

Pathway-Based Analysis of a Melanoma Genome-Wide Association Study: Analysis of Genes Related to Tumour-Immunosuppression

Nils Schoof^{1,2}, Mark M. Iles¹, D. Timothy Bishop¹, Julia A. Newton-Bishop¹, Jennifer H. Barrett^{1*}, GenoMEL consortium^{3¶}

1 Section of Epidemiology and Biostatistics, Leeds Institute of Molecular Medicine, University of Leeds, Leeds, United Kingdom, **2** Department of Medical Epidemiology and Biostatistics, Karolinska Institute, Stockholm, Sweden, **3** The Melanoma Genetics Consortium (GenoMEL)

Abstract

Systemic immunosuppression is a risk factor for melanoma, and sunburn-induced immunosuppression is thought to be causal. Genes in immunosuppression pathways are therefore candidate melanoma-susceptibility genes. If variants within these genes individually have a small effect on disease risk, the association may be undetected in genome-wide association (GWA) studies due to low power to reach a high significance level. Pathway-based approaches have been suggested as a method of incorporating *a priori* knowledge into the analysis of GWA studies. In this study, the association of 1113 single nucleotide polymorphisms (SNPs) in 43 genes (39 genomic regions) related to immunosuppression have been analysed using a gene-set approach in 1539 melanoma cases and 3917 controls from the GenoMEL consortium GWA study. The association between melanoma susceptibility and the whole set of tumour-immunosuppression genes, and also predefined functional subgroups of genes, was considered. The analysis was based on a measure formed by summing the evidence from the most significant SNP in each gene, and significance was evaluated empirically by case-control label permutation. An association was found between melanoma and the complete set of genes ($p_{\text{emp}} = 0.002$), as well as the subgroups related to the generation of tolerogenic dendritic cells ($p_{\text{emp}} = 0.006$) and secretion of suppressive factors ($p_{\text{emp}} = 0.0004$), thus providing preliminary evidence of involvement of tumour-immunosuppression gene polymorphisms in melanoma susceptibility. The analysis was repeated on a second phase of the GenoMEL study, which showed no evidence of an association. As one of the first attempts to replicate a pathway-level association, our results suggest that low power and heterogeneity may present challenges.

Citation: Schoof N, Iles MM, Bishop DT, Newton-Bishop JA, Barrett JH, et al. (2011) Pathway-Based Analysis of a Melanoma Genome-Wide Association Study: Analysis of Genes Related to Tumour-Immunosuppression. PLoS ONE 6(12): e29451. doi:10.1371/journal.pone.0029451

Editor: Kerby Shedden, University of Michigan, United States of America

Received: April 1, 2011; **Accepted:** November 28, 2011; **Published:** December 27, 2011

Copyright: © 2011 Schoof et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: The funding for this project comes from GenoMEL EU FP6 LSHC-CT-2006-018702 and Cancer Research UK Programme Award (C588/A4994 and C588/A10589). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

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

* E-mail: J.H.Barrett@leeds.ac.uk

¶ GenoMEL consortium list can be found in the Acknowledgments section.

Introduction

The incidence of melanoma is increased in chronically immunosuppressed individuals, such as recipients of transplanted organs, indicating that the immune system restricts the outgrowth of melanoma cells [1,2]. Anti-melanoma immune responses seem to be possible, but certain mechanisms probably at the tumour site circumvent these and give rise to tumour development [3]. Furthermore a potent risk factor for melanoma is sunburn [4], and seminal animal studies established that ultraviolet-induced local and systemic immunosuppression is important in the pathogenesis of melanoma. The hypothesis is that intense sun exposure induces both genetic changes, resulting in tumour antigenicity, and an inability of the immune system to detect those changes [5].

Within the concept of tumour immunosurveillance [6], transformed cells are recognized by antigen-presenting cells (APCs) (essentially dendritic cells (DCs)), and the latter differentiate into activated states. The activated APCs further interact with tumour-specific T helper lymphocytes and induce their activation,

which in turn leads to activation of tumour-specific cytotoxic T lymphocytes (cTLs). These cTLs recognise the nascent tumour cells and induce their elimination. Many molecular mechanisms are known to influence immunological capacity. The DCs are known to exist in a state that induces immune tolerance and in an activated state, which induces immunity [7]. It has been shown that melanoma appears to induce tolerogenic DCs (tDCs) capable of inducing immunosuppression [8]. Two major mechanisms are known to prevent T lymphocyte activation and resulting immune responses. Firstly, T lymphocytes can differentiate into a state of anergy indicating their functional inactivation [9]. The analysis of the microenvironment around melanoma cells showed the presence of anergic T-cells, and these might also contribute to the lack of anti-tumoural immune responses [10,11]. Secondly, regulatory T lymphocytes (Treg) have strong immunosuppressive properties through multiple modes of action [12]. Treg cells have also been found in melanoma lesions and could induce immunotolerance [13,14]. The interaction between tumour cells, APCs and T lymphocytes and their respective effects are strongly

dependent on molecules on the surface of each cell. The significance of one of these, the cytotoxic T-lymphocyte-associated protein 4 (*CTLA4*), is indicated by the recent encouraging clinical trials using antibodies directed to this immunosuppressive costimulatory molecule in patients with stage IV melanoma [15]. One major effector function of these immunosuppressive cell types is the secretion of factors with immune regulatory functions. However, the tumour cells are also capable of secretion of these factors and may thereby recruit (e.g. by chemokine (C-C motif) ligand 17/22 (*CCL17/22*)) or lead to the differentiation of immunosuppressive cell types (e.g. by indolamine-2,3-dioxygenase 1 (*IDO1*) or Interleukin 10 (*IL10*)) [16].

Several candidate gene studies have been reported focussing on variants within genes related to these immunosuppressive mechanisms [17–20]. Most of these studies analysed only a limited number of single nucleotide polymorphisms (SNPs) or had relatively small sample sizes, and some of the results are inconsistent.

In a genome-wide association (GWA) study of melanoma carried out by the GenoMEL consortium, association was confirmed between disease susceptibility and variants related to melanocortin-1 receptor (*MCLR*) and tyrosinase (*TYR*), and a new locus at chromosome 9p21 was identified [21]. In GWA study analyses, usually each individual SNP is tested for association with the disease, and only loci approaching “genome-wide” statistical significance (e.g. $p < 5 \times 10^{-7}$) are followed up. GWA studies are powerful at identifying risk variants, which are common in a population and have low to moderate penetrance. However, other loci in the large GWA data sets are also likely to be associated with disease risk but are indistinguishable from false positive results using this approach. Thus, candidate gene approaches and GWA studies have contributed to the understanding of genetic disease risk, but the latter are underpowered to detect weak associations with susceptibility to disease at the genome-wide significance level.

To overcome the limitations of these approaches, the analysis of functional gene sets, so-called pathway-based analysis, has recently been proposed [22–25]. The large data sets from GWA studies can be re-analysed incorporating *a priori* knowledge into the analysis in an attempt to identify new risk factors. The idea behind pathway-based GWA study analysis is that SNPs in a group of genes with a shared biological function may show significant association at the pathway level, even though no individual SNP shows association at a stringent level of statistical significance. Thus further information about disease aetiology may be obtained using the existing data from the GWA study. Recently, pathway-based approaches have been applied to GWA studies of several complex diseases [26–31]. Most analyses have taken an agnostic approach and included a comprehensive pathway search, using databases like Gene Ontology or the Kyoto Encyclopedia of Genes and Genomes. In this study, a candidate pathway analysis is applied to data from the GenoMEL melanoma GWA study [21]. Instead of an analysis of many pathways, the pathway of tumour-immunosuppression was selected, and a comprehensive analysis of SNPs in this pathway in relation to melanoma susceptibility was conducted.

Materials and Methods

Study design and subjects

This study reports a further statistical analysis of the first phase of the melanoma GWA study of the GenoMEL consortium, full details of which have been described elsewhere [21]. Briefly, participating GenoMEL groups (Barcelona, Brisbane, Emilia-Romagna, Genoa, Leiden, Leeds, Lund, Paris, Stockholm and Sydney) contributed 1650 melanoma cases with either a family

history of melanoma (without *CDKN2A* mutations), early disease onset (age < 40 years) or multiple primary sites to enrich for cases with greater genetic predisposition. Controls were provided by the same GenoMEL groups from the same populations, and an additional 2938 controls were contributed by the Centre National de Genotypage (CNG, France) and the Wellcome Trust Case Control Consortium (WTCCC, UK). The anonymised data are stored on a secure server, and personal information is held only by the contributing centre. Each participating group holds local ethical approval for the GWA analysis and written informed consent from the participants [21].

Genotyping and sample exclusion

Genotyping was performed through ServiceXS (Netherlands) using the Illumina HumanHap300 Bead-Chip version 2 duo array and by CNG in Paris using the Illumina HumanCNV370k array. Additional French and WTCCC controls were genotyped on the Illumina HumanHap300 BeadChip version 2 duo array. In total, 1650 cases and 4336 controls were genotyped. Samples were excluded if (i) the overall call rate was less than 97%, (ii) there was evidence of non-European origin from principal components analysis (PCA), (iii) sex as inferred from genotyping did not match reported sex, or (iv) there was evidence of first-degree relationship or genetic identity with another sample (for detailed information see [21]). This quality control led to the exclusion of 111 cases and 419 controls, mainly due to a call rate (<97%) (predominantly from the group of additional French controls).

Selection of genes and SNPs

In this study 43 genes, associated with the suppression of immune responses, were selected, based on an extensive literature review and blind to the results of the GWA study (Table 1). The genes were further divided into subgroups related to suppression by regulatory T-cells (*Treg*), the induction of T-lymphocyte anergy (*Anergy*), regulation by costimulatory receptors (*Costim.*), regulation by dysfunctional, tolerogenic dendritic cells (*tDC*), and the secretion of suppressive factors (*Secreted*). Some genes can be categorised into more than one subgroup, as indicated in Table 1. The chromosomal location of the genes was retrieved from the HapMap database (NCBI build 36) and 100 kilobase flanking regions were included. Overlapping genes were merged into one, leading to 39 genomic regions to be analysed. From within these regions, 1178 SNPs genotyped in this study were obtained using the dbSNP database (build 126).

SNP quality control

Quality control was based on the minor allele frequency, genotype call rate, exact Hardy-Weinberg equilibrium (HWE) test and differences in allele frequencies between six geographical regions of the participating GenoMEL centres (grouped as Sweden, Australia, Italy, United Kingdom/Netherlands, France and Spain) (based on a χ^2 test with 5 degrees of freedom (d.f.)). 60 SNPs with a call rate below 97% were excluded from the analysis. Furthermore, 5 SNPs with an exact HWE p -value < 10^{-5} and no regional differences in the allele frequencies (χ^2 test p -value > 0.001) were excluded.

Pathway-based analysis

The association between susceptibility to melanoma and the immunosuppression gene set (and the respective subgroups (Table 1)) was analysed using an approach in which each gene (or genomic region) in the pathway is represented by the maximally associated SNP within the gene [25,32]. First, logistic

Table 1. Selected genes included into the analysis and division into subgroups.

GeneSymbol	EntrezGeneID	Chromo-some	Treg	Anergy	Costim.	tDC	Secreted
BTLA	151888	3			X		
BTNL2	56244	6			X		
CBLB	868	3		X			
CCL17/CCL22	6361/6367	16	X				X
CD160	11126	1			X		
CD274/PDCD1LG2	29126/80380	9			X		
CD28/ICOS/CTLA4	940/29851/1493	2	X		X		
CD40	958	20			X		
CD40LG	959	X			X		
CD80	941	3			X		
CD86	942	3			X		
DGKA	1606	12		X			
FOXP3	50943	X	X				
ICOSLG	23308	21			X		
IDO1	3620	8	X			X	X
IL10	3586	1	X			X	X
IL10RA	3587	11	X			X	
IL10RB	3588	21	X			X	
IL12A	3592	3					X
IL12B	3593	5					X
IL12RB1	3594	19	X				
IL17A	3605	6	X				X
IL17RA	23765	22	X				
IL17RB	55540	3	X				
ITCH	83737	20		X			
LGALS1	3956	22					X
LGALS3	3958	14					X
LILRB2	10288	19				X	
LILRB4	11006	19				X	
PDCD1	5133	2			X		
RNF128	79589	X		X			
TGFB1	7040	19	X			X	X
TGFB2	7042	1	X			X	X
TGFBR1	7046	9	X			X	
TGFBR2	7048	3	X			X	
TGFBR3	7049	1	X			X	
TNFRSF18	8784	1	X				
TREML2	79865	6			X		
VDR	7421	12	X			X	

doi:10.1371/journal.pone.0029451.t001

regression was performed for each SNP, assuming an additive genetic model and with the GenoMEL regional group (defined above) included as a covariate, using the PLINK software package [33]. Then, 5000 case-control label permuted data sets were analysed in the same way. These data sets were created by permuting case-control status within clusters (formed by the GenoMEL regional groups) in order to retain the original structure of the GWA study. All PLINK result files (based on observed and permuted data) were further analysed using the R software package [34]. The maximal z-value (i.e. the absolute

value of the coefficient for the per-allele SNP effect, divided by its standard error) from the logistic regression analysis was assigned to each of the k genes, for the observed data set and each permuted data set. To evaluate the statistical significance of the gene sets, the SUMSTAT ($\sum_k |z_i|$) and SUMSQ ($\sum_k z_i^2$) statistics were calculated for the observed and permuted gene sets [32]. An empirical p-value for the association of a gene set with melanoma susceptibility is calculated by the number of times the permuted gene statistic exceeds the original test statistic, divided by the number of permutations. This method thus provides a test of the

null hypothesis that none of the genes is associated with melanoma risk.

Random gene sets were tested for association with melanoma using the same approach (Table 5 in Information S1). These gene sets were randomly selected from a list containing 18410 genes (Refseq sequences with status “mRNA” from HG18 downloaded from the HapMap database). For each random gene set, 43 genes were sampled, but several sets included genes containing no SNPs in the GenoMEL GWA data set (mainly from the X chromosome), reducing the number of genes to between 35 and 42 genes per set. With these 100 random gene sets the pathway analysis was conducted as described before.

Results

Overall, the data for 1178 genotyped SNPs in the 39 selected genomic regions in 1539 melanoma cases and 3917 controls were available from the GenoMEL GWA study. 1113 SNPs remained after quality control and were used for the pathway-based analyses. The smallest nominal p-value from the logistic regression analysis was found for SNP rs873061 in the region of the lectin galactoside-binding soluble 3 (*LGALS3*) gene ($p_{\text{logreg}} = 0.00033$, odds ratio 0.84, 95% CI: 0.76–0.92). Thus no SNP showed association with risk of melanoma at a genome-wide significance level or after correcting for the number of SNPs in this study (Bonferroni correction with significance level of 0.05 corrected for 1113 tests).

Analysing the data using the pathway analysis based on the most significant SNP in each gene showed evidence of association of the complete set of immunosuppressive genes with the risk of melanoma (SUMSTAT $p_{\text{emp}} = 0.002$ from 5000 permutations, Table 2). Two subgroups of genes were primarily responsible for this result; the subgroup of secreted factors showed the strongest association (SUMSTAT $p_{\text{emp}} = 0.0004$), followed by the subgroup of genes associated with tDCs (SUMSTAT $p_{\text{emp}} = 0.006$). The results were very similar using the SUMSQ statistic instead of SUMSTAT for the pathway statistic. The three other groups (Treg, Anergy and Costim.) showed no significant results at the 5% level using either statistic.

Several further analyses were conducted to test the validity of these results. First, these results remained stable when including the first three principal components (established to account for population stratification [21]) as covariates in the logistic regression (Table 1 in Information S1, $p_{\text{emp}} = 0.006$ from 1000 permutations). In particular, the subgroup of secreted immunosuppressive factors remained most significantly associated

($p_{\text{emp}} = 0.002$). Secondly, 100 random gene sets were tested, applying the same methodology (Table 5 in Information S1). Only 10 of the 100 random gene sets showed a nominally significant result ($p < 0.05$) whichever test statistic was used (SUMSTAT or SUMSQ). Thirdly, the observed data set was replaced by a permuted data set and the complete procedure was repeated to test for any flaws in the programmed R algorithm (Table 2 in Information S1). In this analysis, there was no evidence of overall association, and only the subgroup of genes related to anergy reached nominally significant results (SUMSTAT $p_{\text{emp}} = 0.036$ from 1000 permutations).

Table 3 shows the detailed results for each gene in the subgroup of secreted factors, which showed the strongest association. In 7 of the 10 genes, SNPs with a p-value below 0.05 were found. Three genes contained only one SNP with a p-value below 0.05. In four genes there were two or more SNPs with a p-value below 0.05, and these genes also contained the two most significant SNPs, found in the *LGALS3* and transforming growth factor beta 2 (*TGFB2*) genes. We tried to replicate the results of this study in the second phase of the GenoMEL GWA study [35], consisting of 1450 melanoma cases and 4047 controls (from Italy, France, Scandinavia, Spain, UK, Netherlands, Poland and Israel), but no evidence of association with the pathway was seen (Table 4).

Discussion

Using pathway-based analysis, preliminary evidence for an association between genes involved in immunosuppression and melanoma risk is provided by this study. The pathway itself and the genes to include within it were chosen completely blind to the results of the GenoMEL study, yet the observed level of evidence for association was only seen 10 times in 5,000 permutations. The approach applied here uses the most significant SNP within each gene to form the pathway statistic, as suggested by Wang and colleagues [25]. Instead of using a weighted Kolmogorov-Smirnov-like running-sum statistic (used in the original gene set enrichment analysis for genome-wide gene expression profiling) [36], SUMSTAT and SUMSQ statistics were used as suggested by Tintle and colleagues [32]. Both statistics show comparable results, although the SUMSQ statistic tends to have larger p-values in our analyses.

Table 2. Pathway analysis for all genes and the gene subgroups.

Set	SUMSTAT	SUMSQ
All genes	<i>0.0020</i>	<i>0.0032</i>
Anergy	0.1378	0.1548
Costim.	0.1022	0.1556
Treg	0.0874	0.0812
Secreted	<i>0.0004</i>	<i>0.0004</i>
tDC	<i>0.0060</i>	<i>0.0082</i>

Empirical p-values established by 5000 in-cluster (GenoMEL regional group) label permutations are shown for the pathway statistics SUMSTAT and SUMSQ. Nominally significant results are shown in italics.
doi:10.1371/journal.pone.0029451.t002

Table 3. Detailed results for the subgroup of secreted immunosuppressive factors.

GeneSymbol	# SNPs	# SNPs $p \leq 0.05$	min. p-value
IDO	26	6	0.00128
IL10	21	1	0.03905
TGFB1	20	1	0.00141
TGFB2	28	2	0.00083
CCL17/CCL22	31	0	0.10590
IL12A	27	0	0.05448
IL12B	20	0	0.07014
IL17A	32	1	0.00631
LGALS1	24	4	0.00565
LGALS3	11	8	0.00033

Number of SNPs in the gene region, number of SNPs with a p-value below 0.05 and the minimal p-value of the SNPs (logistic regression analysis) in the gene region are shown.

doi:10.1371/journal.pone.0029451.t003

Table 4. Attempt to replicate the results of the pathway analysis in the second phase of data from the GenoMEL GWA study.

Set	SUMSTAT	SUMSQ
All genes	0.167	0.280
Anergy	0.001	0.002
Costim.	0.291	0.411
Treg	0.690	0.794
Secreted	0.738	0.783
tDC	0.701	0.761

1450 melanoma cases and 4047 controls were included in the analysis. Empirical p-values established by 1000 in-cluster (GenoMEL regional groups: Israel, Italy, France, Poland, Scandinavia, Spain, UK/Netherlands) label permutations are shown for the pathway statistics SUMSTAT and SUMSQ. doi:10.1371/journal.pone.0029451.t004

For genome-wide expression analysis, it was found that these statistics, together with label permutations, might lead to many significant gene sets [37]. This may be because very many genes are expressed differentially between the groups being compared, so that many pathway-based gene sets will include at least one differentially-expressed gene, even though the pathway itself is not important. This is less likely in a GWA context, but false positive results may arise due to population stratification. Several tests have been performed here to prevent or rule out spurious associations. Within-cluster permutations were used to preserve the geographical structure of the GWA study in the permuted data sets. In addition to the geographical region of the respective GenoMEL groups, an adjustment for the first three principal components (PCs) was performed in the logistic regression analysis to further reduce the potential effect of population stratification (Table 1 in Information S1). Although slightly less significant empirical p-values were achieved by this method, the results remained stable. Furthermore, 100 random gene sets were analysed by the same method (Table 5 in Information S1). Only 10 of 100 random gene set showed a nominally significant result suggest a type 1 error rate of 0.10 (95% CI 0.05, 0.18). A slightly increased type 1 error rate might be explained by the fact that some of the 100 random gene sets are likely to include a gene associated with susceptibility for melanoma. Assuming 0.12% of genes to be associated with melanoma risk (approximately 24 genes of 20,000 genes in the genome), the probability that at least one of these is included in a random gene set (including 43 genes) is 5%, leading to some gene sets showing inflated evidence of association. For instance, random gene set number 81 (Table 5 in Information S1) contains the gene CDK10, which is found in a region of genome-wide significance in melanoma association studies [21].

As pointed out by Wang et al. [25], the use of the most significant SNP within each gene is only one possibility. This approach could be strongly influenced by a few highly significant SNPs (occurring by chance in the GWA study) being present within the gene set. In a recent study, the second most strongly associated SNP in each gene was used to reduce the chance of this [38]. An alternative approach to pathway analysis, comparable to the approach of Holmans et al. [27], is provided by the gene set test within PLINK. Instead of using one SNP per gene, this approach uses a predefined p-value threshold for the inclusion of SNPs from the initial association analysis into the pathway-based analysis. It further removes SNPs in linkage disequilibrium (LD)

based on a predefined criterion. The mean p-value of the selected SNPs is then used as the summary pathway measure, and significance is assessed by case-control label permutation as above. As a secondary analysis, we also applied this method to the data, but found little evidence of association (Table 3 and 4 in Information S1). The results might be expected to be sensitive to selection of both parameters (SNP cut-off p-value and LD criterion), although no differences are found in this study for different p-value thresholds (Table 3 in Information S1). Similarly no difference in the conclusions for the complete gene set was detected by changing the R^2 filter criterion to 0.8 (data not shown). This approach takes all predefined significant SNPs into account but makes no use of the gene level, which is used in the main analysis presented here. The lack of association may be the result of the introduction of too much noise from genes with large numbers of SNPs. Recently, a comparison of different pathway-based analysis methods including the PLINK gene set test as well as the approaches of Holmans et al. [27] and Wang et al., [25] was performed [39]. The simulation studies suggest that the PLINK gene set test has higher power than the two other approaches. However, the methodologies for pathway-based analysis applied to SNP data sets are still under development, and further simulation studies with a broad range of scenarios have to be conducted before firm conclusions can be drawn.

If validated, the results from this study give an interesting insight into the biology of melanoma susceptibility. The subgroup of secreted factors contains several molecules which are crucial in the crosstalk between tumour cells and the host immune system. The *IDO1* gene encodes an enzyme crucial for the tryptophan catabolism and promoting the arrest of T lymphocyte proliferation by tryptophan deprivation [40,41]. Non-synonymous coding gene variants in the *IDO1* gene have been associated with an altered gene expression, [42] and it would be of interest to analyse the LD between these coding variants and the SNPs showing association in this analysis. The two Galectins (*LGALS1* and *LGALS3*) are associated with the survival of effector T lymphocytes and may change the balance of the immune response towards an anti-inflammatory cytokine profile [43]. It has been shown that *LGALS3* is regulated by the microphthalmia-associated transcription factor (MITF), which has a pivotal role in melanocyte development and melanoma [44]. Moreover, the serum level of Galectin-3 has been significantly associated with the prognosis of the melanoma patients [45]. The role of these galectins and respective gene variations in melanoma susceptibility has to be further evaluated. *TGFB1* and *TGFB2* are key immunosuppressive cytokines [46]. Currently, only variations in the *TGFB1* gene have been analysed with regards to the risk of melanoma with conflicting results [18,19]. This study provides some evidence that variants in *TGFB2* might also be associated with melanoma susceptibility.

Further analyses are needed to confirm the results of this study. No evidence of association with the overall pathway was seen when applied to the second phase of the GenoMEL study (Table 4) [35]. There are several possible reasons for this lack of replication. First, it is likely that the immunosuppression pathway is not among the strongest predictors of melanoma risk (which are related to nevus development and skin pigmentation) and that the gene set selected includes some genes not associated with susceptibility, resulting in low power to detect association even at the pathway level. Secondly, it could be that initial result is a statistical false positive. Our analysis could be likened to examining a candidate SNP for association with disease. In each case, although we have only examined one hypothesis, motivated by biological understanding, p-values of this magnitude (0.001) can arise by chance.

Thirdly, the lack of replication could be due to heterogeneity between the two phases of the GenoMEL study, which differed slightly both in the distribution of geographical region of origin and of case ascertainment criteria (see Information S1 for further description). Looking at the results in more detail, the three SNPs showing the highest association in the first analysis showed no evidence of association in the second, but this was not readily explained by differences in geography or reason for ascertainment, although other sources of heterogeneity (e.g. site of melanoma) may exist. This is one of the first pathway-based analyses in which an attempt to replicate the results at the pathway level has been reported. The lack of replication is disappointing but may presage more general difficulties in replicating results for complex hypotheses, that may be susceptible to the effects of heterogeneity and low power. We hope others using these methods will be encouraged to attempt to replicate their own results.

In conclusion, the results presented here suggest that variants in the gene set of immuno-suppressive factors, and especially in the subgroup of secreted factors, may be associated with the susceptibility to melanoma. Although the methodology has to be further evaluated and developed, and we have so far not replicated these results, this study underlines the potential of pathway-based methods for complementary analyses of GWA data sets.

Supporting Information

Information S1 Contains additional detail on comparisons between Phase 1 and Phase 2 of the GenoMEL GWA study, supplementary methods and further tables of results.

(DOCX)

Acknowledgments

This study makes use of data generated by the Wellcome Trust Case Control Consortium (Wellcome Trust award 076113). A full list of the investigators who contributed to the generation of the data is available from www.wtccc.org.uk. Most of all we thank the individuals with melanoma and their families for their cooperation.

GenoMEL Collaboration

Australian Melanoma Family Study: Graham J. Mann, John L. Hopper, Joanne F. Aitken, Bruce K. Armstrong, Graham G. Giles, Richard F. Kefford, Anne Cust, Mark Jenkins.

Barcelona: The participants of GenoMEL in Barcelona: Paula Aguilera, Celia Badenas, Cristina Carrera, Francisco Cuellar, Daniel Gabriel, Estefania Martinez, Melinda Gonzalez, Pablo Iglesias, Josep Malvehy, Rosa Marti-Laborda, Montse Mila, Zighe Ogbah, Joan-Anton Puig Butille, Susana Puig and Other members of the Melanoma Unit: Lúcia Alós, Ana Arance, Pedro Arguís, Antonio Campo, Teresa Castel, Carlos Conill, Jose Palou, Ramon Rull, Marcelo Sánchez, Sergi Vidal-Sicart, Antonio Vilalta, Ramon Vilella.

Brisbane: The Queensland study of Melanoma: Environmental and Genetic Associations (Q-MEGA) Principal Investigators are: Nicholas G. Martin, Grant W. Montgomery, David Duffy, David Whiteman, Stuart MacGregor, Nicholas K. Hayward. **The Australian Cancer Study (ACS) Principal Investigators are:** David Whiteman, Penny Webb, Adele Green, Peter Parsons, David Purdie, Nicholas Hayward.

Emilia-Romagna: Maria Teresa Landi, Donato Calista, Giorgio Landi, Paola Minghetti, Fabio Arcangeli, Pier Alberto Bertazzi.

Genoa: Department of Oncology, Biology and Genetics, University of Genoa: Giovanna Bianchi-Scarra, Paola Ghiorzo, Lorenza Pastorino, William Bruno, Linda Battistuzzi, Sara Gargiulo, Sabina Nasti, Sara

Gliori, Paola Origone; Medical Oncology Unit, National Institute for Cancer Research: Paola Queirolo.

Glasgow: Rona Mackie, Julie Lang.

Leeds: Julia A Newton Bishop, Paul Affleck, Jennifer H Barrett, D Timothy Bishop, Jane Harrison, Mark M Iles, Juliette Randerson-Moor, Mark Harland, John C Taylor, Linda Whittaker, Kairen Kukalich, Susan Leake, Birute Karpavicius, Sue Haynes, Tricia Mack, May Chan, Yvonne Taylor, John Davies, Paul King.

Leiden: Department of Dermatology, Leiden University Medical Centre: Nelleke A Gruis, Frans A van Nieuwpoort, Coby Out, Clasine van der Drift, Wilma Bergman, Nicole Kukutsch, Jan Nico Bouwes Bavinck. Department of Clinical Genetics, Centre of Human and Clinical Genetics, Leiden University Medical Centre: Bert Bakker, Nienke van der Stoep, Jeanet ter Huurne. Department of Dermatology, HAGA Hospital, The Hague: Han van der Rhee. Department of Dermatology, Reinier de Graaf Groep, Delft: Marcel Bekkenk. Department of Dermatology, Sint Franciscus Gasthuis, Rotterdam: Dyon Snels, Marinus van Praag. Department of Dermatology, Ghent University Hospital, Ghent, Belgium: Lieve Brochez and colleagues. Department of Dermatology, St. Radboud University Medical Centre, Nijmegen: Rianne Gerritsen and colleagues. Department of Dermatology, Rijnland Hospital, Leiderdorp: Marijanne Crijns and colleagues. Dutch Patient Organization, Stichting Melanoom, Purmerend. The Netherlands Foundation for the Detection of Hereditary Tumors, Leiden: Hans Vasen.

Lund: Lund Melanoma Study Group: Håkan Olsson, Christian Ingvar, Göran Jönsson, Åke Borg, Anna Måsbäck, Lotta Lundgren, Katja Baeckenhorn, Kari Nielsen, Anita Schmidt Casslén.

Norway: Oslo University Hospital: Per Helsing, Per Arne Andresen, Helge Rootwelt. **University of Bergen:** Lars A. Akslen, Anders Molven.

Paris: Marie-Françoise Avril, Brigitte Bressac-de Paillerets, Valérie Chaudru, Nicolas Chateigner, Eve Corda, Patricia Jeannin, Fabienne Lesueur, Mahaut de Lichy, Eve Maubec, Hamida Mohamdi, Florence Demenais and the French Family Study Group including the following Oncogeneticists and Dermatologists: Pascale Andry-Benzaquen, Bertrand Bachollet, Frédéric Bérard, Pascaline Berthet, Françoise Boitier, Valérie Bonadona, Jean-Louis Bonafé, Jean-Marie Bonnetblanc, Frédéric Cambazard, Olivier Caron, Frédéric Caux, Jacqueline Chevrand-Breton, Agnès Chompret (deceased), Stéphane Dalle, Liliane Demange, Olivier Dereure, Martin-Xavier Doré, Marie-Sylvie Doutre, Catherine Dugast, Laurence Faivre, Florent Grange, Philippe Humbert, Pascal Joly, Delphine Kerob, Christine Lasset, Marie Thérèse Leccia, Gilbert Lenoir, Dominique Leroux, Julien Levang, Dan Lipsker, Sandrine Mansard, Ludovic Martin, Tanguy Martin-Denavit, Christine Mateus, Jean-Loïc Michel, Patrice Morel, Laurence Olivier-Faivre, Jean-Luc Perrot, Caroline Robert, Sandra Ronger-Savle, Bruno Sassolas, Pierre Souteyrand, Dominique Stoppa-Lyonnet, Luc Thomas, Pierre Vabres, Eva Wierzbicka.

Philadelphia: David Elder, Peter Kanetsky, Jillian Knorr, Michael Ming, Nandita Mitra, Althea Ruffin, Patricia Van Belle.

Poland: Tadeusz Dębniak, Jan Lubiński, Aneta Mirecka, Sławomir Ertmański.

Slovenia: Srdjan Novakovic, Marko Hocevar, Barbara Peric, Petra Cerkovnik.

Stockholm: Veronica Höiom, Johan Hansson.

Sydney: Graham J. Mann, Richard F. Kefford, Helen Schmid, Elizabeth A. Holland.

Tel Aviv: Esther Azizi, Gilli Galore-Haskel, Eitan Friedman, Orna Baron-Epel, Alon Scope, Felix Pavlitsky, Emanuel Jakobson, Irit Cohen-Manheim, Yael Laitman, Roni Milgrom, Iris Shimoni, Evgeniya Kozlova.

Author Contributions

Conceived and designed the experiments: JHB NS MMI JANB DTB. Performed the experiments: NS. Analyzed the data: NS JHB. Wrote the paper: NS JHB. Designed study and provided samples generating the original GWA data on which the analysis is based: GenoMEL consortium.

References

- McCann J (1999) Can skin cancers be minimized or prevented in organ transplant patients? *J Natl Cancer Inst* 91: 911–3.
- Mire L, Hollowood K, Gray D, Bordea C, Wojnarowska F (2006) Melanomas in renal transplant recipients. *British Journal of Dermatology* 154: 472–477.

3. Nathanson L (1976) Spontaneous regression of malignant melanoma: a review of the literature on incidence, clinical features, and possible mechanisms. *Natl Cancer Inst Monogr* 44: 67–76.
4. Chang Y, Barrett J, Bishop D, Armstrong B, Bataille V, et al. (2009) Sun exposure and melanoma risk at different latitudes: a pooled analysis of 5700 cases and 7216 controls. *International journal of epidemiology* 38: 814–30.
5. Donawho C, Muller H, Bucana C, Kripke M (1996) Enhanced growth of murine melanoma in ultraviolet-irradiated skin is associated with local inhibition of immune effector mechanisms. *The Journal of Immunology* 157: 781.
6. Dunn G, Bruce A, Ikeda H, Old L, Schreiber R (2002) Cancer immunoeediting: from immunosurveillance to tumor escape. *Nature immunology* 3: 991–998.
7. Suci-Foca N, Berloco P, Cortesini R (2009) Tolerogenic dendritic cells in cancer, transplantation, and autoimmune diseases. *Human immunology* 70: 277–80.
8. Polak ME, Borthwick NJ, Gabriel FG, Johnson P, Higgins B, et al. (2007) Mechanisms of local immunosuppression in cutaneous melanoma. *Br J Cancer* 96: 1879–87.
9. Schwartz R (2003) T cell anergy. *Annual Review of Immunology* 21: 305–34.
10. Boon T, Coulic PG, Van den Eynde BJ, van der Bruggen P (2006) Human t cell responses against melanoma. *Annu Rev Immunol* 24: 175–208.
11. Guilloux Y, Viret C, Gervois N, Le Drean E, Pandolfino MC, et al. (1994) Defective lymphokine production by most cd8+ and cd4+ tumor-specific t cell clones derived from human melanoma-infiltrating lymphocytes in response to autologous tumor cells in vitro. *Eur J Immunol* 24: 1966–73.
12. Zou W (2006) Regulatory t cells, tumour immunity and immunotherapy. *Nat Rev Immunol* 6: 295–307.
13. Kryczek I, Liu R, Wang G, Wu K, Shu X, et al. (2009) Foxp3 defines regulatory t cells in human tumor and autoimmune disease. *Cancer Res* 69: 3995–4000.
14. Mourmouras V, Fimiani M, Rubegni P, Epistolato M, Malagnino V, et al. (2007) Evaluation of tumour-infiltrating cd4+ cd25+ foxp3+ regulatory t cells in human cutaneous benign and atypical naevi, melanomas and melanoma metastases. *British Journal of Dermatology* 157: 531–539.
15. Agarwala S (2010) Novel immunotherapies as potential therapeutic partners for traditional or targeted agents: cytotoxic t-lymphocyte antigen-4 blockade in advanced melanoma. *Melanoma Research* 20: 1.
16. Gajewski TF (2007) Failure at the effector phase: immune barriers at the level of the melanoma tumor microenvironment. *Clin Cancer Res* 13: 5256–61.
17. Bouwhuis MG, Gast A, Figl A, Eggermont AM, Hemminki K, et al. (2009) Polymorphisms in the cd28/ctla4/icos genes: role in malignant melanoma susceptibility and prognosis? *Cancer Immunol Immunother* 59: 303–312.
18. Nikolova PN, Pawelec GP, Mihailova SM, Ivanova MI, Myhailova AP, et al. (2007) Association of cytokine gene polymorphisms with malignant melanoma in caucasian population. *Cancer Immunol Immunother* 56: 371–9.
19. Howell WM, Turner SJ, Theaker JM, Bateman AC (2003) Cytokine gene single nucleotide polymorphisms and susceptibility to and prognosis in cutaneous malignant melanoma. *Eur J Immunogenet* 30: 409–14.
20. Schoof N, Von Bonin F, König I, Mössner R, Krüger U, et al. (2009) Distal and proximal interleukin (il)-10 promoter polymorphisms associated with risk of cutaneous melanoma development: a case– control study. *Genes and immunity* 10: 586–590.
21. Bishop DT, Demenais F, Iles MM, Harland M, Taylor JC, et al. (2009) Genome-wide association study identifies three loci associated with melanoma risk. *Nat Genet* 41: 920–5.
22. Chasman DI (2008) On the utility of gene set methods in genomewide association studies of quantitative traits. *Genet Epidemiol* 32: 658–68.
23. Hardy J, Singleton A (2009) Genomewide association studies and human disease. *N Engl J Med* 360: 1759–68.
24. Peng G, Luo L, Siu H, Zhu Y, Hu P, et al. (2010) Gene and pathway-based second-wave analysis of genome-wide association studies. *Eur J Hum Genet* 18: 111–7.
25. Wang K, Li M, Bucan M (2007) Pathway-based approaches for analysis of genomewide association studies. *Am J Hum Genet* 81.
26. Eleftherohorinou H, Wright V, Hoggart C, Hartikainen AL, Jarvelin MR, et al. (2009) Pathway analysis of GWAs provides new insights into genetic susceptibility to 3 inflammatory diseases. *PLoS One* 4: e8068.
27. Holmans P, Green EK, Pahwa JS, Ferreira MA, Purcell SM, et al. (2009) Gene ontology analysis of GWA study data sets provides insights into the biology of bipolar disorder. *Am J Hum Genet* 85: 13–24.
28. Lesnick TG, Papapetropoulos S, Mash DC, Ffrench-Mullen J, Shehadeh L, et al. (2007) A genomic pathway approach to a complex disease: axon guidance and Parkinson disease. *PLoS Genet* 3: e98.
29. Torkamani A, Topol EJ, Schork NJ (2008) Pathway analysis of seven common diseases assessed by genome-wide association. *Genomics* 92: 265–72.
30. Wang K, Zhang H, Kugathasan S, Annesse V, Bradfield JP, et al. (2009) Diverse genome-wide association studies associate the il12/il23 pathway with crohn disease. *Am J Hum Genet* 84: 399–405.
31. Menashe I, Maeder D, Garcia-Closas M, Figueroa J, Bhattacharjee S, et al. (2010) Pathway analysis of breast cancer genome-wide association study highlights three pathways and one canonical signaling cascade. *Cancer Research* 70: 4453.
32. Tintle NL, Borchers B, Brown M, Bekmetjev A (2009) Comparing gene set analysis methods on single-nucleotide polymorphism data from genetic analysis workshop 16. *BMC Proc* 3 Suppl 7: S96.
33. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, et al. (2007) Plink: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 81: 559–575.
34. R Development Core Team (2009) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. Available: URL <http://www.R-project.org>. ISBN 3-900051-07-0. (Accessed 2011 Dec 2).
35. Barrett JH, Iles MM, Harland M, Taylor JC, Aitken JF, et al. (2011) Genome-wide association study identifies three new melanoma susceptibility loci. *Nature Genetics* 2011; 43: 1108–13.
36. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, et al. (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* 102: 15545–50.
37. Efron B, Tibshirani R (2007) On testing the significance of sets of genes. *Ann Appl Stat* 1: 107–129.
38. Nam D, Kim J, Kim S, Kim S (2010) Gsa-snp: a general approach for gene set analysis of polymorphisms. *Nucleic Acids Research* 38 Suppl: W749–54.
39. Chen L, Hutter C, Potter J, Liu Y, Prentice R, et al. (2010) Insights into colon cancer etiology via a regularized approach to gene set analysis of GWAS data. *The American Journal of Human Genetics* 86: 860–871.
40. Mellor A, Munn D (2004) Ido expression by dendritic cells: tolerance and tryptophan catabolism. *Nature Reviews Immunology* 4: 762–774.
41. Munn D, Mellor A (2004) Ido and tolerance to tumors. *Trends in Molecular Medicine* 10: 15–18.
42. Arefayene M, Mamidipalli S, Philips S, Cao D, Flockhart D, et al. (2006) Identification of functional genetic variants of the indoleamine 2, 3 dioxygenase gene. *Proceedings of the American Association for Cancer Research* 2006: 1119.
43. Liu F, Rabinovich G (2005) Galectins as modulators of tumour progression. *Nature Reviews Cancer* 5: 29–41.
44. Hoek K, Schlegel N, Eichhoff O, Widmer D, Praetorius C, et al. (2008) Novel mitf targets identified using a two-step dna microarray strategy. *Pigment Cell & Melanoma Research* 21: 665–676.
45. Vereecken P, Awada A, Suci S, Castro G, Morandini R, et al. (2009) Evaluation of the prognostic significance of serum galectin-3 in American joint committee on cancer stage iii and stage iv melanoma patients. *Melanoma Research* 19: 316.
46. Gorelik L, Flavell R (2002) Transforming growth factor- β in t-cell biology. *Nature Reviews Immunology* 2: 46–53.