI	C/T	0.983	7.29×10^{-12}	0.80	Y	-	-	-
16:68,822,971	rs4420522	Multiple, CDH1	A/G	0.690	8.34×10^{-14}	0.93	Y	Y	-	-
16:82,217,153	rs2967383	MPHOSPH6	G/T	0.267	2.24×10^{-9}	1.06	-	-	-	Y
17:7,571,752	rs78378222	TP53	T/G	0.989	3.33×10^{-10}	0.76	Y	-	-	-
19:3,540,539	rs12984831 ^b	MFS12	G/C	0.984	3.86×10^{-10}	0.65	Y	-	Y	-
20:62,291,767	rs143190905	RETL1	G/T	0.907	6.54×10^{-13}	1.15	-	-	-	Y
22:45,622,684	rs5766565	KIAA0930	A/G	0.647	1.44×10^{-9}	1.06	Y	Y	Y	-
22:50,722,408	rs79966207	PLXNB2	T/C	0.849	8.68×10^{-9}	0.92	-	Y	-	-

Abbreviations and explanations as for Table 1. ^aAssociated with cutaneous melanoma by non-GWAS-based approaches—*MITF*^{25,26} and *IRF4* (refs. 27,34,35). ^bPreviously associated pleiotropically with cutaneous melanoma and nevus count²⁷. ^cVariant meta-analysis results are heterogeneous ($P > 31\%$) and random-effects estimates are presented. For rs12523094/*GPR98* while the lead SNP selected in conditional mapping is heterogeneous, other SNPs in LD pass this requirement (for example, rs12173258, $r^2_{EUR} = 0.9$, $P_{meta} = 1.09 \times 10^{-11}$, $P = 29.6$). ^dPreviously associated with tanning response⁴³. ^eJoint cutaneous melanoma + hair color P value is greater than multiple-testing corrected threshold of 1.25×10^{-8} (Supplementary Table 10).

We also performed polygenic risk score (PRS) analyses based on the lead independent genome-wide-significant SNPs for nevus count (ten variants; Methods) and hair color (276 variants; Methods) to explore further whether the association of either of these traits with cutaneous melanoma differs across phenotypic subtypes (significance threshold = 0.05/28; Methods). We observed no significant differences in the distribution of the tested PRSs by sex or age at cutaneous melanoma diagnosis. We did, however, detect differences in the distribution of the hair color PRS for the acral lentiginous subtype compared to all nonacral subtypes ($P = 2.1 \times 10^{-4}$). Our analyses indicated that genetically predicted pigmentation in cases of acral lentiginous subtype was no different from that in controls

($P = 0.65$; Extended Data Fig. 4) and darker than that in cases of the subtypes superficial spreading, lentigo maligna and nodular melanoma ($P = 5.3 \times 10^{-5}$, 0.01 and 4.8×10^{-4} , respectively). These findings provide strong genetic evidence that the pigmentation pathway is far less important for risk of acral lentiginous melanoma than for other subtypes of cutaneous melanoma. No significant differences were observed by subtype for the nevus count PRS.

Variant annotation with cutaneous melanoma risk phenotypes.

To investigate possible biological pathways underlying cutaneous melanoma signals, variants independently associated with cutaneous melanoma in the total meta-analysis were evaluated in GWAS

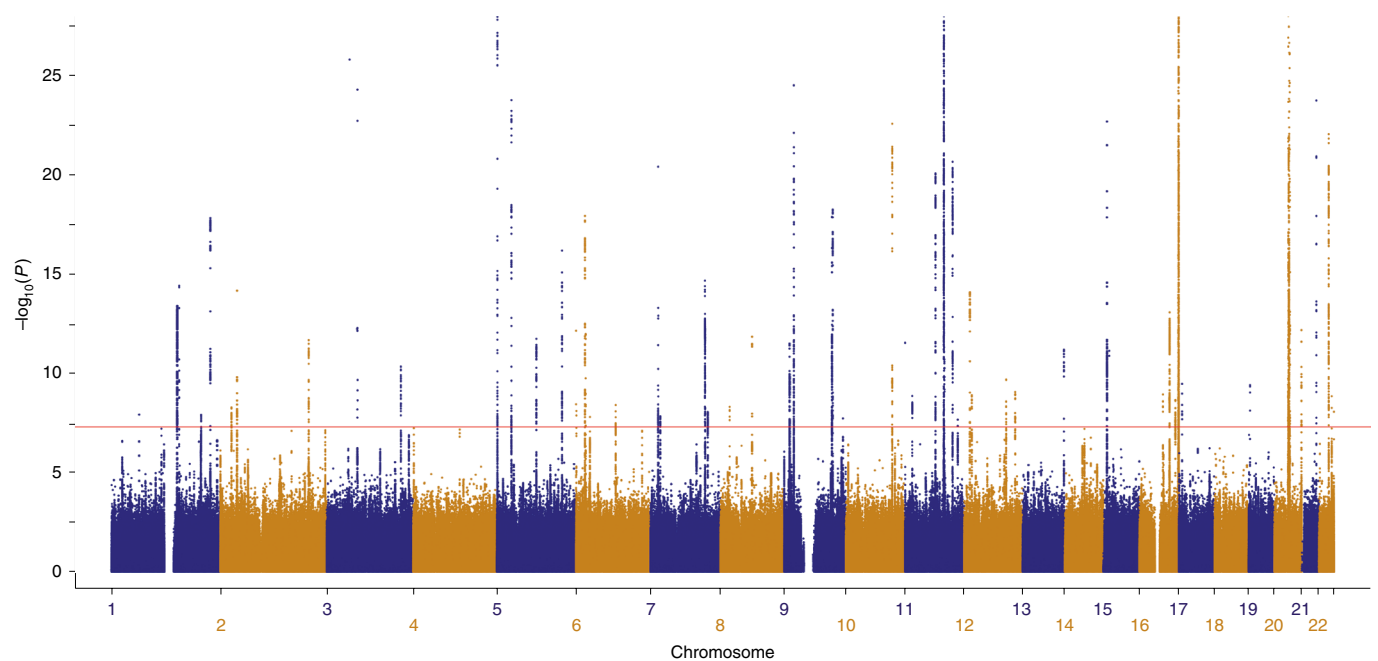


Fig. 1 | Manhattan plot for the total cutaneous melanoma meta-analysis. The $-\log_{10}$ of two-sided P values for SNPs derived from a fixed-effects inverse-variance weighted meta-analysis of logistic regression GWAS (y axis) plotted against SNP chromosome positions for the total meta-analysis (36,760 cases of melanoma and 375,188 controls; for full details of analysis and covariates included see Methods). The y axis is limited to $-\log_{10}(1 \times 10^{-25})$ to truncate strong signals at loci such as *MC1R* and *ASIP*. The full plot is displayed in Extended Data Fig. 2. To account for multiple testing, SNPs with $P < 5 \times 10^{-8}$ are deemed significant.

of telomere length, tanning response, pigmentation and nevus count (Methods; Tables 1 and 2 and Supplementary Tables 5, 7 and 8). Using a Bonferroni-corrected threshold of phenotype $P < 0.00074$ (0.05/68 independent SNPs), 18 of the 35 new loci are associated with tanning response or pigmentation (Table 2 and Supplementary Table 5), further indicating the importance of pigmentation pathways in cutaneous melanoma susceptibility. Several new loci, including rs12473635 near *DTNB* and rs78378222 near *TP53*, are associated with nevus count, reinforcing the role of nevi in cutaneous melanoma susceptibility. Furthermore, four new loci have previously been associated with telomere length (rs3950296/*TERC*, rs4731207/*POT1*, rs2967383/*MPHOSPH6* and rs143190905/*RTEL1*, ref. ³⁶) (Table 2 and Supplementary Table 5) providing additional support for the role of telomere maintenance in cutaneous melanoma susceptibility following earlier findings that genetic determinants of telomere length are generally associated with melanoma risk^{13,14,37}. Other newly discovered lead variants are not associated with these phenotypes, suggesting new pathways.

Additional approaches to identify melanoma risk loci. To identify further loci influencing cutaneous melanoma risk and provide a more nuanced annotation of discovered cutaneous melanoma risk loci, we used a range of secondary approaches with correction for multiple testing (Methods). To explore the overlap between cutaneous melanoma loci and established risk factor phenotypes, we combined our total cutaneous melanoma GWAS meta-analysis with a nevus count GWAS meta-analysis ($n = 65,777$; Methods) and separately with a UK Biobank hair color GWAS ($n = 352,662$; Methods). For the total cutaneous melanoma GWAS meta-analysis and nevus count the R_g is 0.57 (s.e.m. = 0.11, $P = 2.39 \times 10^{-7}$) and for hair color scored from light hair to dark (Methods) the R_g is 0.290 (s.e.m. = 0.096, $P = 0.0025$). Pairwise GWAS (GWAS-PW)³⁸ was used to determine whether loci were associated with only one trait or pleiotropic with both cutaneous melanoma and either nevus count or hair color (Methods). Loci previously reported through the

combination of cutaneous melanoma and nevus GWAS²⁷ are now confirmed by our larger cutaneous melanoma GWAS meta-analysis (Table 2). Together these analyses identified secondary potential loci not associated at genome-wide significance levels in the total cutaneous melanoma GWAS meta-analysis. At the Bonferroni-corrected threshold of 1.25×10^{-8} (Methods), they included eight loci jointly significant for cutaneous melanoma and nevus count, 17 for cutaneous melanoma and hair color, and four with cutaneous melanoma, nevus count and hair color (Table 3 and Supplementary Tables 9 and 10).

In parallel, we examined data from a recently established cell-type-specific melanocyte *cis*-expression quantitative trait loci (eQTL) dataset²⁸ as well as tissue-based *cis*-eQTL datasets available through genotype-tissue expression (GTEx)³⁹ resource to identify additional susceptibility loci using a transcriptome prediction mapping strategy (or TWAS)^{40,41}. TWAS using these expression datasets enabled gene-based testing for significant *cis* genetic correlations between imputed gene expression and cutaneous melanoma risk, aiding identification of additional susceptibility loci (Methods). While identification of significant genes by TWAS does not establish causation, it can indicate plausible gene candidates to be used in pathway analyses and investigated in future functional studies. This analysis built on a previous melanocyte TWAS that analysed data from a prior cutaneous melanoma GWAS meta-analysis²⁸ and identified significant new associations between cutaneous melanoma and imputed gene expression of five genes at four loci. Importantly, the *CBWD1* locus on chromosome 9 was later identified as a genome-wide-significant cutaneous melanoma + nevus count pleiotropic locus²⁷ (Table 3 and Supplementary Table 9) and the other three loci (*ZFP90* on chromosome 16, *HEBP1* on chromosome 12 and *MSC/RP11-383H13.1* on chromosome 8) are now at genome-wide significance with cutaneous melanoma in this larger GWAS meta-analysis (Table 2). This confirmation supports the TWAS approach for both identifying new loci and nominating potentially functional genes at GWAS-discovered loci (Tables 1 and 2).

Table 3 | New pleiotropic associations with cutaneous melanoma and nevus count or hair color

CHR:BP	rsID	Genes	CM P	CM + Nevus P	CM + Hair P
1:24,787,947	rs195720	<i>NIPAL3</i>	7.97×10^{-6}	-	2.24×10^{-12}
1:78,450,517	rs34517439	<i>DNAJB4</i>	2.23×10^{-4}	-	2.17×10^{-12}
1:214,673,271	rs7533482	<i>PTPN14</i>	2.79×10^{-5}	-	2.45×10^{-13}
2:135,430,709	rs6745983	<i>TMEM163</i>	1.69×10^{-3}	-	7.00×10^{-13}
2:214,065,880	rs16849932	<i>IKZF2</i>	1.46×10^{-3}	-	1.18×10^{-10}
2:240,065,356	rs11677464 ^a	<i>HDAC4</i>	4.00×10^{-5}	1.10×10^{-9}	-
4:37,470,753	rs11730662	<i>KIAA1239</i>	1.82×10^{-3}	1.19×10^{-8}	-
5:56,011,357	rs7714232	<i>MAP3K1</i>	6.99×10^{-4}	-	3.32×10^{-22}
6:7,189,567	rs75818295	<i>RREB1</i>	1.87×10^{-3}	-	8.27×10^{-10}
6:11,637,483	rs548304	<i>ADTRP</i>	2.67×10^{-5}	-	1.46×10^{-10}
6:15,503,696	rs10949304	<i>DTNBP1</i>	1.7×10^{-3}	4.96×10^{-9}	-
6:50,790,642	rs2857482	<i>TFAP2B</i>	3.59×10^{-5}	3.44×10^{-10}	-
6:151,577,739	rs10434895	<i>AKAP12</i>	8.17×10^{-8}	-	2.07×10^{-42}
6:151,577,830	rs10434896 ^b	<i>AKAP12</i>	7.88×10^{-8}	7.71×10^{-10}	-
8:131,138,979	rs111595456	<i>ASAP1</i>	3.86×10^{-4}	2.83×10^{-10}	-
9:211,762	rs520015	<i>CBWD1</i>	8.95×10^{-7}	-	1.10×10^{-43}
9:235,201	rs593179 ^{a,c}	<i>CBWD1</i>	3.78×10^{-6}	4.13×10^{-12}	-
10:5,767,177	rs76154345 ^a	<i>GDI2</i>	4.43×10^{-6}	7.80×10^{-11}	-
10:111,889,779	rs11194997	<i>MXI1</i>	3.45×10^{-6}	-	2.70×10^{-11}
11:7,543,519	rs11041426	<i>PPFIBP2</i>	2.73×10^{-4}	-	1.66×10^{-33}
11:62,203,865	rs10897275	<i>AHNAK</i>	6.47×10^{-5}	-	2.47×10^{-33}
11:91,616,691	rs12225068	<i>FAT3</i>	3.80×10^{-5}	-	6.48×10^{-10}
13:76,351,286	rs474240	<i>LMO7</i>	2.53×10^{-4}	-	9.28×10^{-9}
13:114,744,546	rs75414584	<i>RASA3</i>	6.31×10^{-3}	-	4.62×10^{-12}
14:64,390,030	rs10873172 ^{a,d}	<i>SYNE2</i>	6.29×10^{-8}	5.95×10^{-13}	6.47×10^{-27}
14:69,226,931	rs11625064 ^d	<i>ZFP36L1</i>	3.33×10^{-6}	2.09×10^{-10}	1.83×10^{-19}
14:92,795,912	rs4904871	<i>SLC24A4</i>	2.06×10^{-4}	-	2.15×10^{-278}
14:103,923,475	rs2273699	<i>MARK3</i>	5.27×10^{-5}	-	1.21×10^{-16}
15:48,400,199	rs2675345	<i>SLC24A5</i>	4.92×10^{-3}	-	1.09×10^{-9}
16:54,118,132	rs62034121	<i>FTO</i>	1.16×10^{-9}	-	2.56×10^{-8}
16:54,131,939	rs62034139 ^{a,e}	<i>FTO</i>	4.56×10^{-9}	4.69×10^{-14}	-
16:55,322,732	rs12930459 ^a	<i>IRX6</i>	1.82×10^{-5}	4.89×10^{-9}	-

Reported cutaneous melanoma P values are from the total fixed-effects inverse-variance weighted meta-analysis of logistic regression two-sided P values from GWAS representing a total of 36,760 cases of melanoma and 375,188 controls (Methods). Results for the lead variants from pleiotropic loci (lead SNP reaching $P < 5 \times 10^{-8}$ following a Stouffer's sample size weighted meta-analysis of cutaneous melanoma P values and either nevus GWAS meta-analysis ($n = 65,777$) or hair color GWAS ($n = 352,662$) and GWAS-PW model 3 prior probability of association > 0.5 , Methods) distinct from those in the total cutaneous melanoma meta-analysis (Tables 1 and 2). CHR:BP, hg19 positional information. rsID, dbSNP142 rs number. Genes, prioritizing genes that the variant is an eQTL for in GTEx skin datasets or otherwise is the closest protein coding gene. We report the total cutaneous melanoma meta-analysis P (CM P) and the CM + nevus or CM + hair color Stouffer's meta-analysis fixed-effect P value. Full results can be found in Supplementary Tables 7 and 10. ^aLocus previously reported as pleiotropically associated with cutaneous melanoma and nevus count but not significant for cutaneous melanoma alone here. ^bLead SNP for pigment (rs10434895) and nevus (rs10434896) are in LD $r^2_{\text{EUR}} = 1.0$. ^cLead SNP for pigment (rs520015) and nevus (rs593179) are in LD $r^2_{\text{EUR}} = 0.63$. ^dSame lead SNP. ^eLead SNP for pigment (rs62034121) and nevus (rs62034139) are in LD $r^2_{\text{EUR}} = 0.88$.

To empirically identify the target tissues for cutaneous melanoma risk variants, we used partitioned LD score regression¹² to determine the proportion of total cutaneous melanoma GWAS meta-analysis heritability that could be captured by genes expressed in melanocytes and in 50 GTEx tissue types. We found that partitioned cutaneous melanoma heritability was most enriched in genes specifically expressed in melanocytes (2.76-fold, $P = 3.12 \times 10^{-6}$ for top 4,000 genes; Extended Data Fig. 5), followed by three other skin-related tissues (GTEx sun-exposed and not sun-exposed skin and transformed skin fibroblasts). This enrichment was much stronger than the one based on the previously published melanoma GWAS²³. We then focused on these four tissues for discovery of new loci, applying Bonferroni correction for multiple comparisons based on the

number of genes tested within each tissue set (Methods). TWAS using the melanocyte dataset (Supplementary Table 11) identified a total of 40 significant genes. Combining genes within 1 megabase (Mb) of each other into discrete loci, 32 genes were located within 13 formally genome-wide-significant cutaneous melanoma GWAS loci, and eight genes were identified within six new loci. Considering the other skin-related tissues collectively (Supplementary Table 12), TWAS identified a single significant gene at one additional new locus, as well as genes within 15 GWAS-significant loci. The TWAS using all GTEx tissues is reported in Supplementary Table 13.

In aggregate, these complementary approaches identified a total of 85 discrete loci (Fig. 2 and Supplementary Table 14): 54 formally significant at $P < 5 \times 10^{-8}$ in the total cutaneous melanoma

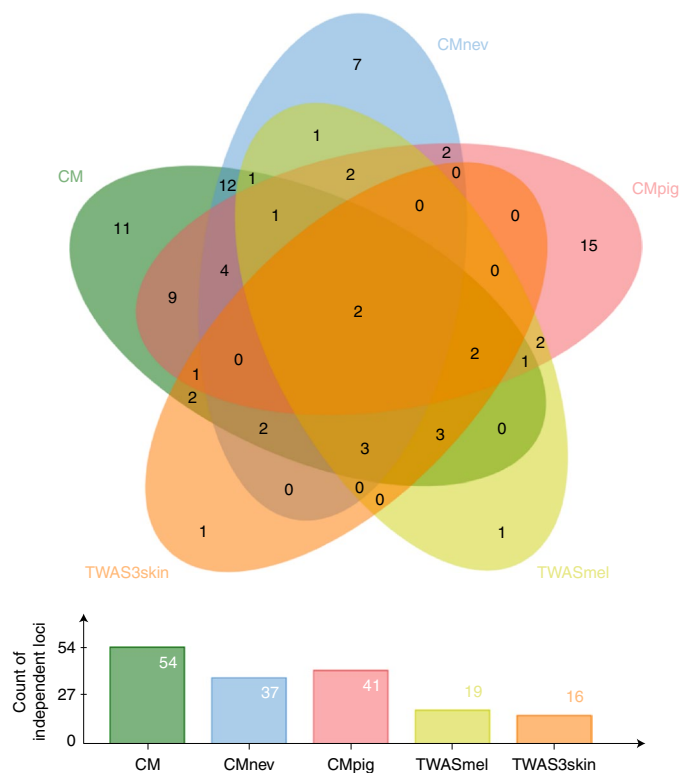


Fig. 2 | Overlap of loci identified by primary and secondary analyses. Loci identified in the total cutaneous melanoma meta-analysis (CM, green; Supplementary Table 3), the pleiotropic analysis with nevus count (CMnev, blue; Supplementary Table 9) and hair color (CMpig, red; Supplementary Table 10), melanocyte TWAS (TWASmel, yellow; Supplementary Table 10) and TWAS using the expression of three skin tissues (TWAS3skin, orange; Supplementary Table 12).

meta-analysis (Tables 1 and 2 and Supplementary Table 3), and the remainder supported by one or more of the secondary analyses (Tables 3 and 4 and Supplementary Tables 9–12) and likely representing additional cutaneous melanoma risk loci but requiring a larger sample size to reach genome-wide significance. To annotate cutaneous melanoma GWAS loci for candidate susceptibility genes for pathway analyses as well as future functional studies, we turned to eQTL colocalization analyses. These approaches identified multiple pathways that may play a role in developing melanoma and are described in the Supplementary Note.

Discussion

This meta-analysis of cutaneous melanoma susceptibility has more than three times the effective sample size in prior analyses, providing power to identify cutaneous melanoma susceptibility variants and enhanced distinction of independent variants in previously reported cutaneous melanoma susceptibility regions^{17–24} (Supplementary Table 1). We identified 68 independent cutaneous melanoma-associated variants across 54 loci. TWAS analysis, eQTL colocalization and multimarker genomic annotations identified promising gene candidates at many of these risk loci. Joint pairwise GWAS with the cutaneous melanoma-related traits of nevus count and hair color, and TWAS, identified a further 31 independent loci that, while not formally reaching genome-wide significance for cutaneous melanoma alone, represent potential additional risk loci. Our cutaneous melanoma meta-analysis also confirmed several loci previously identified only by TWAS²⁸, supporting the value of TWAS in identifying additional genes associated with cutaneous

melanoma (Table 4). In total, our integrative analysis identified 85 loci associated with cutaneous melanoma susceptibility (Tables 1–4 and Fig. 2), constituting a substantial increase from the 21 loci previously identified by cutaneous melanoma susceptibility GWAS alone (Table 1), in addition to those found by family-based approaches or in combination with nevus GWAS data (Table 2).

Our analyses showed strong genetic correlation between self-reported and clinically confirmed cases (Supplementary Table 2, Extended Data Fig. 6 and Supplementary Note), and inclusion of self-reported cases enabled the identification of 11 additional cutaneous melanoma susceptibility loci (Supplementary Tables 3 and 6 and Supplementary Note), indicating that self-reported cases of cutaneous melanoma are a valuable and reliable resource for genomic cutaneous melanoma studies. Furthermore, we assessed cutaneous melanoma genetic susceptibility across several geographic regions, including the often under-represented Mediterranean population. Interestingly, we found little evidence for difference in cutaneous melanoma locus effect estimates by contributing GWAS (Supplementary Fig. 2) or differences in effect size and allele frequency by geographic regions (Supplementary Fig. 3), beyond minor variation in pigmentation genes (for example, rs6059655 near *ASIP* and rs1805007 in *MC1R*). The stratified analysis based on cutaneous melanoma histological subtypes identified acral lentiginous melanomas as being uniquely unassociated with pigmentation loci, in line with observational data²⁹. In contrast, the stratified analyses based on age at diagnosis and gender found no evidence for differences in the distribution of nevus-related or pigmentation-related loci.

The discovery of new loci and genes augments our understanding of cutaneous melanoma risk and provides many new insights into cutaneous melanoma etiology. Many of the loci previously associated with nevus count²⁷ or pigmentation⁴³ are also associated with cutaneous melanoma (Table 2) confirming the close relationship between these traits. Specifically, of ten loci previously significantly associated in a joint analysis of cutaneous melanoma and nevus, but not associated with cutaneous melanoma alone²⁷, six are now associated with cutaneous melanoma alone (Table 2), demonstrating the benefits of conducting joint analyses. The remaining four loci reach $P < 5 \times 10^{-8}$ in the joint cutaneous melanoma + nevus analysis (Supplementary Table 9); three of which are significant at the Bonferroni-corrected threshold of 1.25×10^{-8} (Table 3). In turn, we conducted further pleiotropic analyses and identified secondary loci associated with a combination of both these traits and cutaneous melanoma but not significantly associated with cutaneous melanoma alone (Table 3). Loci found in such joint analyses are of value as they would likely be associated with cutaneous melanoma alone in a sufficiently large GWAS meta-analysis. These joint analyses provide a direct biological interpretation that several GWAS risk loci may act through nevus development, in line with clinical evidence. Interestingly, following these expanded pleiotropic analyses, many loci were associated with neither nevus count nor hair color, indicating that many risk variants act outside these classic cutaneous melanoma risk phenotypes (Tables 1 and 2).

The discovery of many new loci, when added to the existing catalog of melanoma risk loci, augments our understanding of the genetic architecture of cutaneous melanoma, as discussed in the Supplementary Note. It is important to note that confirmation that the genes we have identified are causal for cutaneous melanoma, and the biological understanding of how variants at these loci influence cutaneous melanoma remains to be functionally established. For example, melanocyte eQTL and TWAS analyses indicated *PARP1* expression was associated with cutaneous melanoma risk SNPs at 1q42 (refs. 28,44). While *PARP1* is an established DNA repair gene, extensive functional characterization of the cutaneous melanoma risk locus over *PARP1* demonstrated that its role in cutaneous melanoma appears to be through regulation of melanocyte proliferation,

Table 4 | Genes identified by TWAS outside of regions identified in the total cutaneous melanoma GWAS meta-analysis

Genes	TWAS		Locus peak CM variant		
	Z	P	rsID	CHR:BP	CM P
<i>NIPAL3</i>	4.84	1.28×10^{-6}	rs2294524	1:24,770,594	2.74×10^{-7}
<i>RCAN3</i>	4.83	1.33×10^{-6}	rs2294524	1:24,770,594	2.74×10^{-7}
<i>NOTCH2</i>	4.81	1.50×10^{-6}	rs2793830	1:120,466,108	3.80×10^{-7}
<i>PTPN14</i>	-4.84	1.30×10^{-6}	rs6693492	1:214,685,978	2.68×10^{-5}
<i>CBWD1</i>	-4.81	1.51×10^{-6}	rs478882	9:205,964	1.64×10^{-6}
<i>C9orf66</i>	5.05	4.48×10^{-7}	rs478882	9:205,964	1.64×10^{-6}
<i>SYNE2</i>	5.19	2.06×10^{-7}	rs12881652	14:64,400,120	2.12×10^{-7}
<i>IRX6</i>	-4.80	1.62×10^{-6}	rs12919110	16:55,319,789	1.27×10^{-6}
<i>RP11-676J12.7^a</i>	-5.55	2.79×10^{-8}	rs1703824	17:813,324	1.59×10^{-5}

For each gene with a Bonferroni-corrected *P* value cutoff in melanocytes ($P_{TWAS} < 3.22 \times 10^{-6}$) or skin-related tissue types ($P_{TWAS} < 5.28 \times 10^{-7}$) that does not overlap with an existing cutaneous melanoma region we report the local peak cutaneous melanoma variant from the total confirmed plus self-reported GWAS meta-analysis and TWAS Z score. Full results for all genes with a $P_{TWAS} < 1.48 \times 10^{-5}$ can be found in Supplementary Tables 10 and 12. *CBWD1* and *C9orf66* are within 1 Mb of each other and are merged into a single locus. ^a*RP11-676J12.7* was identified using sun-exposed skin expression data from GTEx (Supplementary Table 12), while all other genes were identified using melanocyte gene expression.

senescence and transcriptional regulation of the key melanoma oncogene *MITF*⁴⁴. Despite the need for follow-up functional studies, a preliminary, complex model of pathways potentially important for the development of melanoma is emerging through the candidate genes suggested by this and prior work, including pathways mediating protection against ultraviolet-induced DNA damage and DNA repair, telomere maintenance, immunity, melanocyte differentiation and cell adhesion.

For example, we identified an association between multiple independent variants at the *TP53* locus, rs78378222 and rs1641548, and cutaneous melanoma, further reinforcing the potential importance of DNA repair and genome integrity for cutaneous melanoma susceptibility (Supplementary Note). Rare germline mutations in *TP53* lead to Li–Fraumeni syndrome⁴⁵ which is associated with early onset of cancer, including cutaneous melanoma⁴⁶. Notably, one of the common sequence variants we found to be associated with cutaneous melanoma has previously been shown to alter *TP53* messenger RNA levels by disruption of *TP53* polyadenylation. *TP53* responds to cellular stresses to regulate target gene expression resulting in DNA repair, cell cycle arrest, apoptosis and cellular senescence^{47,48}; variation resulting in loss of normal *TP53* function could result in clonal expansion of cells that carry accumulated mutations, which may explain the association with both cutaneous melanoma and nevus count.

This study also adds to a growing body of evidence supporting a key role for telomere maintenance in cutaneous melanoma susceptibility^{8,9,13,14,37,49–51}, with cutaneous melanoma risk loci associated with telomere length or located near prominent telomere maintenance genes or loci, including *POT1*, *TERC*, *RTEL1*, *MPHOSPH6* and *OBFC1*. Additional previously identified GWAS loci are located near *CCND1* (rs4354713), *ATM* (rs1801516) and *PARP1* (rs2695237), all genes with established roles in telomere maintenance, DNA repair and regulation of senescence^{44,52}.

The well-established role of immunity in melanoma biology has fueled a search for an association between variation within the HLA region and melanoma risk^{53–55}. While several studies have investigated associations between HLA alleles and cutaneous melanoma, these studies have largely been conducted on small, underpowered datasets and have not been consistently replicated^{56–66}. Here, we report identification of a genome-wide-significant association between cutaneous melanoma susceptibility and rs28986343 at the HLA locus (Supplementary Note). This additional evidence for a role for immunity adds to previous²⁸ and current TWAS and colocalization analyses suggesting association between rs408825 and

expression of the innate immunity gene *MX2*. Additionally, many risk alleles for the autoimmune melanocyte-related disorder vitiligo^{67,68} are protective for cutaneous melanoma with the lead SNPs either identical (rs1126809/*TYR*; rs6059655/*ASIP*), or in strong LD with cutaneous melanoma lead SNPs (rs251464 near *PPARGC1B* for vitiligo; rs32578 for melanoma, LD $r^2_{EUR} = 0.73$; rs72928038 near *BACH2* for vitiligo; rs6908626 for melanoma, $r^2_{EUR} = 0.95$; rs1129038 near *OCA2* for vitiligo; rs12913832 for melanoma, $r^2_{EUR} = 0.99$). While the vitiligo and cutaneous melanoma associations share many similar loci, suggesting a role for immunity, we cannot rule out their action on cutaneous melanoma risk being through pigmentation or protection against ultraviolet damage. Taken as a whole, these data suggest further investigation into these potentially immune-related associations, and more broadly the role of immunity in melanoma risk.

New loci emerging from these analyses suggest a role of genes or networks regulating the development and differentiation of the melanocytic lineage. The cutaneous melanoma meta-analysis identified a locus near *FOXD3*, while the pleiotropic cutaneous melanoma + nevus analysis and TWAS locus identified a new locus significantly associated with allelic expression of *NOTCH2* in melanocytes (Supplementary Note). *FOXD3* participates as a part of a larger gene regulatory network governing the development of melanocytes from the neural crest, at least in part through transcriptional repression of one of the earliest markers of melanoblast development (and melanoma predisposition gene), *MITF*^{69,70}. *NOTCH2*, as well as *NOTCH1*, appear to play roles in both development of the melanocyte lineage as well as maintenance of melanocyte stem cells^{71,72} and *NOTCH* signaling has been shown to lead to de-differentiation of melanocytes to multipotent neural crest stem-like cells⁷³. These two new candidate susceptibility genes join previously identified loci also harboring genes involved in melanocyte fate. Whole-genome and targeted sequencing studies of melanoma-prone families led to the identification of a functional intermediate-penetrance missense mutation of *MITF* associated with both melanoma and nevus count (*MITF* pGlu318Lys)^{25,26}, a variant that was rediscovered by this population-based meta-analysis (rs149617956, $P = 5.17 \times 10^{-25}$, OR = 0.38). Additionally, a previously identified melanoma and nevus risk locus⁷⁴ is located ~200 kilobases (kb) from *SOX10*, another key regulator of melanocyte development and differentiation and direct transcriptional activator of *MITF*. These genes, and others in this gene regulatory network, have likewise been variously implicated in the progression of melanoma^{75–79}.

The identification of a cutaneous melanoma risk locus for which risk genotype strongly correlates with higher melanocyte-specific expression of *CDH1*, encoding E-cadherin, suggests a potential role for cell–cell adhesion in melanoma risk (Supplementary Note). E-cadherin plays a crucial role in cell–cell adhesion, epithelial–mesenchymal transition and carcinoma progression. Germline mutations in this gene are associated with a variety of tumors, including gastric⁸⁰, breast⁸¹ and potentially colorectal cancer⁸². In human skin, E-cadherin is typically expressed on the cell surface of both melanocytes and keratinocytes and is considered the major adhesion molecule between these two cell types^{83–85}. During melanoma progression, expression of E-cadherin is typically lost, with a concurrent switch to expression of N-cadherin, facilitating preferential association with fibroblasts and vascular endothelial cells⁸⁴. In contrast to loss of E-cadherin expression with melanoma progression, we find the cutaneous melanoma risk allele at this locus to be associated with higher expression of *CDH1*. Interestingly, melanocytes in nonlesional skin of vitiligo patients have been found to have loss of, or discontinuously distributed, E-cadherin expression. This loss of E-cadherin induces reduced adhesiveness to the basal layer under oxidative and mechanical stress, leading melanocytes to migrate passively to the exterior of the skin, and to die by apoptosis⁸⁶. Thus, germline variation leading to higher melanocyte *CDH1* could act as a protective mechanism, allowing cells damaged by oxidative stress to remain in the skin and survive without dying. A similar mechanism has been recently identified in breast cancer metastasis, where E-cadherin acts as a survival factor by limiting reactive oxygen-mediated apoptosis⁸⁷.

In summary, our large, international genetic meta-analysis showcases the utility of including self-reported cases of cutaneous melanoma, complementary analytical approaches, and data from multiple sources to expand our understanding of cutaneous melanoma risk. While the biological mechanisms underlying many of the existing and new cutaneous melanoma risk loci remain to be confirmed or discovered by post-GWAS functional studies and even larger GWAS, these data suggest potential pathways new to melanoma susceptibility, and highlight nevus formation, pigmentation and telomere maintenance, the three pathways that appear to dominate the landscape of melanoma susceptibility.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41588-020-0611-8>.

Received: 10 July 2019; Accepted: 9 March 2020;

Published online: 27 April 2020








References

- Karimkhani, C. et al. The global burden of melanoma: results from the global burden of disease study 2015. *Br. J. Dermatol.* **177**, 134–140 (2017).
- Secretan, B. et al. WHO International agency for research on cancer monograph working group a review of human carcinogens—Part E: tobacco, areca nut, alcohol, coal smoke, and salted fish. *Lancet Oncol.* **10**, 1033–1034 (2009).
- Ford, D. et al. Risk of cutaneous melanoma associated with a family history of the disease. *Int. J. Cancer* **62**, 377–381 (1995).
- Olsen, C. M., Carroll, H. J. & Whiteman, D. C. Familial melanoma: a meta-analysis and estimates of attributable fraction. *Cancer Epidemiol. Biomark. Prev.* **19**, 65–73 (2010).
- Olsen, C. M., Carroll, H. J. & Whiteman, D. C. Estimating the attributable fraction for melanoma: a meta-analysis of pigmentary characteristics and freckling. *Int. J. Cancer* **127**, 2430–2445 (2010).
- Chang, Y. M. et al. A pooled analysis of melanocytic nevus phenotype and the risk of cutaneous melanoma at different latitudes. *Int. J. Cancer* **124**, 420–428 (2009).
- Olsen, C. M., Carroll, H. J. & Whiteman, D. C. Estimating the attributable fraction for cancer: a meta-analysis of nevi and melanoma. *Cancer Prev. Res.* **3**, 233–245 (2010).
- Bataille, V. et al. Nevus size and number are associated with telomere length and represent potential markers of a decreased senescence in vivo. *Cancer Epidemiol. Biomark. Prev.* **16**, 1499–1502 (2007).
- Han, J. et al. A prospective study of telomere length and the risk of skin cancer. *J. Invest. Dermatol.* **129**, 415–421 (2009).
- Green, A. C. & Olsen, C. M. Increased risk of melanoma in organ transplant recipients: systematic review and meta-analysis of cohort studies. *Acta Derm. Venereol.* **95**, 923–927 (2015).
- Kamb, A. et al. Analysis of the p16 gene (*CDKN2*) as a candidate for the chromosome 9p melanoma susceptibility locus. *Nat. Genet.* **8**, 23–26 (1994).
- Berwick, M. et al. The prevalence of *CDKN2A* germ-line mutations and relative risk for cutaneous malignant melanoma: an international population-based study. *Cancer Epidemiol. Biomark. Prev.* **15**, 1520–1525 (2006).
- Robles-Espinoza, C. D. et al. *POT1* loss-of-function variants predispose to familial melanoma. *Nat. Genet.* **46**, 478–481 (2014).
- Shi, J. et al. Rare missense variants in *POT1* predispose to familial cutaneous malignant melanoma. *Nat. Genet.* **46**, 482–486 (2014).
- Palmer, J. S. et al. Melanocortin-1 receptor polymorphisms and risk of melanoma: is the association explained solely by pigmentation phenotype? *Am. J. Hum. Genet.* **66**, 176–186 (2000).
- Landi, M. T. et al. MC1R, ASIP, and DNA repair in sporadic and familial melanoma in a Mediterranean population. *J. Natl. Cancer Inst.* **98**, 144–145 (2005).
- Brown, K. M. et al. Common sequence variants on 20q11.22 confer melanoma susceptibility. *Nat. Genet.* **40**, 838–840 (2008).
- Bishop, D. T. et al. Genome-wide association study identifies three loci associated with melanoma risk. *Nat. Genet.* **41**, 920–925 (2009).
- Amos, C. I. et al. Genome-wide association study identifies novel loci predisposing to cutaneous melanoma. *Hum. Mol. Genet.* **20**, 5012–5023 (2011).
- Barrett, J. H. et al. Genome-wide association study identifies three new melanoma susceptibility loci. *Nat. Genet.* **43**, 1108–1113 (2011).
- Macgregor, S. et al. Genome-wide association study identifies a new melanoma susceptibility locus at 1q21.3. *Nat. Genet.* **43**, 1114–1118 (2011).
- Iles, M. M. et al. A variant in *FTO* shows association with melanoma risk not due to BMI. *Nat. Genet.* **45**, 428–432 (2013).
- Law, M. H. et al. Genome-wide meta-analysis identifies five new susceptibility loci for cutaneous malignant melanoma. *Nat. Genet.* **47**, 987–995 (2015).
- Ransohoff, K. J. et al. Two-stage genome-wide association study identifies a novel susceptibility locus associated with melanoma. *Oncotarget* **8**, 17586–17592 (2017).
- Yokoyama, S. et al. A novel recurrent mutation in *MITF* predisposes to familial and sporadic melanoma. *Nature* **480**, 99–103 (2011).
- Bertolotto, C. et al. A SUMOylation-defective *MITF* germline mutation predisposes to melanoma and renal carcinoma. *Nature* **480**, 94–98 (2011).
- Duffy, D. L. et al. Novel pleiotropic risk loci for melanoma and nevus density implicate multiple biological pathways. *Nat. Commun.* **9**, 4774 (2018).
- Zhang, T. et al. Cell-type-specific eQTL of primary melanocytes facilitates identification of melanoma susceptibility genes. *Genome Res.* **28**, 1621–1635 (2018).
- Elder, D. E., Massi, D., Willemze, R. & Scolyer, R. *WHO Classification of Skin Tumours* (International Agency for Research on Cancer, 2018).
- Higgins, J. P. T. & Thompson, S. G. Quantifying heterogeneity in a meta-analysis. *Stat. Med.* **21**, 1539–1558 (2002).
- Bulik-Sullivan, B. et al. An atlas of genetic correlations across human diseases and traits. *Nat. Genet.* **47**, 1236–1241 (2015).
- Zhang, Y., Qi, G., Park, J.-H. & Chatterjee, N. Estimation of complex effect-size distributions using summary-level statistics from genome-wide association studies across 32 complex traits. *Nat. Genet.* **50**, 1318–1326 (2018).
- Yang, J. et al. Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. *Nat. Genet.* **44**, 369–375 (2012).
- Duffy, D. L. et al. Multiple pigmentation gene polymorphisms account for a substantial proportion of risk of cutaneous malignant melanoma. *J. Invest. Dermatol.* **130**, 520–528 (2010).
- Duffy, D. L. et al. *IRF4* variants have age-specific effects on nevus count and predispose to melanoma. *Am. J. Hum. Genet.* **87**, 6–16 (2010).
- Delgado, D. A. et al. Genome-wide association study of telomere length among South Asians identifies a second *RTEL1* association signal. *J. Med. Genet.* **55**, 64–71 (2018).
- Iles, M. M. et al. The effect on melanoma risk of genes previously associated with telomere length. *J. Natl. Cancer Inst.* **106**, dju267 (2014).
- Pickrell, J. K. et al. Detection and interpretation of shared genetic influences on 42 human traits. *Nat. Genet.* **48**, 709–717 (2016).

39. GTEx Consortium Human genomics: the genotype-tissue expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science* **348**, 648–660 (2015).
40. Gamazon, E. R. et al. A gene-based association method for mapping traits using reference transcriptome data. *Nat. Genet.* **47**, 1091–1098 (2015).
41. Gusev, A. et al. Integrative approaches for large-scale transcriptome-wide association studies. *Nat. Genet.* **48**, 245–252 (2016).
42. Finucane, H. K. et al. Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nat. Genet.* **47**, 1228–1235 (2015).
43. Visconti, A. et al. Genome-wide association study in 176,678 Europeans reveals genetic loci for tanning response to sun exposure. *Nat. Commun.* **9**, 1684 (2018).
44. Choi, J. et al. A common intronic variant of *PARP1* confers melanoma risk and mediates melanocyte growth via regulation of *MITF*. *Nat. Genet.* **49**, 1326–1335 (2017).
45. Li, F. P. & Fraumeni, J. F. Jr. Soft-tissue sarcomas, breast cancer, and other neoplasms. A familial syndrome? *Ann. Intern. Med.* **71**, 747–752 (1969).
46. Curriel-Lewandrowski, C., Speetzen, L. S., Cranmer, L., Warneke, J. A. & Loescher, L. J. Multiple primary cutaneous melanomas in Li-Fraumeni syndrome. *Arch. Dermatol.* **147**, 248–250 (2011).
47. Beausejour, C. M. Reversal of human cellular senescence: roles of the p53 and p16 pathways. *EMBO J.* **22**, 4212–4222 (2003).
48. Kuilman, T., Michaloglou, C., Mooi, W. J. & Peeper, D. S. The essence of senescence. *Genes Dev.* **24**, 2463–2479 (2010).
49. Aoude, L. G. et al. Nonsense mutations in the shelterin complex genes *ACD* and *TERF2IP* in familial melanoma. *J. Natl. Cancer Inst.* **107**, dju408 (2015).
50. Rafnar, T. et al. Sequence variants at the *TERT-CLPTM1L* locus associate with many cancer types. *Nat. Genet.* **41**, 221–227 (2009).
51. Rachakonda, S. et al. Telomere length, telomerase reverse transcriptase promoter mutations, and melanoma risk. *Genes Chromosomes Cancer* **57**, 564–572 (2018).
52. Derheimer, F. A. & Kastan, M. B. Multiple roles of ATM in monitoring and maintaining DNA integrity. *FEBS Lett.* **584**, 3675–3681 (2010).
53. Demenais, F. et al. A linkage study between HLA and cutaneous malignant melanoma or precursor lesions or both. *J. Med. Genet.* **21**, 429–435 (1984).
54. Bale, S. J. et al. Hereditary malignant melanoma is not linked to the HLA complex on chromosome 6. *Int. J. Cancer* **36**, 439–443 (1985).
55. Holland, E. A., Beaton, S. C., Kefford, R. F. & Mann, G. J. Linkage analysis of familial melanoma and chromosome 6 in 14 Australian kindreds. *Genes Chromosomes Cancer* **19**, 241–249 (1997).
56. Barger, B. O., Acton, R. T., Soong, S. J., Roseman, J. & Balch, C. Increase of HLA-DR4 in melanoma patients from Alabama. *Cancer Res.* **42**, 4276–4279 (1982).
57. Rovini, D., Sacchini, V., Codazzi, V., Vaglini, M. & Illeni, M. T. HLA antigen frequencies in malignant melanoma patients. A second study. *Tumori* **70**, 29–33 (1984).
58. Hors, J. et al. in *Histocompatibility Testing 1984* (ed. Albert, E. D.) 407–410 (Springer, 1984).
59. Lee, J. E., Reveille, J. D., Ross, M. I. & Platsoucas, C. D. HLA-DQB1*0301 association with increased cutaneous melanoma risk. *Int. J. Cancer* **59**, 510–513 (1994).
60. Muto, M. et al. HLA class I polymorphism and the susceptibility to malignant melanoma. *Tissue Antigens* **47**, 447–449 (1996).
61. Kageshita, T. et al. Molecular genetic analysis of HLA class II alleles in Japanese patients with melanoma. *Tissue Antigens* **49**, 466–470 (1997).
62. Bateman, A. C., Turner, S. J., Theaker, J. M. & Howell, W. M. HLA-DQB1*0303 and *0301 alleles influence susceptibility to and prognosis in cutaneous malignant melanoma in the British Caucasian population. *Tissue Antigens* **52**, 67–73 (1998).
63. Lombardi, M. L. et al. Molecular analysis of HLA DRB1 and DQB1 polymorphism in Italian melanoma patients. *J. Immunother.* **21**, 435–439 (1998).
64. Luongo, V. et al. HLA allele frequency and clinical outcome in Italian patients with cutaneous melanoma. *Tissue Antigens* **64**, 84–87 (2004).
65. Campillo, J. A. et al. HLA class I and class II frequencies in patients with cutaneous malignant melanoma from southeastern Spain: the role of HLA-C in disease prognosis. *Immunogenetics* **57**, 926–933 (2006).
66. Planelles, D. et al. HLA class II polymorphisms in Spanish melanoma patients: homozygosity for HLA-DQA1 locus can be a potential melanoma risk factor. *Br. J. Dermatol.* **154**, 261–266 (2006).
67. Jin, Y. et al. Genome-wide association studies of autoimmune vitiligo identify 23 new risk loci and highlight key pathways and regulatory variants. *Nat. Genet.* **48**, 1418–1424 (2016).
68. Jin, Y. et al. Genome-wide association analyses identify 13 new susceptibility loci for generalized vitiligo. *Nat. Genet.* **44**, 676–680 (2012).
69. Curran, K. et al. Interplay between *Foxd3* and *Mitf* regulates cell fate plasticity in the zebrafish neural crest. *Dev. Biol.* **344**, 107–118 (2010).
70. Thomas, A. J. & Erickson, C. A. *FOXD3* regulates the lineage switch between neural crest-derived glial cells and pigment cells by repressing *MITF* through a non-canonical mechanism. *Development* **136**, 1849–1858 (2009).
71. Kumano, K. et al. Both *Notch1* and *Notch2* contribute to the regulation of melanocyte homeostasis. *Pigment Cell Melanoma Res.* **21**, 70–78 (2008).
72. Schouwey, K., Larue, L., Radtke, F., Delmas, V. & Beermark, F. Transgenic expression of notch in melanocytes demonstrates RBP-Jkappa-dependent signaling. *Pigment Cell Melanoma Res.* **23**, 134–136 (2010).
73. Zabierowski, S. E. et al. Direct reprogramming of melanocytes to neural crest stem-like cells by one defined factor. *Stem Cells* **29**, 1752–1762 (2011).
74. Falchi, M. et al. Genome-wide association study identifies variants at 9p21 and 22q13 associated with development of cutaneous nevi. *Nat. Genet.* **41**, 915–919 (2009).
75. Garraway, L. A. et al. Integrative genomic analyses identify *MITF* as a lineage survival oncogene amplified in malignant melanoma. *Nature* **436**, 117–122 (2005).
76. Abel, E. V. & Aplin, A. E. *FOXD3* is a mutant B-RAF-regulated inhibitor of G(1)-S progression in melanoma cells. *Cancer Res.* **70**, 2891–2900 (2010).
77. Weiss, M. B., Abel, E. V., Dadpey, N. & Aplin, A. E. *FOXD3* modulates migration through direct transcriptional repression of *TWIST1* in melanoma. *Mol. Cancer Res.* **12**, 1314–1323 (2014).
78. Golan, T. et al. Interactions of melanoma cells with distal keratinocytes trigger metastasis via notch signaling inhibition of *MITF*. *Mol. Cell* **59**, 664–676 (2015).
79. Cronin, J. C. et al. *SOX10* ablation arrests cell cycle, induces senescence, and suppresses melanomagenesis. *Cancer Res.* **73**, 5709–5718 (2013).
80. Guilford, P. et al. E-cadherin germline mutations in familial gastric cancer. *Nature* **392**, 402–405 (1998).
81. Hansford, S. et al. Hereditary diffuse gastric cancer syndrome: *CDH1* mutations and beyond. *JAMA Oncol.* **1**, 23–32 (2015).
82. Kim, H. C. et al. The E-cadherin gene (*CDH1*) variants T340A and L599V in gastric and colorectal cancer patients in Korea. *Gut* **47**, 262–267 (2000).
83. Tang, A. et al. E-cadherin is the major mediator of human melanocyte adhesion to keratinocytes in vitro. *J. Cell Sci.* **107**(Pt 4), 983–992 (1994).
84. Hsu, M. Y., Wheelock, M. J., Johnson, K. R. & Herlyn, M. Shifts in cadherin profiles between human normal melanocytes and melanomas. *J. Invest. Dermatol. Symp. Proc.* **1**, 188–194 (1996).
85. Study, C. et al. Meta-analysis of genome-wide association data identifies four new susceptibility loci for colorectal cancer. *Nat. Genet.* **40**, 1426–1435 (2008).
86. Wagner, R. Y. et al. Altered E-cadherin levels and distribution in melanocytes precede clinical manifestations of vitiligo. *J. Invest. Dermatol.* **135**, 1810–1819 (2015).
87. Padmanaban, V. et al. E-cadherin is required for metastasis in multiple models of breast cancer. *Nature* **573**, 439–444 (2019).
88. Peña-Chilet, M. et al. Genetic variants in *PARP1* (rs3219090) and *IRF4* (rs12203592) genes associated with melanoma susceptibility in a Spanish population. *BMC Cancer* **13**, 160 (2013).
89. Law, M. H. et al. Meta-analysis combining new and existing data sets confirms that the *TERT-CLPTM1L* locus influences melanoma risk. *J. Invest. Dermatol.* **132**, 485–487 (2012).
90. Antonopoulou, K. et al. Updated field synopsis and systematic meta-analyses of genetic association studies in cutaneous melanoma: the MelGene database. *J. Invest. Dermatol.* **135**, 1074–1079 (2015).
91. McCarthy, S. et al. A reference panel of 64,976 haplotypes for genotype imputation. *Nat. Genet.* **48**, 1279–1283 (2016).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This is a U.S. government work and not under copyright protection in the U.S.; foreign copyright protection may apply 2020

Maria Teresa Landi^{1,116} , **D. Timothy Bishop**^{2,116} , **Stuart MacGregor**^{3,116} , **Mitchell J. Machiela**^{1,116} , **Alexander J. Stratigos**^{4,117}, **Paola Ghiorzo**^{5,6,117}, **Myriam Brossard**⁷, **Donato Calista**⁸, **Jiyeon Choi**¹, **Maria Concetta Fargnoli**⁹, **Tongwu Zhang**¹ , **Monica Rodolfo**¹⁰ , **Adam J. Trower**¹¹, **Chiara Menin**¹², **Jacobo Martinez**¹³, **Andreas Hadjisavvas**¹⁴, **Lei Song**¹, **Irene Stefanaki**¹⁵, **Richard Scolyer**^{16,17,18,19} 

Rose Yang¹, Alisa M. Goldstein¹, Miriam Potrony²⁰, Katerina P. Kypreou¹⁵, Lorenza Pastorino^{5,6}, Paola Queirolo²¹, Cristina Pellegrini⁹, Laura Cattaneo²², Matthew Zawistowski^{23,24}, Pol Gimenez-Xavier²⁰, Arantxa Rodriguez²⁵, Lisa Elefanti¹², Siranoush Manoukian²⁶, Licia Rivoltini¹⁰, Blair H. Smith²⁷, Maria A. Loizidou¹⁴, Laura Del Regno^{28,29}, Daniela Massi³⁰, Mario Mandala³¹, Kiarash Khosrotehrani^{32,33}, Lars A. Akslen^{34,35}, Christopher I. Amos³⁶, Per A. Andresen³⁷, Marie-Françoise Avril³⁸, Esther Azizi^{39,40}, H. Peter Soyer^{33,41}, Veronique Bataille^{42,43}, Bruna Dalmasso^{5,6}, Lisa M. Bowdler⁴⁴, Kathryn P. Burdon⁴⁵, Wei V. Chen⁴⁶, Veryan Codd^{47,48}, Jamie E. Craig⁴⁹, Tadeusz Dębniak⁵⁰, Mario Falchi^{42,43}, Shenyang Fang⁵¹, Eitan Friedman⁴⁰, Sarah Simi³⁰, Pilar Galan⁵², Zaida Garcia-Casado²⁵, Elizabeth M. Gillanders⁵³, Scott Gordon⁵⁴, Adele Green^{55,56}, Nelleke A. Gruis⁵⁷, Johan Hansson⁵⁸, Mark Harland⁵⁹, Jessica Harris⁶⁰, Per Helsing⁶¹, Anjali Henders⁶², Marko Hočvar⁶³, Veronica Höiom⁵⁸, David Hunter^{64,65}, Christian Ingvar⁶⁶, Rajiv Kumar⁶⁷, Julie Lang⁶⁸, G. Mark Lathrop⁶⁹, Jeffrey E. Lee⁵¹, Xin Li⁷⁰, Jan Lubiński⁷¹, Rona M. Mackie^{68,72}, Maryrose Malt⁵⁵, Josep Malvehy²⁰, Kerrie McAloney⁵⁴, Hamida Mohamdi⁷, Anders Molven^{35,73}, Eric K. Moses⁷⁴, Rachel E. Neale⁷⁵, Srdjan Novaković⁷⁶, Dale R. Nyholt^{54,77}, Håkan Olsson^{78,79}, Nicholas Orr⁸⁰, Lars G. Fritsche⁸¹, Joan Anton Puig-Butille⁸², Abrar A. Qureshi⁸³, Graham L. Radford-Smith^{84,85,86}, Juliette Randerson-Moor⁵⁹, Celia Requena²⁵, Casey Rowe³², Nilesh J. Samani^{47,48}, Marianna Sanna^{42,43}, Dirk Schadendorf^{87,88}, Hans-Joachim Schulze⁸⁹, Lisa A. Simms⁸⁴, Mark Smithers^{90,91}, Fengju Song⁹², Anthony J. Swerdlow^{93,94}, Nienke van der Stoep⁹⁵, Nicole A. Kukutsch⁵⁷, Alessia Visconti^{42,43}, Leanne Wallace⁶², Sarah V. Ward^{96,97}, Lawrie Wheeler⁶⁰, Richard A. Sturm⁴¹, Amy Hutchinson^{1,98}, Kristine Jones^{1,98}, Michael Malasky^{1,98}, Aurelie Vogt^{1,98}, Weiyin Zhou^{1,98}, Karen A. Pooley⁹⁹, David E. Elder¹⁰⁰, Jiali Han⁷⁰, Belynda Hicks^{1,98}, Nicholas K. Hayward¹⁰¹, Peter A. Kanetsky¹⁰², Chad Brummett¹⁰³, Grant W. Montgomery⁶², Catherine M. Olsen¹⁰⁴, Caroline Hayward¹⁰⁵, Alison M. Dunning¹⁰⁶, Nicholas G. Martin⁵⁴, Evangelos Evangelou^{107,108}, Graham J. Mann^{16,109,110}, Georgina Long^{16,111}, Paul D. P. Pharoah¹⁰⁶, Douglas F. Easton⁹⁹, Jennifer H. Barrett¹¹, Anne E. Cust^{16,112}, Goncalo Abecasis¹¹³, David L. Duffy^{41,54}, David C. Whiteman¹⁰⁴, Helen Gogas¹¹⁴, Arcangela De Nicolo¹¹⁵, Margaret A. Tucker¹, Julia A. Newton-Bishop⁵⁹, GenoMEL Consortium*, Q-MEGA and QTWIN Investigators*, ATHENS Melanoma Study Group*, 23andMe*, The SDH Study Group*, IBD Investigators*, Essen-Heidelberg Investigators*, AMFS Investigators*, MelaNostrum Consortium*, Ketty Peris^{28,29}, Stephen J. Chanock¹, Florence Demenais⁷, Kevin M. Brown^{1,117}, Susana Puig^{20,117}, Eduardo Nagore^{25,117}, Jianxin Shi^{1,116}, Mark M. Iles^{11,116} and Matthew H. Law^{3,116} ✉

¹Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA. ²Leeds Institute of Medical Research at St James's, Leeds Institute for Data Analytics, University of Leeds, Leeds, UK. ³Statistical Genetics, QIMR Berghofer Medical Research Institute, Brisbane, Queensland, Australia. ⁴Department of Dermatology, Andreas Syggros Hospital, Medical School, National and Kapodistrian University of Athens, Athens, Greece. ⁵Genetics of Rare Cancers, Ospedale Policlinico San Martino, Genoa, Italy. ⁶Department of Internal Medicine and Medical Specialties (DIMI), University of Genoa, Genoa, Italy. ⁷Genetic Epidemiology and Functional Genomics of Multifactorial Diseases Team, Institut National de la Santé et de la Recherche Médicale (INSERM), UMR5-1124, Université Paris Descartes, Paris, France. ⁸Department of Dermatology, Maurizio Bufalini Hospital, Cesena, Italy. ⁹Department of Dermatology & Department of Biotechnological and Applied Clinical Sciences, University of L'Aquila, L'Aquila, Italy. ¹⁰Unit of Immunotherapy of Human Tumors, Department of Research, Fondazione IRCCS Istituto Nazionale dei Tumori di Milano, Milan, Italy. ¹¹Leeds Institute for Data Analytics, University of Leeds, Leeds, UK. ¹²Immunology and Molecular Oncology Unit, Veneto Institute of Oncology IOV-IRCCS, Padua, Italy. ¹³Red Valenciana de Biobancos, FISABIO, Valencia, Spain. ¹⁴Department of EM/Molecular Pathology & The Cyprus School of Molecular Medicine, The Cyprus Institute of Neurology and Genetics, Nicosia, Cyprus. ¹⁵Department of Dermatology, University of Athens School of Medicine, Andreas Syggros Hospital, Athens, Greece. ¹⁶Melanoma Institute Australia, The University of Sydney, Sydney, New South Wales, Australia. ¹⁷Royal Prince Alfred Hospital, Sydney, New South Wales, Australia. ¹⁸Central Clinical School, The University of Sydney, Sydney, New South Wales, Australia. ¹⁹New South Wales Health Pathology, Sydney, New South Wales, Australia. ²⁰Dermatology Department, Melanoma Unit, Hospital Clínic de Barcelona, IDIBAPS, Universitat de Barcelona, CIBERER, Barcelona, Spain. ²¹Medical Oncology Unit, IRCCS Ospedale Policlinico San Martino, Genoa, Italy. ²²Pathology Unit, Azienda Socio-Sanitaria Territoriale Papa Giovanni XXIII, Bergamo, Italy. ²³Department of Biostatistics, University of Michigan School of Public Health, Ann Arbor, MI, USA. ²⁴Center for Statistical Genetics, University of Michigan School of Public Health, Ann Arbor, MI, USA. ²⁵Department of Dermatology, Instituto

Valenciano de Oncología, Valencia, Spain. ²⁶Unit of Medical Genetics, Department of Medical Oncology and Hematology, Fondazione IRCCS Istituto Nazionale dei Tumori di Milano, Milan, Italy. ²⁷Division of Population Health and Genomics, Ninewells Hospital and Medical School, University of Dundee, Dundee, UK. ²⁸Institute of Dermatology, Catholic University, Rome, Italy. ²⁹Fondazione Policlinico Universitario A. Gemelli, IRCCS, Rome, Italy. ³⁰Section of Anatomic Pathology, Department of Health Sciences, University of Florence, Florence, Italy. ³¹Department of Oncology, Giovanni XXIII Hospital, Bergamo, Italy. ³²UQ Diamantina Institute, The University of Queensland, Brisbane, Queensland, Australia. ³³Department of Dermatology, Princess Alexandra Hospital, Brisbane, Queensland, Australia. ³⁴Centre for Cancer Biomarkers CCBIO, Department of Clinical Medicine, University of Bergen, Bergen, Norway. ³⁵Department of Pathology, Haukeland University Hospital, Bergen, Norway. ³⁶Department of Community and Family Medicine, Geisel School of Medicine, Dartmouth College, Hanover, NH, USA. ³⁷Department of Pathology, Molecular Pathology, Oslo University Hospital, Rikshospitalet, Oslo, Norway. ³⁸Assistance Publique-Hôpitaux de Paris, Hôpital Cochin, Service de Dermatologie, Université Paris Descartes, Paris, France. ³⁹Department of Dermatology, Sheba Medical Center, Tel Hashomer, Sackler Faculty of Medicine, Tel Aviv, Israel. ⁴⁰Oncogenetics Unit, Sheba Medical Center, Tel Hashomer, Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel. ⁴¹Dermatology Research Centre, The University of Queensland Diamantina Institute, Translational Research Institute, Brisbane, Queensland, Australia. ⁴²Department of Twin Research and Genetic Epidemiology, King's College London, London, UK. ⁴³Department of Dermatology, West Herts NHS Trust, Herts, UK. ⁴⁴Sample Processing, QIMR Berghofer Medical Research Institute, Brisbane, Queensland, Australia. ⁴⁵Menzies Institute for Medical Research, University of Tasmania, Hobart, Tasmania, Australia. ⁴⁶Department of Genetics, The University of Texas MD Anderson Cancer Center, Houston, TX, USA. ⁴⁷Department of Cardiovascular Sciences, University of Leicester, Leicester, UK. ⁴⁸NIHR Leicester Biomedical Research Centre, Glenfield Hospital, Leicester, UK. ⁴⁹Department of Ophthalmology, Flinders University, Adelaide, South Australia, Australia. ⁵⁰Department of Genetics and Pathology, International Hereditary Cancer Center, Pomeranian Medical University, Szczecin, Poland. ⁵¹Department of Surgical Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA. ⁵²Université Paris 13, Equipe de Recherche en Epidémiologie Nutritionnelle (EREN), Centre de Recherche en Epidémiologie et Statistiques, Institut National de la Santé et de la Recherche Médicale (INSERM U1153), Institut National de la Recherche Agronomique (INRA U1125), Conservatoire National des Arts et Métiers, Communauté d'Université Sorbonne Paris Cité, Bobigny, France. ⁵³Inherited Disease Research Branch, National Human Genome Research Institute, National Institutes of Health, Baltimore, MD, USA. ⁵⁴Genetic Epidemiology, QIMR Berghofer Medical Research Institute, Brisbane, Queensland, Australia. ⁵⁵Cancer and Population Studies, QIMR Berghofer Medical Research Institute, Brisbane, Queensland, Australia. ⁵⁶CRUK Manchester Institute, Institute of Inflammation and Repair, University of Manchester, Manchester, UK. ⁵⁷Department of Dermatology, Leiden University Medical Center, Leiden, the Netherlands. ⁵⁸Department of Oncology-Pathology, Karolinska Institutet, Karolinska University Hospital, Stockholm, Sweden. ⁵⁹Leeds Institute of Medical Research at St James's, University of Leeds, Leeds, UK. ⁶⁰Translational Research Institute, Institute of Health and Biomedical Innovation, Princess Alexandra Hospital, Queensland University of Technology, Brisbane, Queensland, Australia. ⁶¹Department of Dermatology, Oslo University Hospital, Rikshospitalet, Oslo, Norway. ⁶²Institute for Molecular Bioscience, The University of Queensland, Brisbane, Queensland, Australia. ⁶³Department of Surgical Oncology, Institute of Oncology Ljubljana, Ljubljana, Slovenia. ⁶⁴Nuffield Department of Population Health, University of Oxford, Oxford, UK. ⁶⁵Program in Genetic Epidemiology and Statistical Genetics, Harvard T.H. Chan School of Public Health, Boston, MA, USA. ⁶⁶Department of Surgery, Clinical Sciences, Lund University, Lund, Sweden. ⁶⁷Division of Molecular Genetic Epidemiology, German Cancer Research Center, Heidelberg, Germany. ⁶⁸Department of Medical Genetics, University of Glasgow, Glasgow, UK. ⁶⁹McGill University and Genome Quebec Innovation Centre, Montreal, Canada. ⁷⁰Department of Epidemiology, Richard M. Fairbanks School of Public Health, Melvin and Bren Simon Cancer Center, Indiana University, Indianapolis, IN, USA. ⁷¹International Hereditary Cancer Center, Pomeranian Medical University, Szczecin, Poland. ⁷²Department of Public Health, University of Glasgow, Glasgow, UK. ⁷³Gade Laboratory for Pathology, Department of Clinical Medicine, University of Bergen, Bergen, Norway. ⁷⁴Centre for Genetic Origins of Health and Disease, Faculty of Medicine, Dentistry and Health Sciences, The University of Western Australia, Crawley, Western Australia, Australia. ⁷⁵Cancer Aetiology & Prevention, QIMR Berghofer Medical Research Institute, Brisbane, Queensland, Australia. ⁷⁶Department of Molecular Diagnostics, Institute of Oncology Ljubljana, Ljubljana, Slovenia. ⁷⁷School of Biomedical Sciences and Institute of Health and Biomedical Innovation, Queensland University of Technology, Brisbane, Queensland, Australia. ⁷⁸Department of Oncology/Pathology, Clinical Sciences, Lund University, Lund, Sweden. ⁷⁹Department of Cancer Epidemiology, Clinical Sciences, Lund University, Lund, Sweden. ⁸⁰Breakthrough Breast Cancer Research Centre, The Institute of Cancer Research, London, UK. ⁸¹Center for Statistical Genetics, Department of Biostatistics, University of Michigan School of Public Health, Ann Arbor, MI, USA. ⁸²Biochemistry and Molecular Genetics Department, Melanoma Unit, Hospital Clínic de Barcelona, IDIBAPS, Universitat de Barcelona, CIBERER, Barcelona, Spain. ⁸³Department of Dermatology, The Warren Alpert Medical School of Brown University, Providence, RI, USA. ⁸⁴Inflammatory Bowel Diseases, QIMR Berghofer Medical Research Institute, Brisbane, Queensland, Australia. ⁸⁵Department of Gastroenterology and Hepatology, Royal Brisbane & Women's Hospital, Brisbane, Queensland, Australia. ⁸⁶University of Queensland School of Medicine, Herston Campus, Brisbane, Queensland, Australia. ⁸⁷Department of Dermatology, University Hospital Essen, Essen, Germany. ⁸⁸German Consortium Translational Cancer Research (DKTK), Heidelberg, Germany. ⁸⁹Department of Dermatology, Fachklinik Hornheide, Institute for Tumors of the Skin, University of Münster, Münster, Germany. ⁹⁰Queensland Melanoma Project, Princess Alexandra Hospital, The University of Queensland, St Lucia, Queensland, Australia. ⁹¹Mater Research Institute, The University of Queensland, St Lucia, Queensland, Australia. ⁹²Departments of Epidemiology and Biostatistics, Key Laboratory of Cancer Prevention and Therapy, Tianjin, National Clinical Research Center of Cancer, Tianjin Medical University Cancer Institute and Hospital, Tianjin, P. R. China. ⁹³Division of Genetics and Epidemiology, The Institute of Cancer Research, London, UK. ⁹⁴Division of Breast Cancer Research, The Institute of Cancer Research, London, UK. ⁹⁵Department of Clinical Genetics, Center of Human and Clinical Genetics, Leiden University Medical Center, Leiden, the Netherlands. ⁹⁶Centre for Genetic Origins of Health and Disease, School of Biomedical Sciences, The University of Western Australia, Perth, Western Australia, Australia. ⁹⁷Department of Epidemiology and Biostatistics, Memorial Sloan Kettering Cancer Center, New York, NY, USA. ⁹⁸Cancer Genome Research Laboratory, Leidos Biomedical Research, Bethesda, MD, USA. ⁹⁹Department of Public Health and Primary Care, Centre for Cancer Genetic Epidemiology, University of Cambridge, Cambridge, UK. ¹⁰⁰Department of Pathology and Laboratory Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA. ¹⁰¹Oncogenomics, QIMR Berghofer Medical Research Institute, Brisbane, Queensland, Australia. ¹⁰²Department of Cancer Epidemiology, H. Lee Moffitt Cancer Center and Research Institute, Tampa, FL, USA. ¹⁰³Department of Anesthesiology, University of Michigan, Ann Arbor, MI, USA. ¹⁰⁴Cancer Control Group, QIMR Berghofer Medical Research Institute, Brisbane, Queensland, Australia. ¹⁰⁵MRC Human Genetics Unit, Institute of Genetics and Molecular Medicine, University of Edinburgh, Western General Hospital, Edinburgh, UK. ¹⁰⁶Department of Oncology, Centre for Cancer Genetic Epidemiology, University of Cambridge, Cambridge, UK. ¹⁰⁷Department of Hygiene and Epidemiology, University of Ioannina Medical School, Ioannina, Greece. ¹⁰⁸Department of Epidemiology and Biostatistics, Imperial College London, London, UK. ¹⁰⁹Centre for Cancer Research, Westmead Institute for Medical Research, Sydney, Australia. ¹¹⁰John Curtin School of Medical Research, Australian National University, Canberra, Australian Capital Territory, Australia. ¹¹¹Royal North Shore Hospital, Sydney, Australia. ¹¹²Cancer Epidemiology and Prevention Research, Sydney School of Public Health, Sydney, Australia. ¹¹³Department of Biostatistics, University of Michigan, Ann Arbor, MI, USA. ¹¹⁴First Department of Internal Medicine, Laikon General Hospital Greece, National and Kapodistrian University of Athens, Athens, Greece. ¹¹⁵Cancer Genomics Program, Veneto Institute of Oncology IOV-IRCCS, Padua, Italy. ¹¹⁶These authors jointly supervised this work: Maria Teresa Landi, D. Timothy Bishop, Stuart MacGregor, Mitchell J. Machiela, Jianxin Shi, Mark M. Iles, Matthew H. Law. ¹¹⁷These authors contributed equally: Alexander J. Stratigos, Paola Ghiorzo, Kevin M. Brown, Susana Puig, Eduardo Nagore. *A list of members and affiliations appears in the Supplementary Note. ¹¹⁸e-mail: landim@mail.nih.gov; M.M.Iles@leeds.ac.uk; Matthew.Law@qimrberghofer.edu.au

Methods

Quality control metrics, imputation and association analysis. Data cleaning was performed using Illumina GenomeStudio/BeadStudio (v.2.0.4) and PLINK (v.1.90b5.4)^{92,93}. Full details of the sample collections and genotyping arrays used for each GWAS are reported in the Supplementary Note. Before imputation, any SNP with minor allele frequency (MAF) < 0.01, Hardy–Weinberg equilibrium (HWE) $P < 5 \times 10^{-4}$ in controls or $P < 5 \times 10^{-10}$ in cases was removed. Similarly, any individual was removed who was missing > 3% of variants, had heterozygosity values either > 0.05 or < -0.05 or 3 s.d. from the mean, whose genetically predicted sex did not match their recorded sex, or who was determined to be non-European based on principal component analysis (PCA). In addition, one of any pair of individuals estimated to be related with identity by descent (IBD) $\text{pihat} > 0.15$ was removed.

The Harvard, Brisbane Nevus Morphology Study and 23andMe GWAS were imputed to 1000 Genomes Project phase 1 v.3; for all other sets (Supplementary Table 1) imputation was conducted using the Michigan Imputation Server with the Haplotype Reference Consortium panel (HRC v.1) and run using Minimac3 (ref. ⁹⁴). Following imputation, any imputed variant with imputation quality score $r^2 < 0.5$ or MAF < 0.0001 was rejected. As rare SNPs where one allele is missing in the case or control group can lead to very large (or infinite) OR estimates, variants with an OR < 1×10^{-4} (the minimum reported by PLINK) or OR > 1×10^6 were also filtered. To handle variants with the same name (for example, triplicate SNPs), variant IDs were converted to the format CHR:BP:A1A2 before meta-analysis.

Logistic regression under an additive model with ORs calculated on a per-allele basis was then conducted using PLINK (v.1.90b5.4)^{92,93} with either geographic region (in GenoMEL Phase 1 and 2 data) or principal components as covariates to account for potential population stratification. Individual studies were checked for evidence of inflation by producing Q–Q plots (Supplementary Fig. 1) and calculating the corresponding inflation factor λ and LDSC intercept (Supplementary Table 1).

Where individual studies have deviated from this protocol, details are included in the study description in the Supplementary Note. All reported tests are two-sided.

Meta-analysis and conditional- and joint-analysis to identify independent loci. Meta-analyses of the GWAS were conducted in one stage using both inverse-variance weighted fixed-effects and random-effects meta-analysis⁹⁵ as implemented in PLINK v.1.90b5.4 (ref. ^{92,93}). Meta-analyses were conducted for confirmed-only cases and in the total set including self-report sets (23andMe and a portion of UK Biobank).

Conditional- and joint-analyses of summary GWAS meta-analysis data were performed using genome-wide complex trait analysis (GCTA, v.1.26.0) to identify independently associated variants³³. To ensure we were only detecting completely independent SNPs, the collinearity threshold (--cojo-collinear) was set to $R^2 = 0.05$. The threshold for genome-wide significance 5×10^{-8} and fixed-effect meta-analysis P values and log(OR) effect sizes were analysed.

Linkage disequilibrium (LD) between SNPs for the conditional- and joint-analyses of summary data in GCTA (v.1.26.0) reported in the manuscript was calculated using a reference population of 5,000 individuals selected randomly from the portion of the UK Biobank population determined to be European by PCA (LD_{EUR}). Variants were converted to best-guess genotype (threshold 0.3). Best-guess data were cleaned for missingness > 3%, HWE $P < 1 \times 10^{-6}$ and MAF < 0.001.

To limit the chance of false-positive claims of new SNP/loci, we further filtered the list of 77 conditionally independent variants (Supplementary Table 4) to those (1) genome-wide-significant ($P < 5 \times 10^{-5}$) in single SNP and conditional- and joint-analysis, and (2) as recommended³⁰ where there was evidence of heterogeneity between studies ($I^2 > 31\%$) the random-effect P value also needed to be $< 5 \times 10^{-8}$. Passing variants were further checked to ensure that MAFs and effect sizes were consistent across studies and that the result was not driven by a single study (Supplementary Figs. 2 and 3). The 68 retained variants were combined into 54 loci using a concatenating 1 Mb window (Supplementary Table 3). Regional association plots for all 54 loci were interactively plotted by LDassoc (<https://ldlink.nci.nih.gov/>)⁹⁶ and are included in Supplementary Data 1.

Multiple-testing corrections. The primary aim of our study was to perform a GWAS meta-analysis of cutaneous melanoma risk. For this primary analysis our significance threshold was set at $P < 5 \times 10^{-8}$. Following this primary analysis, we conducted two classes of secondary analyses: (1) joint analysis of melanoma with risk phenotype (nevus or pigmentation) and (2) TWAS.

To ensure robust adjustment for multiple testing, within the joint cutaneous melanoma–nevus and cutaneous melanoma–pigmentation GWAS analyses we Bonferroni-corrected for each of the two risk factor phenotypes (pigmentation and nevus count), as well as accounting for the two classes of secondary analysis (joint GWAS and TWAS). The resulting significance threshold was $(5 \times 10^{-8}) / (2 \times 2) = 1.25 \times 10^{-8}$. Loci reaching this corrected threshold are indicated in bold in Supplementary Tables 7 and 10.

TWAS was performed on expression data from melanocytes and then separately on the three skin tissues within GTEx (sun-exposed, not sun-exposed and fibroblasts) as these were the most enriched tissues in terms of enrichment for

cutaneous melanoma heritability after melanocytes (Extended Data Fig. 5) and are likely to be involved in cutaneous melanoma development.

For the melanocyte TWAS analysis, we Bonferroni-corrected the significance threshold by the number of tested genes in melanocytes multiplied by the two classes of secondary tests and further for the two tissue sets; $0.05 / (3878 \text{ genes} \times \text{two classes} \times \text{two tissue sets}) = 3.22 \times 10^{-6}$.

For the GTEx skin TWAS analysis we Bonferroni-corrected for the total number of tested genes across the tissues multiplied by two classes of secondary tests and further for the two tissue sets; $0.05 / ((8879 + 7458 + 7353 \text{ genes}) \times \text{two classes} \times \text{two tissue sets}) = 5.28 \times 10^{-7}$.

The accuracy of P value calculation for rare SNPs where case/control numbers are imbalanced. The non-normality of the test statistics may cause severely inflated P values due to violation of asymptotic approximations, particularly for imbalanced case–control ratios. While we addressed this for extreme cases by filtering very rare SNPs (Methods), we also investigated whether this could be inflating the P value of rare SNPs included in the meta-analysis by performing 5×10^8 simulations. For each simulation, we first generated genotype data for 21 studies with the same sample size as in our meta-analysis (Supplementary Table 1) assuming Hardy–Weinberg equilibrium for variants with MAF = 0.01.

We then performed association testing for each study and calculated the test statistics to derive an empirical P value of 6.4×10^{-8} when using an asymptotic P value of 5×10^{-8} as the threshold. While imbalanced case–control ratios had minimal impact on the calculation of asymptotic P values for SNPs with MAF = 0.01, as the empirical P value was slightly larger than genome-wide significance we further explored the results of our meta-analysis. Three of our 68 reported variants have a MAF less than 0.01: rs149617956 with MAF = 0.002, rs79356439 with MAF = 0.008 and rs3212371 with MAF = 0.003. All three variants had asymptotic $P < 5 \times 10^{-12}$. We performed 5×10^8 simulations for each of the variants using their MAF and found no simulations had a nominal $P < 5 \times 10^{-12}$. These simulations indicate that the actual P values for these three SNPs are less than $1 / (5 \times 10^8) = 2 \times 10^{-9}$ and have reached genome-wide significance.

Joint analyses of cutaneous melanoma and nevus count and pigmentation. Nevus GWAS meta-analysis. Using beta meta-analysis weighted by s.e.m. as implemented in PLINK v.1.90b5.4, we combined the recently published nevus meta-analysis ($n = 52,506$) (ref. ²⁷) which excluded samples with melanoma but may include a small portion of overlap with the controls used for some melanoma GWAS datasets; participants of the QSkin study with nevus count that are nonoverlapping and unrelated (IBD $\text{pihat} < 0.15$) to the QSkin melanoma case–control set ($n = 12,930$) and the final set of participants not previously included from the Brisbane Twin Nevus Morphology study ($n = 341$) (ref. ²⁷). The total sample size was 65,777.

Pigmentation GWAS. A GWAS for hair color was performed on 352,662 UK Biobank samples not included in the melanoma GWAS individuals who self-reported having blonde, light brown, dark brown or black hair (coded as 1, 2, 3 and 4). Hair color was then treated as a continuous variable and regressed on imputed genotype adjusting for principal components using the same approach as for the melanoma GWAS.

Joint analyses. The melanoma results were then jointly analysed first with nevus count and then with hair color. Two approaches were taken. First, the total confirmed plus self-reported cutaneous melanoma GWAS meta-analysis results were combined with the separate nevus and pigmentation GWAS data using Stouffer's method (P value weighted by per SNP sample n) as implemented in METAL v.2011-03-25 (ref. ⁹⁷). The FUMA platform v.1.3.5 (ref. ⁹⁸) was used to identify independent SNPs with $P < 5 \times 10^{-8}$ on the basis of LD calculations using a reference panel of 10,000 white British UK Biobank individuals; independent SNPs within 1 Mb were considered to be single loci. Secondly, the melanoma and pigmentation/nevus GWAS results were analysed using GWAS-PW v.0.21 (ref. ⁹⁸), which estimates the posterior probability of four possible models for each genetic region: (1) association with cutaneous melanoma only, (2) association with the second trait only, (3) association with both traits (pleiotropic), (4) association with both traits but colocated and independent and (5) no association with either trait. Given that nevus count and pigmentation are believed to act directly on melanoma risk, model 4 seemed unrealistic so we only considered models 1, 2, 3 and 5. For nevus count, SNPs were assigned to blocks using the recommended boundaries for GWAS-PW (<https://bitbucket.org/nygcresearch/ldetect-data>). For cutaneous melanoma and hair color, 50 SNP windows were used for blocks as the default LD blocks contained multiple independent hair color loci. Following the approach described in ref. ²⁷, any locus with a lead SNP reaching $P < 1.25 \times 10^{-8}$ for the combined cutaneous melanoma and nevus/hair color analysis and with a posterior probability > 0.5 that the locus is associated with both traits (model 3) to ensure that the association is not driven by a single trait was declared to be pleiotropically associated with both traits.

Analysis of pigmentation and nevi PRS across melanoma subtypes. For each subject in our study, we calculated two PRSs, using 276 genetic variants

associated with pigmentation and ten genetic variants associated with nevus count. Nevus count SNPs were derived from the same nevus GWAS meta-analysis used for the pleiotropic analysis ($n=65,597$), with independent lead SNPs with $P < 5 \times 10^{-8}$ identified using LD calculations performed in PLINK using a reference panel of 10,000 white British UK Biobank individuals as implemented in the FUMA platform v.1.3.5 (ref. ⁹⁸), with the LD r^2 cutoff for independence < 0.05 . Pigmentation PRS SNPs were selected from the hair color GWAS used for the pleiotropic analysis ($n=352,662$), with independent lead SNPs with $P < 5 \times 10^{-8}$ and LD calculations performed in PLINK using a reference panel of 10,000 white British UK Biobank individuals as implemented in the FUMA platform, with the LD r^2 cutoff for independence < 0.025 . PRS were calculated for each subject by applying the regression coefficient (from the GWAS of pigmentation or nevus count) to the genotype dosages. We then tested whether PRS distribution differed between males and females, across age groups and histology subtypes. In total, we performed 27 comparisons and thus any comparison with P value $< 0.05/27$ ($= 0.00186$) was declared as statistically significant.

GENESIS estimation of heritability and polygenic risk. We used GENESIS (<https://github.com/yandorazhang/GENESIS>; v.2019-06-01; ref. ³²) to estimate the genetic architecture (number of causal SNPs and their effect-size distribution) using the summary-level statistics from the GWAS meta-analysis. Q-Q plot comparing the P values generated from this fitted distribution against the observed P values suggested a three-component Gaussian mixture model for the effect-size distribution. Based on this estimated genetic architecture, we calculated the heritability at the observational scale and the number of SNPs reaching genome-wide significance for a given GWAS with known sample size. Similarly, GENESIS calculated the AUC for an additive polygenic risk prediction model based on a discovery GWAS of known sample size.

UK Biobank melanoma risk phenotype GWAS. Four pigmentary GWAS were performed on UK Biobank participants not included in the melanoma GWAS. (1) Ease of tanning with 367,229 UK Biobank samples who self-reported their ability to tan as 'get very tanned', 'get moderately tanned', 'get mildly or occasionally tanned' or 'never tan, only burn' (coded as 1, 2, 3 and 4). Ease of tanning was treated as a continuous variable and regressed on imputed genotypes adjusting for principal components using the same approach as for the melanoma GWAS of UK Biobank data. (2) Skin color with 370,260 UK Biobank samples who self-reported having 'very fair', 'fair', 'light olive', 'dark olive', 'brown' or 'black' skin color (coded as 1, 2, 3, 4, 5 and 6); skin color was treated as a continuous variable and regressed on imputed genotype adjusting for principal components using the same approach as for the melanoma GWAS of UK Biobank data. (3) Number of childhood sunburns with 320,345 UK Biobank samples who self-reported their sunburn incidents pre-16-years-old. The data were dichotomized into none and at least one pre-16-years-old sunburn incident categories (coded as 1 and 2). Number of childhood sunburns was treated as a binary variable and regressed using a logistic model on imputed genotype adjusting for principal components using the same approach as for the melanoma GWAS of UK Biobank data. (4) Red hair with 120,925 UK Biobank samples who self-reported having either 'red hair' or 'other' (coded as 1 or 2). Red hair was treated as a binary variable and regressed using a logistic model on imputed genotype adjusting for principal components using the same approach as for the melanoma GWAS of UK Biobank data.

LD score regression. As LDSC is sensitive to the quality of input SNPs, GWAS or meta-analysis variants were filtered to the list of high-quality HapMap SNPs provided⁹⁹. Using LDSC program v.1.0.0 (<https://github.com/bulik/ldsc>), genomic inflation (λ), intercept and SNP heritability (h^2) were estimated. The LDSC h^2 estimates were converted to the liability scale by using the proportion of cutaneous melanoma cases in the UK Biobank (3,499 confirmed CM cases and 1,802 self-reported cases divided by 361,194 European ancestry samples = 0.0147).

LDSC of tissue-specific genes. Cutaneous melanoma heritability enrichment for SNPs around tissue-specific genes was assessed by stratified LD score regression as described previously^{28,42} and implemented in the LDSC program v.1.0.0. Briefly, RNA-seq data for all 50 GTEx (v.7) tissue types and primary melanocyte were quantified as reads per kilobase of transcript, per million mapped reads (RPKM) by using RNA-SeQC v.1.18 (ref. ¹⁰⁰) and quantile normalized to reduce batch effect. Tissue-specific genes were defined by calculating the t -statistic of each gene for a given tissue, excluding all samples from the same tissue category. Tissue category assignment for GTEx tissue types was based on the previous publications^{28,101} and melanocytes were defined as 'skin' category together with two types of skin and transformed skin fibroblasts from the GTEx. We selected the top 1,000, 2,000 and 4,000 tissue-specific genes from the t -statistic analysis and added 100 kb each to the transcription start site and transcription end site to define tissue-specific genes annotation. Stratified LDSC was then applied on a joint SNP annotation to estimate the heritability enrichment against the total cutaneous melanoma GWAS data from the current study.

Colocalization of cutaneous melanoma GWAS and eQTLs. We performed colocalization analyses of cutaneous melanoma GWAS signals with eQTL signals

from our melanocyte and 48 GTEx (v.7) tissue eQTL datasets (note that two tissue types that were included for LDSC using expression data were not included here as well as in TWAS analyses due to lack of eQTL data from GTEx), using eQTL and GWAS Causal Variants Identification in Associated Regions (eCAVIAR, v.2.0, <http://genetics.cs.ucla.edu/caviar/> and <https://github.com/thormoz/caviar>)¹⁰². Consistent with the previous study, we used 50 SNPs upstream and downstream of each cutaneous melanoma GWAS lead SNP to extract both GWAS and eQTL summary statistics to be used as the input for eCAVIAR analysis. The LD matrix was calculated using the unphased 1000 Genomes reference set. For the colocalization posterior probability score calculation, we allowed a maximum number of two causal SNPs in each locus. For a given cutaneous melanoma GWAS locus, an eGene with a CLPP score above 1% (0.01) was considered to display a positive colocalization. To avoid reporting spurious effects, we applied a conservative criterion and only reported variants displaying LD $r^2 > 0.9$ with the cutaneous melanoma GWAS lead SNP and eQTL P value below a Bonferroni-corrected cutoff of each dataset (0.05/number of eGenes tested for each tissue dataset)^{67,103–108}.

TWAS. We performed TWAS for the cutaneous melanoma GWAS meta-analysis data using TWAS/FUSION (<http://gusevlab.org/projects/fusion/>) as previously described^{28,41}. TWAS was performed in three separate groups, using eQTL datasets from (1) melanocytes, (2) three skin tissues (sun-exposed, not sun-exposed and fibroblasts) within GTEx v.7 and (3) the rest of GTEx tissue types (a total of 45) by imputing the gene expression phenotypes for the total cutaneous melanoma GWAS meta-analysis data. The analysis parameters were set to allow for multiple prediction models, independent reference LD, additional feature statistics and cross-validation results⁴¹. The total cutaneous melanoma GWAS meta-analysis summary statistics were included with no significance thresholding. For GTEx data, we downloaded the precomputed expression reference weights for GTEx gene expression v.7 RNA-seq across 48 tissue types from the TWAS/FUSION website (<http://gusevlab.org/projects/fusion/>). We computed functional weights from the primary melanocyte RNA-seq data²⁸ one gene at a time. Genes that failed quality control during the heritability check (using minimum heritability P value 0.01) were excluded from further analyses. We restricted the *cis*-locus to 500 kb on either side of the gene boundary.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

Genome-wide summary statistics for the confirmed meta-analysis have been made publicly available at dbGaP ([phs001868.v1.p1](https://www.ncbi.nlm.nih.gov/bioproject/5000000)), with the exclusion of self-reported data from 23andMe and UK Biobank. Results for SNPs with a fixed or random $P < 5 \times 10^{-7}$ from the total meta-analysis are reported in Supplementary Table 7. The total meta-analysis includes self-reported cutaneous melanoma GWAS data from the UK Biobank and 23andMe. The raw genetic and phenotypic UK Biobank data used in this study, which were used under license, are available from: <http://www.ukbiobank.ac.uk/>. The genome-wide summary statistics from 23andMe data were obtained under a data transfer agreement. Further information about obtaining access to the 23andMe summary statistics is available from <https://research.23andme.com/collaborate/>. Source data for Fig. 2, Extended Data Figs. 4–6 and Supplementary Figs. 2 and 3 are available with the paper.

References

- Purcell, S. et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
- Chang, C. C. et al. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* **4**, 7 (2015).
- Das, S. et al. Next-generation genotype imputation service and methods. *Nat. Genet.* **48**, 1284–1287 (2016).
- DerSimonian, R. & Laird, N. Meta-analysis in clinical trials. *Control. Clin. Trials* **7**, 177–188 (1986).
- Machiela, M. J. & Chanock, S. J. LDassoc: an online tool for interactively exploring genome-wide association study results and prioritizing variants for functional investigation. *Bioinformatics* **34**, 887–889 (2018).
- Willer, C. J., Li, Y. & Abecasis, G. R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190–2191 (2010).
- Watanabe, K., Taskesen, E., Van Bochoven, A. & Posthuma, D. Functional mapping and annotation of genetic associations with FUMA. *Nat. Commun.* **8**, 1826 (2017).
- Bulik-Sullivan, B. K. et al. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.* **47**, 291–295 (2015).
- DeLuca, D. S. et al. RNA-SeQC: RNA-seq metrics for quality control and process optimization. *Bioinformatics* **28**, 1530–1532 (2012).
- Finucane, H. K. et al. Heritability enrichment of specifically expressed genes identifies disease-relevant tissues and cell types. *Nat. Genet.* **50**, 621–629 (2018).

102. Hormozdiari, F. et al. Colocalization of GWAS and eQTL signals detects target genes. *Am. J. Hum. Genet.* **99**, 1245–1260 (2016).
103. Stacey, S. N. et al. A germline variant in the TP53 polyadenylation signal confers cancer susceptibility. *Nat. Genet.* **43**, 1098–1103 (2011).
104. Chahal, H. S. et al. Genome-wide association study identifies 14 novel risk alleles associated with basal cell carcinoma. *Nat. Commun.* **7**, 12510 (2016).
105. Ostrom, Q. T. et al. Sex-specific glioma genome-wide association study identifies new risk locus at 3p21.31 in females, and finds sex-differences in risk at 8q24.21. *Sci. Rep.* **8**, 7352 (2018).
106. Melin, B. S. et al. Genome-wide association study of glioma subtypes identifies specific differences in genetic susceptibility to glioblastoma and non-glioblastoma tumors. *Nat. Genet.* **49**, 789–794 (2017).
107. Zhou, Y., Wu, H., Zhao, M., Chang, C. & Lu, Q. The Bach family of transcription factors: a comprehensive review. *Clin. Rev. Allergy Immunol.* **50**, 345–356 (2016).
108. Milovic-Holm, K., Kriehhoff, E., Jensen, K., Will, H. & Hofmann, T. G. FLASH links the CD95 signaling pathway to the cell nucleus and nuclear bodies. *EMBO J.* **26**, 391–401 (2007).

Acknowledgements

NCI: This study was supported by the Intramural Research Program of the Division of Cancer Epidemiology and Genetics, National Cancer Institute (NCI), National Institutes of Health (NIH) and Department of Health and Human Services (DHHS). **AOCS/OCAC/SEARCH:** AOCS/OCAC/SEARCH is accessible via European Genome-Phenome Archive. We acknowledge their support and data, and the contribution of the study nurses, research assistants and all clinical and scientific collaborators in generation of these data. We also acknowledge their funding sources: OCAC (NIH grant no. U19CA148112), SEARCH team (Cancer Research UK grant no. C490/A16561), AOCS (US Army Medical Research and Materiel Command under grant no. DAMD17-01-1-0729, The Cancer Council Victoria, Queensland Cancer Fund, The Cancer Council New South Wales, The Cancer Council South Australia, The Cancer Foundation of Western Australia, The Cancer Council Tasmania and the National Health and Medical Research Council of Australia (NHMRC) (grant nos. ID400413 and ID400281, as well as support from S. Boldeman, the Agar family, Ovarian Cancer Action (UK), Ovarian Cancer Australia and the Peter MacCallum Foundation). **MelaNostrum Consortium:** We thank the participants of the MelaNostrum Consortium from Italy (Genoa, L'Aquila, Rome, Padua, Milan, Florence and Bergamo), Spain (Valencia and Barcelona), Greece (Athens) and Cyprus (Nicosia) who provided data and biospecimens for this study. The Consortium is partially supported by the Intramural Research Program of the Division of Cancer Epidemiology and Genetics, NCI, NIH, DHHS. Funding for the University of Genoa and Genetics of Rare Cancers, Ospedale Policlinico San Martino came from Italian Ministry of Health 5 × 1000 per la Ricerca Corrente to Ospedale Policlinico San Martino and AIRC IG 15460. The research at the Melanoma Unit in Barcelona was supported by the Spanish Fondo de Investigaciones Sanitarias grant nos. PI15/00716 and PI15/00956 cofinanced by FEDER 'Una manera de hacer Europa'; CIBER de Enfermedades Raras of the Instituto de Salud Carlos III, Spain, cofinanced by European Development Regional Fund 'A way to achieve Europe' ERDF; AGAUR 2014_SGR_603 of the Catalan Government, Spain; European Commission, contract no. LSHC-CT-2006-018702 (GenoMEL) and by the European Commission under the 7th Framework Programme, Diagnostics; 'Fundació La Marató de TV3' grant no. 201331-30, Catalonia, Spain; 'Fundación Científica de la Asociación Española Contra

el Cáncer' grant no. GCB15152978SOEN, Spain, and CERCA Programme/Generalitat de Catalunya. Melanoma research at the Department of Dermatology, University of L'Aquila, Italy was supported by the Italian Ministry of the University and Scientific Research (PRIN-2012 grant no. 2012JJX494). **Q-MEGA/QTWIN:** The Q-MEGA/QTWIN study was supported by the Melanoma Research Alliance, the NIH NCI (grant nos. CA88363, CA83115, CA122838, CA87969, CA055075, CA100264, CA133996 and CA49449), the NHMRC (grant nos. 200071, 241944, 339462, 380385, 389927, 389875, 389891, 389892, 389938, 443036, 442915, 442981, 496610, 496675, 496739, 552485, 552498 and APP1049894), the Cancer Councils New South Wales, Victoria and Queensland, the Cancer Institute New South Wales, the Cooperative Research Centre for Discovery of Genes for Common Human Diseases, Cerylid Biosciences (Melbourne), the Australian Cancer Research Foundation, The Wellcome Trust (grant no. WT084766/Z/08/Z) and donations from N. and S. Hawkins. S. MacGregor acknowledges fellowship support from the Australian National Health and Medical Research Council and from the Australian Research Council.

Please see the Supplementary Note for additional acknowledgments.

Author contributions

M.T.L., M.M.I. and M.H.L. conceptualized and designed the project. D.T.B., S. MacGregor, M.T.L. and S.J.C. provided funding support. M.T.L., D.T.B., S. MacGregor, M.J.M., J.S., M.M.I. and M.H.L. interpreted the results and supervised the study. M.J.M., M.T.L., M.M.I., K.B., J.C. and M.H.L. wrote the manuscript. J.S., M.M.I., K.B., T.Z., J.C., D.L.D. and M.H.L. analyzed the data. A.J. Stratigos, P. Ghiorzo, S.P. and E.N. coordinated the study and collected the data. M.T.L., D.T.B., S. MacGregor, M.J.M., A.J. Stratigos, P. Ghiorzo, M.B., D.C., J.C., M.C.F., T.Z., M.R., A.J.T., C.M., J. Martinez, A. Hadjisavvas, L.S., I.S., R.S., X.R.Y., A.M.G., M.P., K.P.K., L.P., P.Q., C.P., L.C., M.Z., P. Gimenez-Xavier, A.R., L.E., S. Manoukian, L.R., B.H.S., M.A.L., L.D.R., D.M., M. Mandala, K.K., L.A.A., C.I.A., P.A.A., M.A., E.A., H.P.S., V.B., B.D., L.M.B., K.P.B., W.V.C., V.C., J.E.C., T.D., M.F., S.F., E.F., S.S., P. Galan, Z.G., E.M.G., S.G., A.G., N.A.G., J. Hansson, M. Harland, J. Harris, P.H., A. Henders, M. Hočevár, V.H., D.H., C.I., R.K., J. Lang, G.M.L., J.E.L., X.L., J. Lubiński, R.M.M., M. Malt, J. Malvey, K.M., H.M., A.M., E.K.M., R.E.N., S.N., D.R.N., H.O., N.O., L.G.F., J.A.P., A.A.Q., G.L.R., J.R., C. Requena, C. Rowe, N.J.S., M. Sanna, D.S., H.S., L.A.S., M. Smithers, F.S., A.J. Swerdlow, N.V.D.S., N.A.K., A. Visconti, L. Wallace, S.V.W., L. Wheeler, R.A.S., A. Hutchinson, K.J., M. Malasky, A. Vogt, W.Z., K.A.P., D.E.E., J. Han, B.H., N.K.H., P.A.K., C.B., G.W.M., C.M.O., C.H., A.M.D., N.G.M., E.E., G.J.M., G.L., P.D.P.P., D.F.E., J.H.B., A.E.C., G.A., D.L.D., D.C.W., H.G., A.D.N., M.A.T., J.A.N.B., K.P., S.J.C., K.M.B., F.D., S.P., E.N., J.S., M.M.I. and M.H.L. participated in data collection, results interpretation and manuscript review.

Competing interests

The authors declare no competing interests.

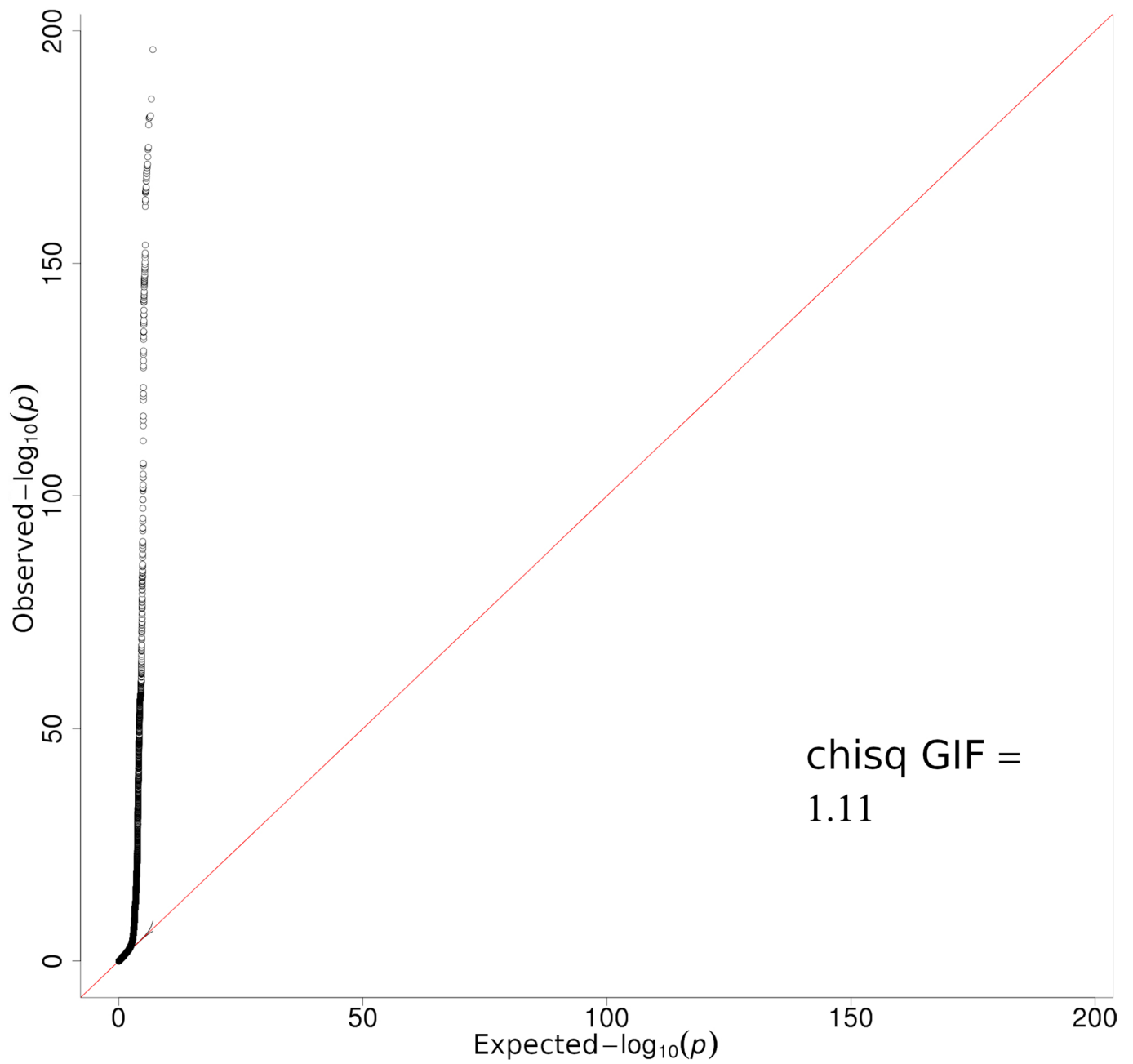
Additional information

Extended data is available for this paper at <https://doi.org/10.1038/s41588-020-0611-8>.

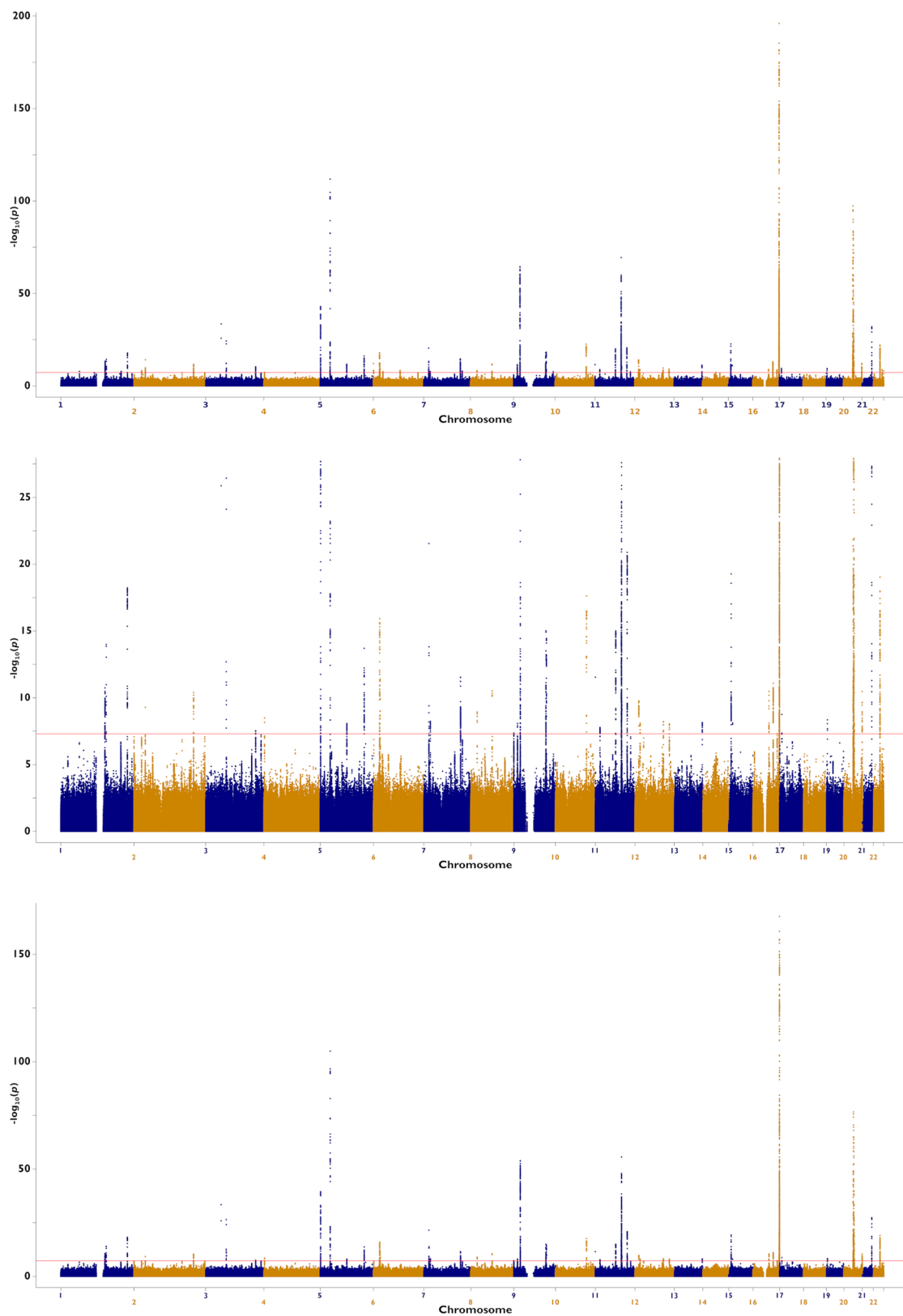
Supplementary information is available for this paper at <https://doi.org/10.1038/s41588-020-0611-8>.

Correspondence and requests for materials should be addressed to M.T.L., M.M.I. or M.H.L.

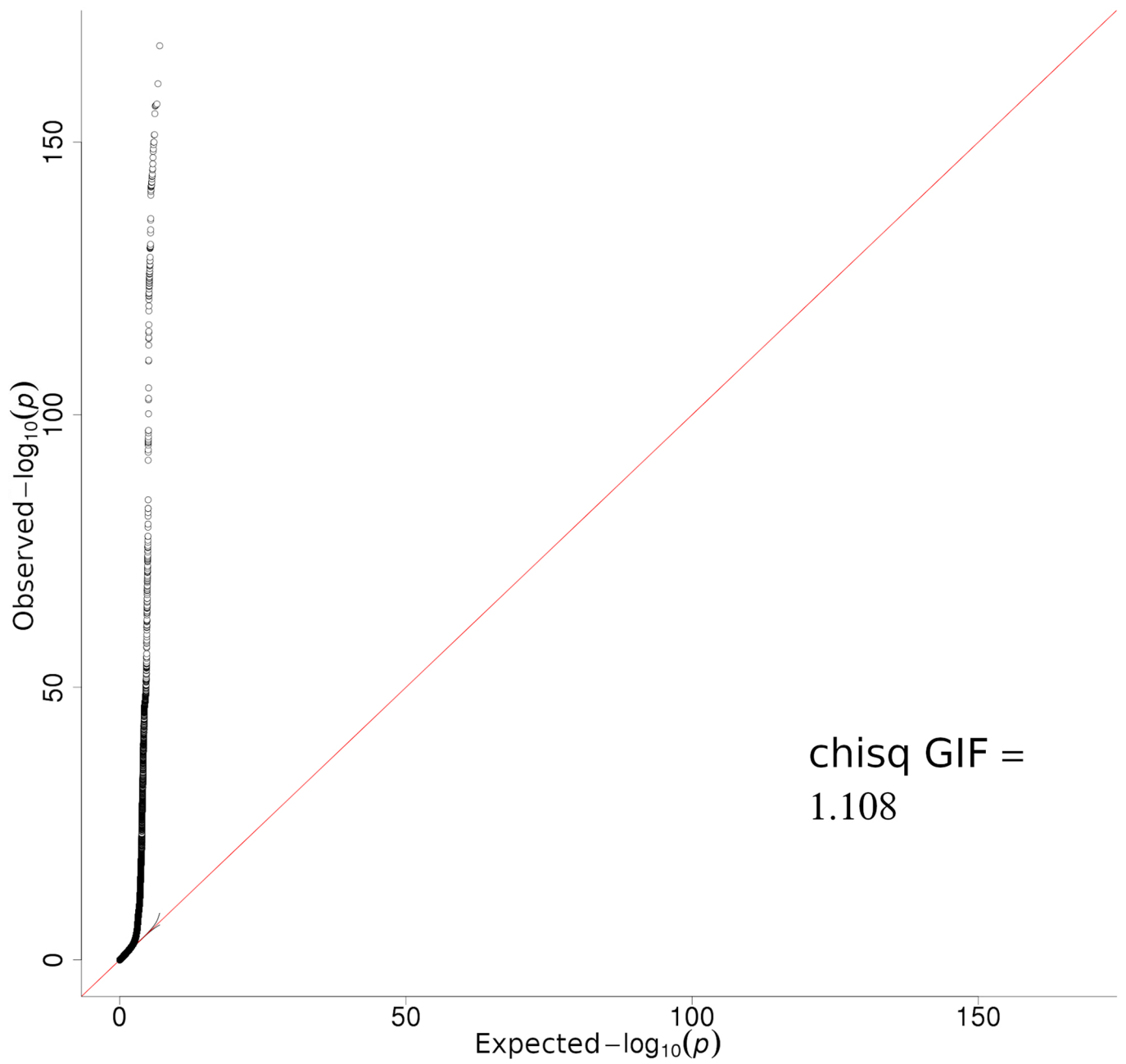
Reprints and permissions information is available at www.nature.com/reprints.



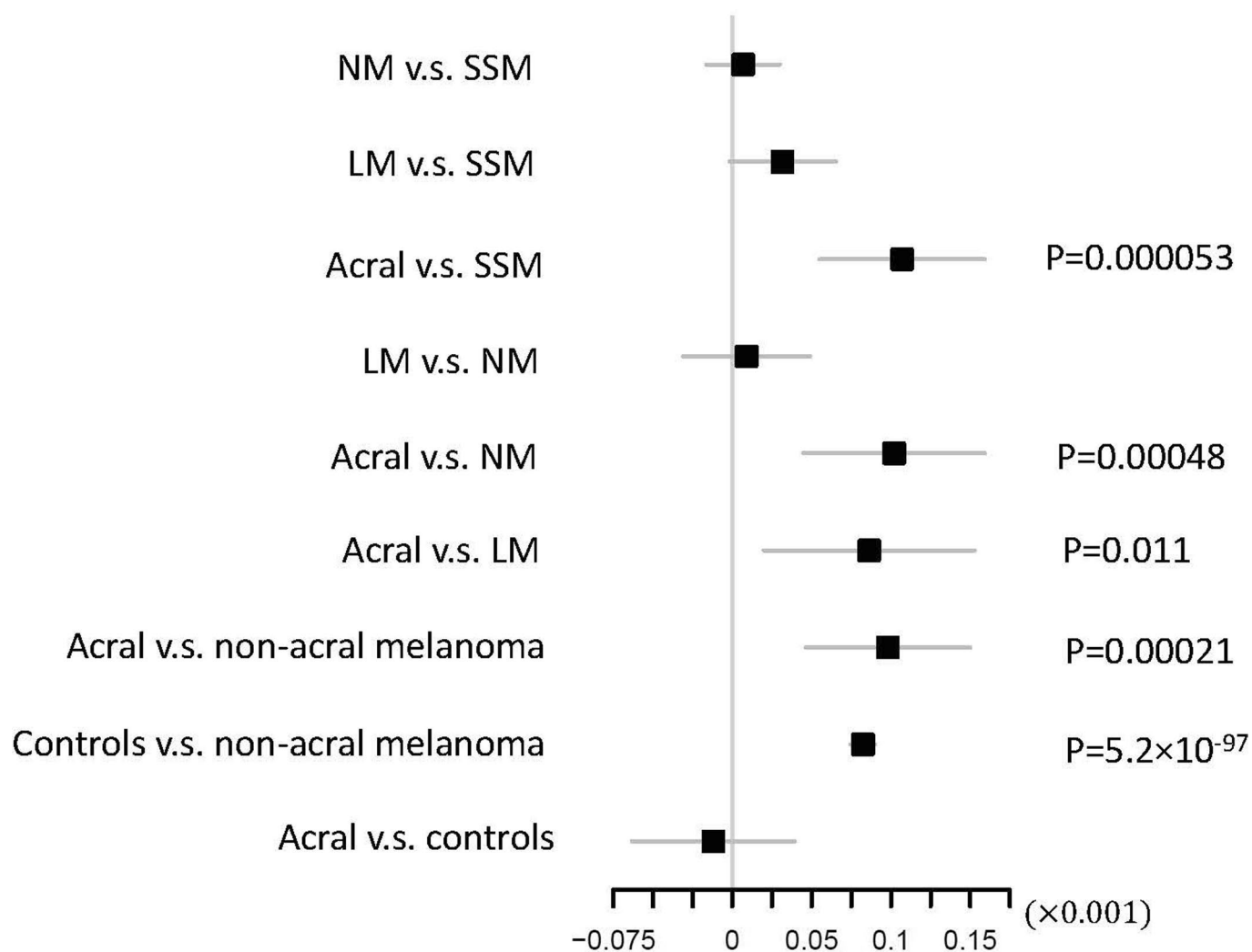
Extended Data Fig. 1 | Quantile-Quantile plot of total CM meta-analysis. Quantile-quantile plots of negative \log_{10} two-sided P value derived from a fixed-effects inverse-variance weighted meta-analysis of $\log(\text{OR})$ effect-sizes derived from the logistic regression GWAS listed in Supplementary Table 1. All confirmed and self-report cases are included, with a total sample size of 36,760 melanoma cases and 375,188 controls.



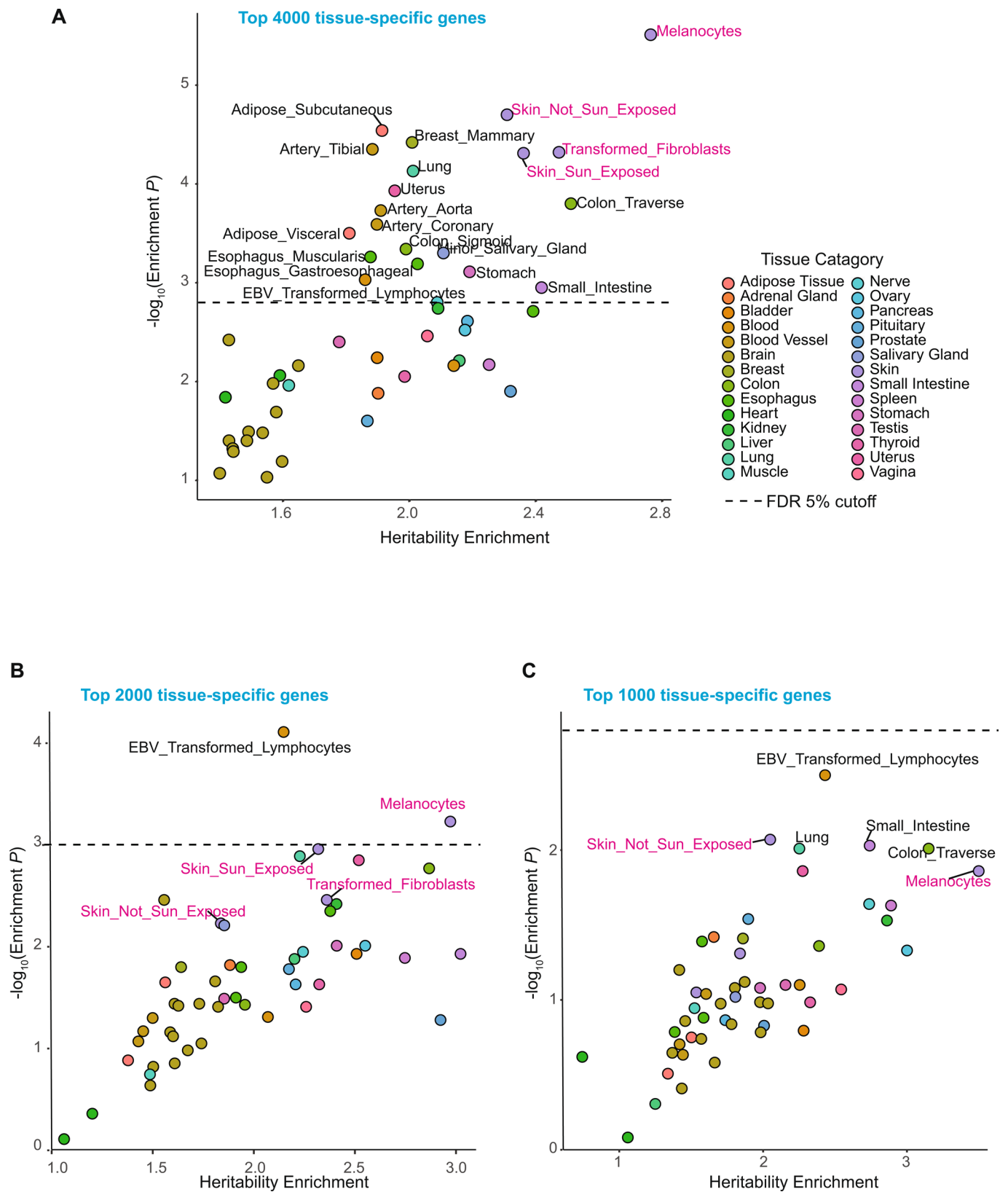
Extended Data Fig. 2 | Manhattan plots of melanoma risk loci from total and confirmed-only GWAS meta-analyses. Negative log₁₀ two-sided P value derived from a fixed-effects inverse-variance weighted meta-analysis of log(OR) effect-sizes derived from the logistic regression GWAS (y-axis) are plotted by their chromosome position. The confirmed-only analysis included 30,134 cases with histopathologically confirmed CM, and 81,415 controls. The total CM meta-analysis includes all confirmed and self-report cases, with a total sample size of 36,760 CM cases and 375,188 controls. Multiple-testing corrected genome-wide significance threshold was $P < 5 \times 10^{-8}$. We display in order the total CM meta-analysis without limiting the y-axis; the pathologically confirmed CM cases only meta-analysis with the y-axis limited to 1×10^{-25} and without a limit to more clearly display loci other than *MC1R*.



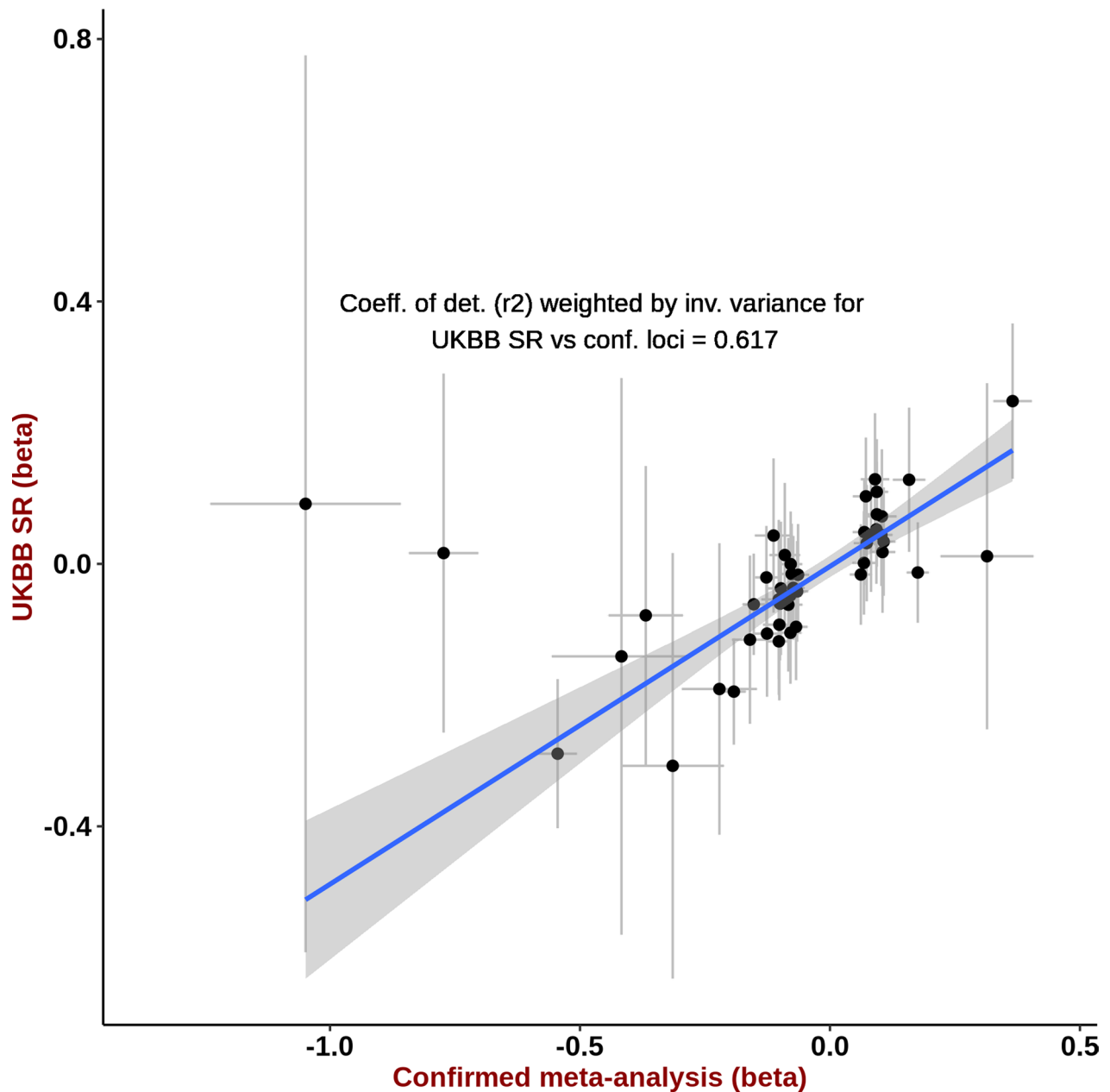
Extended Data Fig. 3 | Quantile-Quantile plot of confirmed-only CM meta-analysis. Quantile-quantile plots of negative \log_{10} two-sided P value derived from a fixed-effects inverse-variance weighted meta-analysis of $\log(\text{OR})$ effect-sizes derived from the logistic regression GWAS listed in Supplementary Table 1. Only cases with histopathologically confirmed CM are included, with a total sample size of 30,134 melanoma cases and 81,415 controls.



Extended Data Fig. 4 | Distribution of pigmentation polygenic risk scores across melanoma histological subtypes. The figure shows whether PRS defined based on SNPs associated with hair color differ across CM histological types (Methods; SSM: superficial spreading melanoma; NM: nodular melanoma; LM: lentigo melanoma; Acral: acral lentiginous melanoma). The higher the PRS the lighter the hair color. When comparing subtype 1 vs. subtype 2, we report the effect size for the linear regression of PRS on subtype 1, including study and principal components as covariates to control for population stratification. The regression coefficient, 95% confidence interval, and statistical significance are shown. The positive beta indicates the PRS is higher in subtype 2 (for example, nonacral melanomas). This analysis included 9828 SSM, 2137 NM, 900 LM, 353 acral melanoma cases and 44676 controls. Two-sided t-statistic was used for testing significance. P values reported were not adjusted for multiple comparison.



Extended Data Fig. 5 | LD score regression plots. LD score regression was performed for the top 4000 (A) 2000 (B) and 1000 (C) tissue-specific genes from melanocyte and GTEx tissue types (v7 datasets), to assess the enrichment of melanoma heritability in these genomic regions using summary statistics from Total CM GWAS meta-analysis. The level of enrichment and P values are shown, with an FDR = 0.05 cutoff marked as a dashed horizontal line (See Methods for statistical test). Tissue categories are color-coded, and a subset of top individual tissue types are shown on the plot. Tissue types from “Skin” category including melanocytes are highlighted in magenta.



Extended Data Fig. 6 | Effect sizes for confirmed-only meta-analysis versus UKBB self-report set. For each independent genome-wide significant ($P < 5 \times 10^{-8}$) lead SNP from the confirmed-only meta-analysis (30,134 melanoma cases and 81,415 controls), we plot on the Y-axis UK Biobank self-report GWAS (UKBB SR) $\log(\text{OR})$ and standard error from a logistic regression GWAS (1,802 self-report CM cases and 7,208 controls) and on the X-axis we plot the $\log(\text{OR})$ and standard error from a fixed-effects inverse-variance weighted meta-analysis of $\log(\text{OR})$ effect-sizes derived from the logistic regression GWAS for confirmed melanoma cases listed in Supplementary Table 1. We also report the r^2 correlation from the linear regression of UKBB SR $\log(\text{OR})$ on the confirmed meta-analysis estimates, weighted by their standard error.

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection

Data were generated at many sites using standard genotype calling settings in Illumina GenomeStudio/BeadStudio (v2.0.4); see Online Methods for full description. RNA-seq data for all the GTEx (v7) tissue types and primary melanocyte were quantified as RPKM using RNA-SeQC (v1.18)

Data analysis

Imputation was performed using the Michigan Imputation Server and Minimac3. GWAS analyses were performed in PLINK (v1.90b5.4). Specific QC and any additional software used for analyses, including version numbers include: GCTA (v1.26.0); GENESIS (v 2019-06-01, <https://github.com/yandorazhang/GENESIS>); eQTL and GWAS CAusal Variants Identification in Associated Regions (eCAVIAR, <http://genetics.cs.ucla.edu/caviar/>, <https://github.com/fhormoz/caviar>); TWAS/FUSION (<http://gusevlab.org/projects/fusion/>); Ingenuity Pathway Analysis; Illumina GenomeStudio/BeadStudio (v2.0.4); Minimac3; METAL (version 2011-03-25); FUMA (v1.3.5); GWAS-PW (v0.21); LD boundaries for GWAS-PW (<https://bitbucket.org/nygresearch/ldetect-data>); LD score regression (LDSC, v1.0.0 <https://github.com/bulik/ldsc>); RNA-SeQC (v1.18).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Genome-wide summary statistics for the confirmed meta-analysis have been made publicly available at dbGaP (phs001868.v1.p1). Results for SNPs with a fixed or random $P < 5 \times 10^{-7}$, from the total meta-analysis are reported in Supplementary Table 15. The total meta-analysis includes self-report cutaneous melanoma GWAS data from the UK Biobank and 23andMe. The raw genetic and phenotypic UK Biobank data used in this study, which were used under license, are available from:

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences Behavioural & social sciences Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	The previous largest GWAS meta-analysis of melanoma included GWAS data 12,874 cases and 23,203 controls plus a replication set of ~ 3,116 CMM cases and 3,206 controls. This identified 20 risk loci. As loci discovery scales proportionally to sample size (Visscher AJHG 2012 PMID 22243964), and we assembled all available melanoma GWAS datasets (total sample size 36,760 melanoma cases and 375,188 controls) we determined our sample size was sufficient to identify additional new loci for melanoma.
Data exclusions	As genotype imputation, and genome-wide association studies, are sensitive to data quality both per SNP and per individual, as well as population stratification and unaccounted for relationships, we used pre-established criteria to filter both SNPs and sample. These criteria included but were not limited to high missingness, high heterozygosity, first degree relationship with another sample, or being an ancestry outliers.
Replication	This meta-analysis utilises a combined melanoma GWAS dataset far larger than any other in existence. Therefore any attempted replication would be grossly underpowered. However all associated SNPs were explored for consistency across contributing GWAS (Supplementary Figure 7).
Randomization	GWAS were analysed by contributing cohort. Principal components or country of original covariates (See Methods) were included to control for ancestry.
Blinding	GWAS data were derived from population based case-control observational studies and all available post-QC/cleaning samples were used; there is no meaningful blinding strategy that could be applied, nor required as this was not a randomised control trial.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

Methods

n/a	Involved in the study	n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies	<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines	<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology	<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms		
<input type="checkbox"/>	<input checked="" type="checkbox"/> Human research participants		
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data		

Human research participants

Policy information about [studies involving human research participants](#)

Population characteristics	21 genome-wide association studies of histopathologically confirmed cutaneous melanoma were included, with the total numbers and country of origin included in Supplementary Table 1. 2 further self-reported melanoma datasets were included; these are also detailed in Supplementary Table 1. The only covariates included were country of origin variables, or principal components derived from PCA with the 1000 Genomes Project samples, to control for population structure. See Online Methods for details. GWAS were performed using a range of genotyping arrays; these are reported in Supplementary Table 1 and the Online Methods.
Recruitment	Included samples were recruited using different approaches in each of the contributing GWAS, and included both samples collected from the general population, as well as samples recruited from hospitals and treatment clinics. In addition to the description of each cohorts collection in the Online Methods, full recruitment details for each study can be found in the following papers:

GenoMEL Phase 1 and 2 - Barrett et al. PMID 21983787 and Law et al. 26237428

MDACC - Amos et al. PMID 21926416

AMFS - Cust et al. PMID 19887461

Q-MEGA - Baxter et al. PMID 18361720

GSEdinCIDRulcer - see references 2,34-38 of the Supplementary Note.

WAMHS - Ward et al. PMID 21474410

Essen-Heidelberg - Law et al. PMID 26237428

MELARISK - Barrett et al. PMID 21983787; Law et al. PMID 26237428; Newton-Bishop et al. PMID 20647408

Harvard - Song et al. PMID 22230721

NCI_CPSII+PLCO+Rose - Goldstein et al. PMID 29036293

UK Biobank confirmed - Bycroft et al. PMID 30305743

MIA_PAH - reported here for the first time, details in the Supplementary Note.

EPIGENE - Kvaskoff et al. PMID 24083994

QSKIN - Olsen et al. PMID 22933644

Greek - Kypreou et al. PMID 27015455; Pectasides et al. PMID 19139440

Italy - Ghiorzo et al. PMID 16893909; Ghiorzo et al. PMID 23167872; Bruno et al. PMID 26775776; Fagnoli et al. PMID

15057047; Fagnoli et al. PMID 16567973; Fagnoli et al. PMID 18368129; Pellegrini et al. PMID 28146043; Menin et al. PMID

21672182; Fitzpatrick et al. PMID 3377516; Schlafly et al. PMID 31730655; Gu et al. PMID 30060076; Shi et al. PMID 24686846;

Landi et al. PMID 16809487; Landi et al. PMID 15998953.

Spain - Puig et al. PMID 15860862; Potrony et al. PMID 28103633; Nagore et al. PMID 26875008.

Michigan - Fritsche et al. PMID 29779563

BNMS - reported here for the first time, details in the Supplementary Note.

UK Biobank Self-report - Bycroft et al. PMID 30305743

23andMe - Ransohoff et al. PMID 28212542

Ethics oversight

All contributing GWAS were overseen by their contributing human research ethics committee; see Supplementary Notes for specifics (Human Research Ethics Committee of QIMR Berghofer Medical Research Institute, Brisbane, Australia; National Cancer Institute, National Institutes of Health; Each participating center in MelaNostrum; and Leeds University).

Note that full information on the approval of the study protocol must also be provided in the manuscript.