

Genome-wide association study identifies variants at 9p21 and 22q13 associated with development of cutaneous nevi

Mario Falchi^{1,2,12}, Veronique Bataille^{1,3,12}, Nicholas K Hayward^{4,12}, David L Duffy⁵, Julia A Newton Bishop⁶, Tomi Pastinen^{7,8}, Alessandra Cervino¹, Zhen Z Zhao⁹, Panos Deloukas¹⁰, Nicole Soranzo^{1,10}, David E Elder¹¹, Jennifer H Barrett⁶, Nicholas G Martin⁵, D Timothy Bishop^{6,12}, Grant W Montgomery^{9,12} & Timothy D Spector^{1,12}

A high melanocytic nevi count is the strongest known risk factor for cutaneous melanoma. We conducted a genome-wide association study for nevus count using 297,108 SNPs in 1,524 twins, with validation in an independent cohort of 4,107 individuals. We identified strongly associated variants in *MTAP*, a gene adjacent to the familial melanoma susceptibility locus *CDKN2A* on 9p21 (rs4636294, combined $P = 3.4 \times 10^{-15}$), as well as in *PLA2G6* on 22q13.1 (rs2284063, combined $P = 3.4 \times 10^{-8}$). In addition, variants in these two loci showed association with melanoma risk in 3,131 melanoma cases from two independent studies, including rs10757257 at 9p21, combined $P = 3.4 \times 10^{-8}$, OR = 1.23 (95% CI = 1.15–1.30) and rs132985 at 22q13.1, combined $P = 2.6 \times 10^{-7}$, OR = 1.23 (95% CI = 1.15–1.30). This provides the first report of common variants associated to nevus number and demonstrates association of these variants with melanoma susceptibility.

The worldwide incidence of cutaneous melanoma in populations of European descent has risen rapidly over the past 30 years, more so than that of any other malignancy¹. The World Health Organization estimates 132,000 new cases of melanoma per year globally. Approximately 10% of malignant melanoma cases occur in multiplex families, with the *CDKN2A* locus in 9p21 accounting for susceptibility in 20–57% of all melanoma families². However, remaining cases of familial melanoma do not carry mutations in *CDKN2A* and additional tumor suppressor genes at 9p21 and elsewhere have been proposed^{3,4}. Large numbers of nevi are the strongest known risk factor for melanoma⁵ and in many familial cases susceptibility to melanoma is associated with an excess of nevi. Linkage to 9p21 has been found in two

genome-wide analyses of nevus count^{6,7}, suggesting that shared genetic factors might be involved in melanoma susceptibility and neovogenesis. Understanding the genetic bases of nevus formation is therefore an important step in understanding melanoma etiology.

To search for loci involved in nevus count, we carried out a genome-wide association study (GWAS) in 1,524 healthy adult female twins from the TwinsUK registry using 297,108 SNPs with minor allele frequency (MAF) > 1% on Illumina HumanHap 300k duo chips (see Online Methods). Total body nevus count was evaluated by trained research nurses and defined as the sum of all nevi at least 2 mm in diameter. Comparison of the observed and expected distributions (Supplementary Fig. 1) showed no evidence for inflation of the test statistics (inflation factor = 1.006) and highlighted 46 SNPs having P values < 10^{-4} against the expected 29. We selected 12 SNPs for

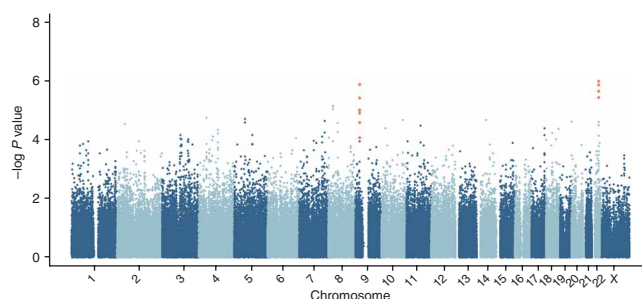


Figure 1 Genome-wide association plot for total nevus count in the TwinsUK sample. The SNPs typed and replicated in the Brisbane Twin Nevus Study sample are highlighted in red.

¹Department of Twin Research & Genetic Epidemiology, Kings College London, St. Thomas' Hospital Campus, London, UK. ²Genomic Medicine, Hammersmith Hospital, Imperial College London, London, UK. ³Dermatology Department, West Hertfordshire NHS Trust, Hemel Hempstead General Hospital, Herts, UK. ⁴Oncogenomics, Queensland Institute of Medical Research, Brisbane, Australia. ⁵Genetic Epidemiology, Queensland Institute of Medical Research, Brisbane, Australia. ⁶Section of Epidemiology and Biostatistics, Leeds Institute of Molecular Medicine, University of Leeds, Leeds, UK. ⁷Department of Human and Medical Genetics, McGill University, Montréal, Canada. ⁸McGill University and Genome Québec Innovation Centre, Montréal, Canada. ⁹Molecular Epidemiology, Queensland Institute of Medical Research, Brisbane, Australia. ¹⁰Human Genetics Department, Wellcome Trust Sanger Institute, Hinxton, Cambridge, UK. ¹¹Department of Pathology and Laboratory Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, USA. ¹²These authors contributed equally to this work. Correspondence should be addressed to T.D.S. (Tim.spector@kcl.ac.uk) or M.F. (m.falchi@imperial.ac.uk).

Table 1 Total nevus count summary association results and follow-up for the 9p21 and 22q13 regions

Marker	Chr.	MAF Allele TwinsUK	Beta (s.e.) TwinsUK	<i>P</i> TwinsUK	Beta (s.e.) BTNS ^a	<i>P</i> BTNS	Combined <i>P</i>	Heterogeneity <i>P</i>
rs4636294	9	0.48 (A)	0.188 (0.042)	9.8×10^{-6}	0.197 (0.037)	6.6×10^{-11}	3.4×10^{-15}	0.48
rs2218220	9	0.48 (C)	0.188 (0.043)	1.0×10^{-5}	0.194 (0.037)	1.1×10^{-10}	5.7×10^{-15}	0.57
rs751173	9	0.46 (G)	0.166 (0.042)	8.5×10^{-5}	0.168 (0.037)	2.6×10^{-10}	9.9×10^{-14}	0.95
rs7023329	9	0.50 (A)	0.184 (0.042)	1.3×10^{-5}	0.164 (0.036)	2.6×10^{-7}	2.6×10^{-11}	0.53
rs10757257	9	0.40 (A)	-0.205 (0.042)	1.3×10^{-6}	-0.147 (0.037)	1.3×10^{-6}	2.9×10^{-11}	0.11
rs1335510	9	0.41 (G)	-0.198 (0.043)	3.9×10^{-6}	-0.145 (0.037)	5.3×10^{-6}	3.2×10^{-10}	0.12
rs1341866	9	0.41 (G)	-0.189 (0.043)	1.0×10^{-5}	-0.143 (0.037)	5.2×10^{-6}	6.2×10^{-10}	0.16
rs10811629	9	0.41 (G)	-0.175 (0.042)	2.7×10^{-5}	-0.089 (0.037)	4.7×10^{-3}	4.3×10^{-6}	0.03
rs2284063	22	0.35 (G)	-0.209 (0.045)	2.3×10^{-6}	-0.078 (0.039)	3.4×10^{-4}	3.4×10^{-8}	0.03
rs6001027	22	0.35 (G)	-0.209 (0.045)	3.8×10^{-6}	-0.078 (0.039)	3.6×10^{-4}	5.0×10^{-8}	0.04
rs132985	22	0.46 (T)	-0.209 (0.043)	1.0×10^{-6}	-0.055 (0.037)	1.5×10^{-2}	4.0×10^{-6}	4.0×10^{-3}
rs738322	22	0.47 (A)	-0.209 (0.043)	1.4×10^{-6}	-0.056 (0.037)	4.6×10^{-2}	2.5×10^{-5}	2.0×10^{-3}

Effect size (beta) and standard error (s.e.) are expressed in log (nevi). The meta-analytic *P* values are calculated using a weighted Z-score method.

^aBTNS Beta and s.e. are reported on the log scale to make them comparable to the TwinsUK results. *P* values have been evaluated using cube-root transformation, which better fits the BTNS data (Online Methods).

follow-up that clustered in either the 9p21 or 22q13 chromosomal regions (Fig. 1). We replicated these SNP associations in an independent sample of 4,107 adolescent subjects of European ancestry from the Brisbane Twin Nevus Study (BTNS) for whom nevus count had been evaluated using a similar protocol (see Online Methods).

At the replication stage (Table 1), the strongest association signal was observed for rs4636294 (combined $P = 3.4 \times 10^{-15}$). This marker is located in the 5' UTR of the *MTAP* (methylthioadenosine phosphorylase) gene located at 9p21, next to *CDKN2A* (Fig. 2a). The other replicated region on 22q13 (Fig. 2b) showed the highest combined association signal with rs2284063 (combined $P = 3.4 \times 10^{-8}$). This SNP lies in the intron between the second and third exon of *PLA2G6*, belonging to the phospholipase A2 (PLA2) superfamily of genes⁸. The 9p21 and 22q13 loci accounted for 1.5% (UK) and 3.0% (Australia) and 1.9% (UK) and 0.7% (Australia) of nevus count variance, respectively.

The combined additive effect of rs4636294 and rs2284063 markedly increases the observed number of nevi. Indeed, subjects in the TwinsUK discovery sample homozygous for both variants (7%) had more than double the number of nevi (median number of nevi = 30) compared to subjects homozygous for the protective alleles (median number of nevi = 12.5). Subjects in the Australian BTNS sample are characterized by an overall larger number of nevi (median number of nevi = 93 compared to 21 in the TwinsUK sample). Homozygous subjects for both variants in the Australian population (11%) also showed greater median number of nevi (116) than homozygous

individuals for the protective alleles (79) (Fig. 3 and Supplementary Table 1). Sun exposure varies greatly between the UK and Australian samples. The number of sunburns over a lifetime and the skin type (measured according to the Fitzpatrick classification) were not related to the number of nevi in either sample (see also Supplementary Table 2 for the distributions in the two countries). Therefore, adjusting for these two measures did not affect our results.

We explored whether these two genetic regions were also involved directly or indirectly in melanoma susceptibility by analyzing two melanoma case-control samples: the Queensland study of Melanoma, Environment and Genetic Associations (Q-MEGA), which included 1,734 cases and 1,811 controls of northern European origin, and the Leeds melanoma case-control study of North UK, which included 1,274 incident melanoma cases, 123 cases with a family history of melanoma, and 1,070 controls, plus 1,395 controls from the Wellcome Trust Case Control Consortium (WTCCC) 1958 birth cohort samples (see Online Methods for cohort details).

We derived joint odds ratios (OR) and confidence intervals under a fixed-effects model for each SNP, and associated *P* values from the standard normal distribution. As expected, we observed a small degree of heterogeneity for some of the SNPs shown by the lack of complete agreement for the highest association signals in the two samples. The highest combined ORs were 1.23 (95% CI = 1.15–1.30) for rs10757257 at 9p21 and 1.23 (95% CI = 1.15–1.30) for rs132985 at 22q13 (Table 2 and Supplementary Fig. 2). The top SNP-associated results in the nevus count analyses at each of these loci were as follows:

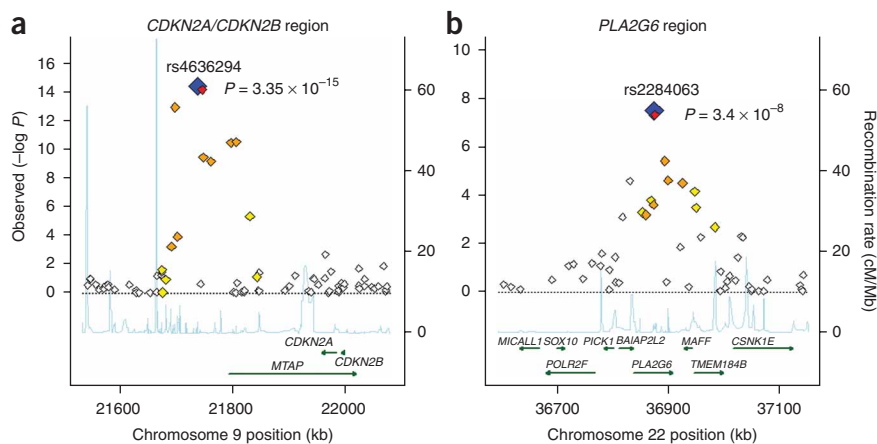


Figure 2 Regional plots. (a) Chromosome 9p21. (b) Chromosome 22q13. Meta analysis \log_{10} *P* values are plotted as a function of genomic position (build 36). The *P* values for the lead SNPs are denoted by large blue (combined discovery and replication) diamonds. Proxies are indicated with diamonds of smaller size, with colors assigned based on the pairwise r^2 values with the lead SNP in the HapMap CEU sample: red ($r^2 > 0.8$), orange ($0.5 < r^2 < 0.8$) or yellow ($0.2 < r^2 < 0.5$). White indicates either no LD with the lead SNP ($r^2 < 0.2$) or loci where such information was not available. Recombination rate estimates (HapMap Phase II) are given in light blue, RefSeq genes (NCBI) in green.

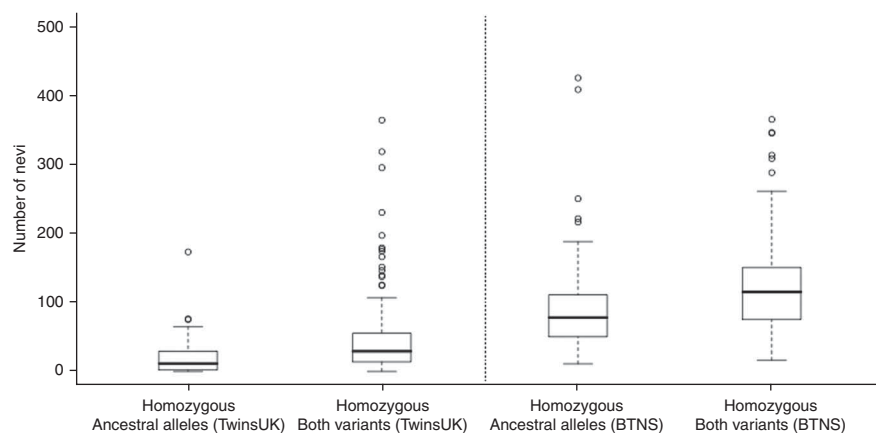


Figure 3 Box plot of number of nevi for subjects in the TwinsUK (left) and BNTS (right) samples homozygous for both *MTAP* SNP rs4636294 and *PLA2G6* SNP rs2284063 variants against the subjects homozygous for the protective alleles.

for 9p21 rs4636294, combined $P = 3.7 \times 10^{-8}$, OR = 1.21 (95% CI = 1.14–1.28) and for 22q13 rs2284063, combined $P = 1.5 \times 10^{-6}$, OR = 1.19 (95% CI = 1.12–1.27). Population attributable risk of the two loci reflected their different effect in the Australian and UK populations, ranging from 20.0% (Australia) to 9.0% (UK) for 9p21 and from 14.8 (Australia) to 29.3 (UK) for 22q13.

We used the Leeds sample to evaluate the combined risk for melanoma considering these two loci together (**Supplementary**

for melanoma status were then compared with those obtained when the number of nevi was included in the logistic regression (**Supplementary Table 4**). When the association between melanoma and SNP was adjusted for nevus count, the OR for melanoma risk was reduced toward 1.0 and the P value became nonsignificant, consistent with melanoma risk being mediated via nevus count. It is notable that the OR was midway between the unadjusted OR and 1.0, which might be expected as nevus count would not be a perfect measure of risk

Table 3). Individuals homozygous at the two loci showed nearly twice the risk for developing melanoma, in comparison of cases and controls in the Leeds cohort, as well as the controls from the WTCCC cohort (OR = 1.94; 95% CI = 1.19–3.16).

From the Leeds case-control study, 903 cases and 485 controls were assessed for their nevus phenotype through a standardized nevus count protocol, similar to that used in the TwinsUK sample. We used this subsample to explore whether the association between the identified loci and melanoma was direct or mediated by nevus count. Using these 485 controls, we verified that the effect of the nevus count loci was consistent with what was observed with the TwinsUK and BTNS samples. The results using this subsample were still consistent with those reported in **Table 2**. Odd ratios

Table 2 Summary results for the lead SNPs from the 9p21 and 22q13 loci for cutaneous melanoma

Marker	Chr.	Sample	Allele reference		Risk allele frequency		OR per risk allele (95% CI)	P	Combined OR per risk allele (95% CI)	Combined P	Heterogeneity P
			Base	Risk	Controls	Cases					
rs4636294	9	Q-MEGA Leeds2	G	A	0.48	0.55	1.31 (1.19, 1.46)	1.7×10^{-7}	1.21 (1.14, 1.28)	3.7×10^{-8}	4.0×10^{-3}
					0.48	0.51	1.11 (1.01, 1.22)	0.029			
rs2218220	9	Q-MEGA Leeds2	T	C	0.49	0.54	1.29 (1.17, 1.44)	5.7×10^{-7}	1.21 (1.14, 1.27)	4.8×10^{-8}	0.02
					0.48	0.51	1.13 (1.03, 1.25)	0.011			
rs7023329	9	Q-MEGA Leeds2	G	A	0.50	0.54	1.22 (1.11, 1.37)	1.2×10^{-4}	1.20 (1.13, 1.26)	2.7×10^{-7}	0.48
					0.49	0.53	1.17 (1.06, 1.29)	0.001			
rs10757257	9	Q-MEGA Leeds1	A	G	0.58	0.64	1.30 (1.17, 1.44)	7.7×10^{-7}	1.23 (1.15, 1.30)	3.4×10^{-8}	0.06
					0.58	0.62	1.16 (1.05, 1.28)	0.003			
rs751173	9	Q-MEGA Leeds1	A	G	0.46	0.51	1.28 (1.16, 1.42)	1.9×10^{-6}	1.17 (1.11, 1.24)	5.6×10^{-6}	3.0×10^{-3}
					0.46	0.48	1.07 (0.97, 1.18)	0.152			
rs1335510	9	Q-MEGA Leeds1	G	T	0.58	0.63	1.26 (1.14, 1.40)	1.1×10^{-5}	1.19 (1.12, 1.25)	1.1×10^{-7}	0.05
					0.57	0.60	1.12 (1.02, 1.24)	0.02			
rs1341866	9	Q-MEGA Leeds1	G	A	0.58	0.63	1.24 (1.13, 1.39)	2.6×10^{-5}	1.19 (1.12, 1.26)	6.9×10^{-7}	0.12
					0.58	0.60	1.13 (1.02, 1.25)	0.015			
rs10811629	9	Q-MEGA Leeds1	G	A	0.57	0.61	1.19 (1.08, 1.32)	1.3×10^{-4}	1.21 (1.12, 1.30)	1.6×10^{-5}	0.41
					0.56	0.62	1.28 (1.06, 1.55)	0.01			
rs2284063	22	Q-MEGA Leeds2	G	A	0.65	0.66	1.06 (0.95, 1.19)	0.275	1.19 (1.12, 1.27)	1.5×10^{-6}	1.0×10^{-3}
					0.63	0.69	1.30 (1.18, 1.44)	2.3×10^{-7}			
rs6001027	22	Q-MEGA Leeds2	G	A	0.65	0.66	1.06 (0.95, 1.19)	0.275	1.16 (1.09, 1.24)	6.8×10^{-5}	0.01
					0.64	0.69	1.25 (1.13, 1.38)	2.0×10^{-5}			
rs132985	22	Q-MEGA Leeds1	T	C	0.53	0.56	1.18 (1.05, 1.30)	2.7×10^{-3}	1.23 (1.15, 1.30)	2.6×10^{-7}	0.31
					0.52	0.58	1.26 (1.14, 1.38)	2.8×10^{-6}			
rs738322	22	Q-MEGA Leeds1	G	A	0.53	0.55	1.18 (1.05, 1.30)	2.8×10^{-3}	1.21 (1.14, 1.28)	3.0×10^{-7}	0.51
					0.52	0.57	1.23 (1.12, 1.36)	1.4×10^{-6}			

Joint odds ratios and confidence intervals were evaluated under a fixed-effects model for each SNP. Associated P values were estimated from the standard normal distribution. The Leeds1 and Leeds2 datasets are described in the Online Methods.

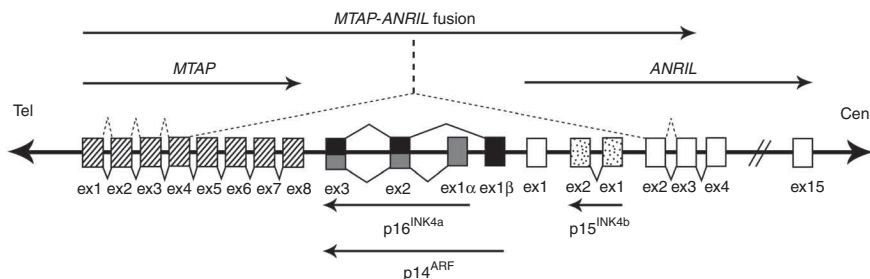


Figure 4 Relative positions of genes in the vicinity of the 9p21 association signal (not drawn to scale). Shaded boxes represent exons (ex), which are numbered, and solid lines connecting exons show the normal splicing patterns. Dotted lines between exons (above the figure) denote alternative splicing that generates an *MTAP-ANRIL* fusion transcript. Arrows indicate the direction of transcription. Gene product names are indicated. Tel, telomere; Cen, centromere.

(including error in its measurement). Indeed, nevus count is likely modified by other genetic influences as well as lifestyle factors. **Supplementary Figure 3** shows the distribution of nevus count in the Leeds case-control sample stratified by genotype for the highest associations on 9p21 and 22q13.

The location of the key SNPs in the 9p21 region is puzzling. They lie in the proximal promoter, 5' UTR and upstream from the coding region of the housekeeping gene *MTAP*, which encodes a key enzyme in the adenine and methionine salvage pathways. *MTAP* is often co-deleted^{9,10} with the adjacent melanoma susceptibility locus *CDKN2A*, which encodes p16INK4A and p14ARF (**Fig. 4**). The loss of *MTAP* in some tumors has generally been attributed to its proximity to *CDKN2A*, but some studies suggest *MTAP* might function as an independent tumor suppressor^{11,12}. A novel fusion transcript has been reported that comprises the first four exons of *MTAP* plus exons 2 and 3 of the noncoding RNA *ANRIL*, which has its first exon in the *CDKN2A* promoter¹³. This points to complex regulation of several genes within this region (**Fig. 4**). A possible hypothesis is that this scenario parallels that of common cancer susceptibility variants found earlier, where long-range regulatory effects on established oncogenes (*MYC* and *FAM84B*) have been postulated as a possible mechanism underlying multiple cancer associations mapping to a 1.2-Mb gene desert on 8q24 (ref. 14).

The *PLA2G6* gene on 22q13 belongs to the PLA2 superfamily of genes which encode esterases that cleave glycerophospholipids⁸. These enzymes are normally involved in the maintenance of membrane phospholipids, but have recently been shown to be associated with lung cancer susceptibility¹⁵ and to regulate cell growth¹⁶ and apoptosis¹⁷, as well as cell proliferation in human tumors such as insulinoma, colon cancer, ovarian cancer and prostate cancer^{18–20}.

We attempted to refine the location of chromosomes 9 and 22 association signals by imputing genotypes from the HapMap in the TwinsUK sample, and identified ten and four imputed SNPs on 9p21 and 22q13, respectively, that showed *P* values smaller than the strongest peak observed with the typed genotypes and might be helpful to identify some causal variants (**Supplementary Table 5**).

In the same issue of this journal, Bishop *et al.*²¹ report a genome-wide association study for melanoma conducted by the GenoMEL consortium. In their initial GWAS cohort, melanoma cases from Europe and Australia were carefully selected to enrich for those with family history and earlier onset of disease, in order to increase power to detect genetic associations to susceptibility. They report three loci associated with melanoma (on chromosome 16q24, 11q14, and 9p21), with replication of these loci in two independent cohorts. The associated locus on chromosome 9p21, with a top association at rs7023329 ($P = 4.03 \times 10^{-7}$), is also reported in our study. Bishop *et al.* also successfully attempted replication of our association on 22q13 at rs2284063 ($P = 2.40 \times 10^{-9}$). The GenoMEL study involves melanoma cases and controls

of diverse European ancestry, suggesting that the results presented in this paper generalize beyond English and Anglo-Celtic (Australian) ancestry.

Understanding how these genetic loci influence the development of nevi could provide important clues into the pathogenesis and treatment of melanoma, and may also prove of relevance for other cancers.

METHODS

Methods and any associated references are available in the online version of the paper at <http://www.nature.com/naturegenetics/>.

Database accession numbers. Fusion transcript reported in **Figure 4** - Ensemble ID: OTTHUMT00000051928.

Note: Supplementary information is available on the Nature Genetics website.

ACKNOWLEDGMENTS

The KCL authors acknowledge financial support from the Wellcome Trust, the Department of Health via the National Institute for Health Research (NIHR) comprehensive Biomedical Research Centre award to Guy's & St. Thomas' NHS Foundation Trust in partnership with King's College London, EC Framework 7 Health-2007-A ENGAGE project and the Chronic Disease Research Foundation (CDRF). T.D.S. is an NIHR senior investigator. We would also like to thank all the nurses and research assistants who collected the skin data, including U. Perks, G. Clement, D. Glass and E. Qweitin as well as the volunteer twins who gave their time.

The Leeds case-control study and genotyping was supported by Cancer Research UK Programme Award (C588/A4994), NIH (R01 CA83115) and EU FP6 to GenoMEL (LSHC-CT-2006-018702).

This study makes use of data generated by the Wellcome Trust Case Control Consortium. A full list of the investigators who contributed to the generation of the data are available from <http://www.wtccc.org.uk>. Funding for the project was provided by the Wellcome Trust under award 076113. The Leeds second control series were recruited through an award from the Department of Health in conjunction with I. dos Santos Silva and A. Swerdlow. We would like to thank J. Taylor for statistical analyses relating to the Leeds dataset. T.P. holds a Canada Research Chair and is supported by Genome Canada/Quebec and the CIHR.

The Australian Studies were supported by the National Institutes of Health/National Cancer Institute (CA88363), the National Health and Medical Research Council of Australia (NHMRC) and the Cancer Council Queensland. D.L.D., G.W.M. and N.K.H. are supported by the NHMRC Fellowships scheme. We thank D. Whiteman, A. Green, J. Aitken, A. Henders, M. Campbell, M. Stark, A. Baxter, M. de Nooyer, I. Gardner, D. Statham, B. Haddon, J. Palmer, L. Bardsley, D. Smyth and H. Beeby for their input into project management, sample processing, database and questionnaire development. We are grateful to all the participants of the BTNS and Q-MEGA as well as the research interviewers and examiners for these studies.

AUTHOR CONTRIBUTIONS

M.F., T.D.S., N.G.M., D.T.B. and N.K.H. designed the study. M.F., D.L.D., J.H.B. and A.C. analysed the data. V.B., N.G.M. and J.A.N.B. contributed to data collection and phenotype definitions. T.D.S., P.D., D.E.E., N.G.M., N.K.H. and J.A.N.B. obtained funding. Z.Z.Z., P.D., N.S. and G.W.M. contributed to genotyping. M.F., V.B. and T.D.S. wrote the first draft of the paper. All authors made important contributions to the final version of the paper.

Published online at <http://www.nature.com/naturegenetics/>.

Reprints and permissions information is available online at <http://npg.nature.com/reprintsandpermissions/>.

1. Karim-Kos, H.E. *et al.* Recent trends of cancer in Europe: a combined approach of incidence, survival and mortality for 17 cancer sites since the 1990s. *Eur. J. Cancer* **44**, 1345–1389 (2008).
2. Goldstein, A.M. *et al.* Features associated with germline CDKN2A mutations: a GenoMEL study of melanoma-prone families from three continents. *J. Med. Genet.* **44**, 99–106 (2007).
3. Palmieri, G. *et al.* Definition of the role of chromosome 9p21 in sporadic melanoma through genetic analysis of primary tumours and their metastases. The Melanoma Cooperative Group. *Br. J. Cancer* **83**, 1707–1714 (2000).
4. Fountain, J.W., Bale, S.J., Housman, D.E. & Dracopoli, N.C. Genetics of melanoma. *Cancer Surv.* **9**, 645–671 (1990).
5. Gandini, S. *et al.* Meta-analysis of risk factors for cutaneous melanoma: I. Common and atypical naevi. *Eur. J. Cancer* **41**, 28–44 (2005).
6. Falchi, M., Spector, T.D., Perks, U., Kato, B.S. & Bataille, V. Genome-wide search for nevus density shows linkage to two melanoma loci on chromosome 9 and identifies a new QTL on 5q31 in an adult twin cohort. *Hum. Mol. Genet.* **15**, 2975–2979 (2006).
7. Zhu, G. *et al.* A genome-wide scan for naevus count: linkage to CDKN2A and to other chromosome regions. *Eur. J. Hum. Genet.* **15**, 94–102 (2007).
8. Ma, Z., Wang, X., Nowatzke, W., Ramanadham, S. & Turk, J. Human pancreatic islets express mRNA species encoding two distinct catalytically active isoforms of group VI phospholipase A2 (iPLA2) that arise from an exon-skipping mechanism of alternative splicing of the transcript from the iPLA2 gene on chromosome 22q13.1. *J. Biol. Chem.* **274**, 9607–9616 (1999).
9. Garcia-Castellano, J.M. *et al.* Methylthioadenosine phosphorylase gene deletions are common in osteosarcoma. *Clin. Cancer Res.* **8**, 782–787 (2002).
10. Hori, Y. *et al.* The methylthioadenosine phosphorylase gene is frequently co-deleted with the p16INK4a gene in acute type adult T-cell leukemia. *Int. J. Cancer* **75**, 51–56 (1998).
11. Christopher, S.A., Diegelman, P., Porter, C.W. & Kruger, W.D. Methylthioadenosine phosphorylase, a gene frequently codeleted with p16(cdkn2a/ARF), acts as a tumor suppressor in a breast cancer cell line. *Cancer Res.* **62**, 6639–6644 (2002).
12. Behrmann, I. *et al.* Characterization of methylthioadenosine phosphorylase (MTAP) expression in malignant melanoma. *Am. J. Pathol.* **163**, 683–690 (2003).
13. Pasmant, E. *et al.* Characterization of a germ-line deletion, including the entire INK4/ARF locus, in a melanoma-neural system tumor family: identification of ANRIL, an antisense noncoding RNA whose expression coclusters with ARF. *Cancer Res.* **67**, 3963–3969 (2007).
14. Easton, D.F. & Eeles, R.A. Genome-wide association studies in cancer. *Hum. Mol. Genet.* **17**, R109–R115 (2008).
15. Hosgood, H.D. III *et al.* Pathway-based evaluation of 380 candidate genes and lung cancer susceptibility suggests the importance of the cell cycle pathway. *Carcinogenesis* **29**, 1938–1943 (2008).
16. Hooks, S.B. & Cummings, B.S. Role of Ca²⁺-independent phospholipase A2 in cell growth and signaling. *Biochem. Pharmacol.* **76**, 1059–1067 (2008).
17. Atsumi, G. *et al.* Fas-induced arachidonic acid release is mediated by Ca²⁺-independent phospholipase A2 but not cytosolic phospholipase A2, which undergoes proteolytic inactivation. *J. Biol. Chem.* **273**, 13870–13877 (1998).
18. Bao, S. *et al.* Effects of stable suppression of Group VIA phospholipase A2 expression on phospholipid content and composition, insulin secretion, and proliferation of INS-1 insulinoma cells. *J. Biol. Chem.* **281**, 187–198 (2006).
19. Song, Y. *et al.* Inhibition of calcium-independent phospholipase A2 suppresses proliferation and tumorigenicity of ovarian carcinoma cells. *Biochem. J.* **406**, 427–436 (2007).
20. Sun, B., Zhang, X., Talathi, S. & Cummings, B.S. Inhibition of Ca²⁺-independent phospholipase A2 decreases prostate cancer cell growth by p53-dependent and independent mechanisms. *J. Pharmacol. Exp. Ther.* **326**, 59–68 (2008).
21. Bishop, D.T. *et al.* Genome-wide association study identifies three loci associated with melanoma risk. *Nat. Genet.* advance online publication, doi:10.1038/ng.411 (5 July 2009).

ONLINE METHODS

Subjects and phenotypes. *TwinsUK.* The St. Thomas' UK adult twin registry (TwinsUK) cohort is unselected for any disease and is representative of the general UK population. Nevus counts were collected at St. Thomas Hospital in London and the study was approved by the St. Thomas' Hospital Research Ethics Committee. Written informed consent was obtained from every participant in the study. Examination was performed by trained research nurses following a standardized and reproducible nevus count protocol²². The total body nevus count (excluding the genital area, breasts and posterior scalp) was defined as the sum of all nevi >2 mm in diameter

Genotypic data were available for 1,524 subjects: 206 singletons, 832 dizygous (DZ) and 486 monozygous (MZ) twins. The median (first to third quartiles) age was 47 (37–54). The median (first to third) number of nevi was 21 (8–47).

BTNS and Q-MEGA. The Brisbane Twin Nevus Study (BTNS) is described in detail elsewhere²³ and comprised 4,107 adolescent twins, their siblings and parents (1,148 nuclear families) which have been recruited over 16 years into an ongoing study of genetic and environmental factors contributing to the development of nevi and other risk factors for skin cancer. The twins are recruited at age 12 years via schools around Brisbane, Australia, and followed up at age 14. All controls are screened to be unaffected by cutaneous malignant melanoma (CMM). The sample is overwhelmingly (>95%) of northern European origin (mainly Anglo-Celtic).

Total body nevus counting has been performed on the adolescent twins and singleton siblings on two occasions by a trained nurse (at ages 12 and 14). There were 1,979 twins and singleton siblings where nevus counts and genotypes were available. Of the parents of the twins, 1,470 had self-reported nevus score and genotypes available.

The Australian melanoma cases (Q-MEGA) were a stratified sample of 1,734 subjects including all cases of melanoma diagnosed in the state of Queensland in the period 1982–1990, as described in detail elsewhere²⁴. These individuals were originally studied between 1991 and 1994, but were recontacted and interviewed in 2002–2004.

Controls for the present analysis come from the parents of the BTNS families. An additional 658 individuals were genotyped and contributed as controls for the melanoma case-control analysis.

All cases and controls gave informed consent to participation in this study, and the study protocol was approved by appropriate institutional review boards.

Leeds Case-Control Study. In this study, 1,274 population-based incident melanoma cases were recruited in a geographically defined area of Yorkshire and the Northern region of the UK. Overall, the study had a 63% response rate. Subjects were diagnosed between September 2000 and December 2006. Cases were identified by clinicians, pathology registers and via the Northern and Yorkshire Cancer Registry and Information Service cancer registry to ensure overall ascertainment. For the first three years of the study all individuals with invasive melanoma were invited to participate; subsequently, only individuals with Breslow thickness of at least 0.75 mm were invited as cases. The case series was supplemented by 123 melanoma cases with a family history of melanoma (at most two affected relatives) who were known to be negative for a *CDKN2A* mutation. *CDKN2A* is the most common gene mutated in the germline of multicase melanoma families; such mutations are found in an estimated 21% of three-case families (data not shown).

Controls were ascertained by contacting general practitioners to identify eligible individuals. These controls were frequency matched with cases for age and sex from general practitioners who had also had cases as part of their patient register. Overall there was a 55% response rate among controls (496 subjects). Additional controls were supplemented by a population-based group of 574 women who following informed consent agreed to participate in a study involving recording their history of sun exposure and sun bathing including sun-bed usage and for whom various measures of skin aging were recorded. DNA samples were also provided.

All participating individuals gave written informed consent and regional ethical committee approval was obtained.

The first 903 cases and all the controls were asked about UV exposure and other aspects of lifestyle and provided detailed information on the history of residency. A trained nurse interviewed all cases and controls and formed a

standardized examination for skin nevi. Nevi were examined by body sector; nevi greater than 2 mm in diameter were counted.

Genotyping. The TwinsUK sample was genotyped using the Illumina Human-Hap 300k duo chip, and population substructure was detected and removed as described previously²⁵. The SNPs rs751173, rs4636294, rs2218220, rs1335510, rs1341866, rs10757257, rs7023329 and rs10811629 for *MTAP* and rs2284063, rs6001027, rs132985 and rs738322 for *PLA2G6* were genotyped both in the Australian BTNS and Q-MEGA samples in multiplex assays designed using the Sequenom MassARRAY Assay Design software (version 3.0). SNPs were typed using Sequenom iPLEX chemistry on a MALDI-TOF Compact Mass Spectrometer (Sequenom). The 2.5 μ l PCR reactions were performed in standard 384-well plates using 12.5 ng genomic DNA, 0.8 unit of Taq polymerase (HotStarTaq), 500 μ mol of each dNTP, 1.625 mM of MgCl₂, and 100 nmol of each PCRprimers (Bioneer). Standard PCR thermal cycling conditions and post-PCR extension reactions were carried out as described previously²⁶. The iPLEX reaction products were desalted by diluting samples with 15 μ l of water and adding 3 μ l of resin. The products were spotted on a SpectroChip (Sequenom), and data were processed and analyzed by MassARRAY TYPER 3.4 software (Sequenom). None of SNPs showed Hardy-Weinberg disequilibrium, and there were no effects of ethnicity in both the control and case groups.

For the Leeds sample, genotyping data from the Illumina HumanHap300 v2 array were available for the 123 familial cases and for 148 cases and 167 controls (Leeds1 sample). The SNPs rs4636294, rs2218220 and rs7023329 for *MTAP* and rs2284063 and rs6001027 for *PLA2G6* were genotyped in further 1,126 cases and 903 controls from the Leeds case-control study by Taqman with overall call rate greater than 97%. The Leeds1 sample plus the additional genotypes were part of the Leeds2 sample. None of SNPs showed deviation from Hardy-Weinberg equilibrium. Participants in the study were all of UK ancestry and principal component analysis (PCA) involving a subset of the samples included in this analysis found no evidence of population stratification. Both Leeds1 and Leeds2 case-control sample were supplemented with genotyping information from the Wellcome Trust Case Control Consortium (WTCCC) 1958 birth cohort control samples analyzed on the Illumina 550k array. Although data on 1,447 controls were provided by the WTCCC, these controls were reduced to 1,395 because of low call rate (call rate < 97%, $n = 51$) and for apparent non-European ancestry on the basis of PCA ($n = 1$).

Statistical analyses. Power analysis for the TwinsUK discovery sample was performed using the Genetic Power Calculator. Assuming that the tested SNP is a good proxy for an untyped functional variant, we have 70% to 90% power of identifying a QTN responsible for 1% to 2% of the total genetic variance showing a P value required for a significant result after Bonferroni correction ($P = 0.05/297,108 = 1.68 \times 10^{-7}$). The GWA was performed in the TwinsUK sample using a rapid variance component association test implemented in MERLIN software version 1.1.2 (ref. 27) on the natural log-transformed nevus counts and including age in the model. The method estimates the additive effect of each SNP by fitting a simple regression model to the trait while using a variance component approach to model the correlation between different observed phenotypes within pairs. A quantile-quantile (Q-Q) plot was generated using R and visually inspected to detect strong inflation of the test statistics. The genomic inflation factor was calculated as the ratio of the median of the empirically observed distribution of the test statistic to the expected median to quantify the excess false-positive rate²⁸ and was found to be 1.006, close to its expected value of 1.0. Simulation analysis by gene dropping was carried out with MERLIN in the TwinsUK sample to assess the empirical P values for the most significant SNPs on chromosome 9 (rs10757257) and 22 (rs132985). By using 1×10^{-6} , we obtained empirical P values of 2×10^{-6} for rs10757257 and 2×10^{-6} for rs132985. Replication in the BTNS was carried out using MENDEL²⁹, following cube-root transformation of the counts, and including age, age², year studied, body surface area, estimated cumulative UV exposure, ancestry, hair and skin color as covariates.

For the Australian data, cube-root transformation was significantly better than log transformation based on the Box-Cox analysis, and it closely matched the results of a negative binomial generalized linear mixed model analyses of the BTNS data, which also improved significantly in terms of model-fit diagnostics.



Empirical P values were evaluated through straight regression of log(nevus count) versus allele count using MENDEL for gene dropping of the most significant SNPs on chromosomes 9 and 22. By using 1×10^6 iterations, we obtained empirical P values of 1×10^{-6} for rs4636294 and 3.6×10^{-3} for rs2284063.

Combined P values from the BTNS and the TwinsUK studies were obtained using METAL. A z statistic is calculated for each marker using the P value and direction of the effect observed in each study. The combined P value is evaluated from the overall z statistic weighting each study proportionally to the square root of the number of subjects. The graphs in **Figure 2** were generated by adapting a freely available R script³⁰.

Given that some cases and all the controls have some relatives in the analysis, the case-control analysis in the Q-MEGA sample was carried out using MENDEL penetrance analysis option, and also the Sib-pair program. The latter was used to calculate simulation-based P values allowing for residual familial resemblance. We selected cases and controls where all four grandparents were reported to be of Northern European ancestry and we also performed quantitative-trait transmission disequilibrium test (TDT) analysis of nevus count, confirming that the observed associations were not due to population stratification.

For the Leeds analysis, case-control comparisons by SNP for melanoma were made by unconditional logistic regression. Analyses including nevus count involved age and sex adjusted nevus count. Tests of nevus count by SNP genotype were based on a test of linear trend of age and sex adjusted nevus count.

Meta-analysis for the melanoma case control association results was performed using the rmeta package of R. To refine the location of chromosomes 9 and 22 association signals and help the identification of some causal variants we imputed genotypes from the HapMap in the TwinsUK sample, as described

previously³¹. An additional 665 and 309 SNPs in the 9p21 and 22q13 regions respectively were tested for association with nevus count.

URLs. Genetic Power Calculator, <http://pngu.mgh.harvard.edu/~purcell/gpc/>; R, <http://www.r-project.org/>; METAL, <http://www.sph.umich.edu/csg/abecasis/metal/>; GenoMEL, <http://www.genomel.org/>.

22. Bataille, V., Snieder, H., MacGregor, A.J., Sasiemi, P. & Spector, D.T. Genetics of risk factors for melanoma: an adult twin study of nevi and freckles. *J. Natl. Cancer Inst.* **92**, 457–463 (2000).
23. Sturm, R.A. *et al.* A single SNP in an evolutionary conserved region within intron 86 of the HERC2 gene determines human blue-brown eye color. *Am. J. Hum. Genet.* **82**, 424–431 (2008).
24. Baxter, A.J. *et al.* The Queensland Study of Melanoma: environmental and genetic associations (Q-MEGA); study design, baseline characteristics, and repeatability of phenotype and sun exposure measures. *Twin Res. Hum. Genet.* **11**, 183–196 (2008).
25. Richards, J.B. *et al.* Bone mineral density, osteoporosis, and osteoporotic fractures: a genome-wide association study. *Lancet* **371**, 1505–1512 (2008).
26. Zhao, Z.Z. *et al.* Genetic variation in tumour necrosis factor and lymphotoxin is not associated with endometriosis in an Australian sample. *Hum. Reprod.* **22**, 2389–2397 (2007).
27. Chen, W.M. & Abecasis, G.R. Family-based association tests for genomewide association scans. *Am. J. Hum. Genet.* **81**, 913–926 (2007).
28. Devlin, B., Roeder, K. & Wasserman, L. Genomic control, a new approach to genetic-based association studies. *Theor. Popul. Biol.* **60**, 155–166 (2001).
29. Lange, K., Weeks, D. & Boehnke, M. Programs for Pedigree Analysis: MENDEL, FISHER, and dGENE. *Genet. Epidemiol.* **5**, 471–472 (1988).
30. Saxena, R. *et al.* Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. *Science* **316**, 1331–1336 (2007).
31. Prokopenko, I. *et al.* Variants in MTNR1B influence fasting glucose levels. *Nat. Genet.* **41**, 77–81 (2009).