



Published in final edited form as:

Biochim Biophys Acta Mol Basis Dis. 2018 June ; 1864(6 Pt B): 2247–2254. doi:10.1016/j.bbadis.2018.01.007.

Integrated case-control and somatic-germline interaction analyses of melanoma susceptibility genes

Yao Yu¹, Hao Hu¹, Jiun-Sheng Chen^{1,2}, Fulan Hu^{1,3}, Jerry Fowler¹, Paul Scheet¹, Hua Zhao¹, and Chad D. Huff¹

¹Department of Epidemiology, The University of Texas MD Anderson Cancer Center, Houston, Texas, USA

²The University of Texas MD Anderson Cancer Center UTHealth Graduate School of Biomedical Sciences, Houston, Texas, USA

³Department of Epidemiology, Public Health College, Harbin Medical University, Harbin, Heilongjiang 150081, China

Abstract

While a number of genes have been implicated in melanoma susceptibility, the role of protein-coding variation in melanoma development and progression remains underexplored. To better characterize the role of germline coding variation in melanoma, we conducted a whole-exome case-control and somatic-germline interaction study involving 322 skin cutaneous melanoma cases from The Cancer Genome Atlas and 3,607 controls of European ancestry. We controlled for cross-platform technological stratification using XPAT and conducted gene-based association tests using VAAST 2. Four established melanoma susceptibility genes achieved nominal statistical significance, *MC1R* ($p = 0.0014$), *MITF* ($p = 0.0165$), *BRCA2* ($p = 0.0206$), and *MTAP* ($p = 0.0393$). We also observed a suggestive association for *FANCA* ($p=0.002$), a gene previously implicated in melanoma survival. The association signal for *BRCA2* was driven primarily by likely gene disrupting (LGD) variants, with an Odds Ratio (OR) of 5.62 (95% Confidence Interval (CI) 1.03 – 30.1). In contrast, the association signals for *MC1R* and *MITF* were driven primarily by predicted pathogenic non-LGD coding variants, with estimated ORs of 1.4 to 3.0 for *MC1R* and 4.1 for *MITF*. *MTAP* exhibited an excess of both LGD and predicted damaging non-LGD variants among cases, with ORs of 5.62 and 3.72, respectively, although neither category was significant. For individuals with known or predicted damaging variants, age of disease onset was significantly lower for two of the four genes, *MC1R* ($p=0.005$) and *MTAP* ($p=0.035$). In an analysis of germline carrier status and overlapping copy number alterations, we observed no

Correspondence should be addressed to C.H. (chad@hufflab.org).

Author contribution

CH conceived of the study; YY, HH, JC, FH, and CH conducted association analyses; JF, PS, YY, and CH conducted copy number alteration analyses; CH, YY, PS, and HZ interpreted results; YY prepared figures; and YY and CH prepared and wrote the manuscript. All authors approved submission of the manuscript in its present form.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Competing interests

The authors declare that they have no competing interests.

evidence to support a two-hit model of carcinogenesis in any of the four genes. Although *MC1R* carriers were represented proportionally among the four molecular tumor subtypes, these individuals accounted for 69% of ultraviolet (UV) radiation mutational signatures among triple-wild type tumors ($p = 0.040$), highlighting the increased sensitivity to UV exposure among individuals with loss-of-function variants in *MC1R*.

Keywords

case-control study; whole exome association analysis; skin cutaneous melanoma; cancer susceptibility gene; somatic-germline interaction

1. Introduction

Despite the potential effectiveness of early detection efforts, cutaneous melanoma remains a deadly disease that continues to claim over 7,000 lives per year in the United States alone [1]. This juxtaposition highlights the growing need for improved identification of individuals at high risk of developing melanoma. Although the disease has a strong hereditary component, the preponderance of familial clustering in melanoma remains unexplained. Familial studies have identified multiple high-risk melanoma genes, including *CDKN2A* [2] and *CDK4* [3], but the proportion of sporadic cases attributable to susceptibility genes identified in a familial context has been underexplored. A number of intermediate risk melanoma-susceptibility genes, with variants conferring a 2-fold to 8-fold increased risk, have been identified through a combination of case-control and familial studies, including *MC1R* [4, 5], *MITF* [6, 7], *BRCA2* [8, 9], *TERT* [10], *BAP1* [11], *POT1* [12], and *MTAP* [13]. The full spectrum of disease risk conferred by rare variants in these genes remains to be elucidated.

Detailed characterization of genomic alterations in melanoma has revealed four major tumor subtypes defined by the occurrence of somatic point mutations, mutant *BRAF*, mutant *RAS*, mutant *NFI*, and triple-WT (wild-type) [14]. Due to reduced DNA repair activity, *BRAF*, *RAS*, and *NFI* subtypes are characterized by high-rates of mutation from ultraviolet (UV) radiation exposure, with over 90% of tumors harboring a UV mutational signature. In contrast, only 30% of triple-WT tumors are known to harbor a UV mutational signature. No prior study has evaluated the relationship between established melanoma risk genes, tumor subtypes, and tumor mutational patterns.

To evaluate the contribution of rare, protein-coding variation to melanoma risk, development, and progression, we conducted a gene-based whole-exome case-control study involving 322 cases from The Cancer Genome Atlas (TCGA) project and 3,607 controls of European ancestry. We then combined the case-control results with whole-exome sequencing data of tumor-derived DNA to test for patterns of somatic-germline interaction.

2. Material and Methods

2.1 Data sources

The initial set of cases consisted of 387 individuals diagnosed with skin cutaneous melanoma (SKCM) from The Cancer Genome Atlas (TCGA) project, with approval for the access to the Cancer Genomics Hub (<https://cghub.ucsc.edu/>). Controls consisted of 4,674 unaffected parents from autism parent-offspring trios in the Simons Simplex Collection (SSC) [15, 16], downloaded from National Database for Autism Research (NDAR). All germline variant calls were generated using sequencing data from blood derived DNA. Other than ASD status, phenotype information was unavailable for the controls and thus a small proportion may have been misclassified, although the potential misclassification bias would have only a modest influence on power and effect size estimates [17].

2.2 Sequencing data processing and cross-platform quality control

We downloaded raw read data in FASTQ format for all cases and controls, where available. If only BAM files were available, we extracted the sequencing reads from BAM files to generate FASTQ files for each individual. We conducted data alignment and variant calling steps using the XPAT pipeline [18], which included the following steps: (1) alignment of FASTQ data for each individual to the human reference genome (HG19) using BWA [19] (v0.7.9a), Samtools [20] (version 0.1.19), Picard (v1.118), sample level variant calling using GATK HaplotypeCaller[21] (v3.3) to generate gvcf files, (3) joint genotype calling of all samples from gvcf files using GATK HaplotypeCaller [21], and (4) variant recalibration using GATK Variant Recalibrator metrics with the tranche sensitivity score 99.9 for SNVs and 98.0 for INDELS. We converted variant calls with genotype quality scores less than five to missing genotypes.

We used the XQC module in XPAT to perform sample level and variant level quality control (QC) with default parameters [18], described briefly as follows: For sample level QC, we excluded samples with NC90 > 0.05%, defined as the proportion of sites called in this individual among the subset of sites with 90% or greater call rates for each platform. Due to insufficient sample sizes, we also excluded any sample with ethnicity reported as Hispanic or Latino. For variant level QC, we excluded variants based on the default criteria in XPAT, which include the following: (1) allelic balance (proportion of reads supporting minor allele among all reads) < 20%, (2) missing genotype calling rate > 10% in any platform, (4) Hardy-Weinberg disequilibrium test with $p < 10^{-6}$ in controls, (4) $p < 0.05$ in a cross-platform differential missing genotype rate test among both cases and controls, and (5) $p < 0.05$ in a cross-platform differential allele frequency test within cases and controls.

2.3 Case-control analyses

To identify population outliers, we conducted an initial principal component analysis (PCA) using an external reference panel of 427 individuals from the 1000 Genomes project (phase3–20130502) [22]. We included a total of 19,053 SNPs that passed QC with minor allele frequency (MAF) > 10%, LD-pruned to $r^2 < 0.2$. We projected cases and controls onto the PC space and selected samples that clustered with the European group to identify our final set of 322 cases and 3,607 controls (Supplementary Figure S1 and S2). We then

conducted a second PCA using the combined set of cases and controls, excluding the reference panel. We included a total of 15,071 SNPs that passed QC with MAF >10% among non-Finnish Europeans reported in ExAC database, LD-pruned to $r^2 < 0.2$ [23] (Supplementary Figure S3).

We conducted gene-based association tests using VAAST 2 with the allele frequency among controls constrained to 0.1 or lower in the likelihood model (parameter: $r = 0.1$). Variants were weighted by the conservation-controlled amino acid substitution matrix (CASM) scores, as previously described [24, 25]. VAAST 2 uses a one-sided burden test to test whether damaging variants are overrepresented in the cases versus controls. We used the covariate matrix from PCA with a biasedUrn sampling method [26] implemented in VAAST 2, with the first two PCs included as covariates. For genes with multiple isoforms, we calculated gene-based p-values using the Multiple Gene Isoform Test (MGIT) implemented in XPAT [18], which is a permutation-based test that jointly evaluates all isoforms in a given gene.

2.4 Copy number alterations

To detect somatic Copy Number Alterations (CNA), we downloaded Affymetrix SNP-6 CEL files from 474 tumor-normal pairs; we used the 318 of these pairs that overlapped with selected germline sequencing data in the subsequent analysis. We ran Birdsuite [27] on CEL files of all normal samples to get the reference SNP genotypes uninfluenced by somatic variation for the phasing. We then ran Birdsuite again on all CEL files of all samples together to obtain B allele frequencies and logR ratios for the tumors. As part of standard processing, we excluded probes from mitochondrial and Y chromosomes, as well as markers spanning the HLA region, which exhibits gene duplications. We applied Mach [28] to phase the samples in the germline set. We injected the normal phasing into each (paired) tumor sample file and applied hapLOH [29] to infer allelic imbalance and copy number alterations in all tumor samples. We reported regions of allelic imbalance where a posterior probability of 95% of imbalance was exceeded. We then combined all hapLOH CNA calls with copy number gain and loss events reported in cBioPortal (<http://www.cbioportal.org/>), with segment mean somatic event $> \log_2(2.5/2)$ for copy number gains and $< \log_2(1.5/2)$ for copy number losses.

3. Results

We conducted a gene-based case-control association analysis of protein-coding variants across the exome using the Variant Annotation, Analysis and Search Tool (VAAST 2, version 2.1) [24, 25]. The analysis considered both LGD (stop gained, stop lost, splicing region variants and frame-shift INDELS) and non-LGD (missense and inframe INDELS) variants, weighting each variant according to the estimated degree of protein dysfunction conferred based on the CASM score in VAAST [24]. The initial level of technological stratification between cases and controls was substantial, due to differences in target capture and sequencing protocols. However, after applying the cross-platform QC pipeline in XPAT [18], we observed no overall inflation in Type I error (Supplementary Figure S4). Although no gene reached genome-wide significance, the association analysis identified four genes

previously implicated in melanoma risk at a nominal significance level of 0.05 (Table 1): *MC1R* [4] ($p = 0.0014$), *MITF* [30] ($p = 0.0165$), *BRCA2* [9] ($p = 0.0206$), and *MTAP* [13] ($p = 0.0393$), Supplementary Table 1 presents the full list of gene-based association results.

To assess whether the association signals in these genes were driven by known pathogenic variants or variants expected to disrupt protein expression, we excluded LGD variants and repeated the association analysis. *MC1R* and *MITF* exhibited no meaningful change in association signal, while *BRCA2* exhibited only a modest attenuation in signal, (Table 1), suggesting that missense variants with unconfirmed pathogenicity in these genes increase melanoma risk.

To estimate variant effect sizes, we calculated crude odds ratios (ORs) according to the classical method [31–33] and adjusted ORs with logistic regression, incorporating the first two PCs as covariates. Table 2 reports variant-specific ORs for variants with five or more allele counts among both cases and controls. For rare variants with $MAF < 0.005$, we calculated collapsed gene-based effect sizes based on the following four functional categories: all coding variants, LGD variants, predicted damaging non-LGD variants (CASM score ≥ 1.0 and PolyPhen-2 score ≥ 0.909), and predicted benign non-LGD variants (CASM score < 1.0 and PolyPhen-2 score ≤ 0.446), as shown in Figure 1 and Table 3.

To conduct somatic-germline interaction analyses, we first classified cases according to their carrier status for *MC1R*, *MITF*, *BRCA2*, and *MTAP*. Here, we defined a carrier as any individual with a germline variant which received a VAAST score > 1.0 in the case-control analysis. The VAAST score incorporates both the CASM score and case-control allele frequency information to identify variants predicted to increase disease risk.

We evaluated two-hit models of carcinogenesis by testing the hypothesis that, for a given gene, somatic mutational events were more frequent among carriers than non-carriers [34]. We considered four classes of mutational events: coding point mutations, copy number gains, copy number losses, and undefined CNAs. The waterfall plot in Figure 2 depicted the landscape of germline variants and somatic events in the four selected melanoma susceptibility genes. The undefined category represents allelic imbalance events that could only be detected by hapLOH; these include CNAs with copy neutral loss-of-heterozygosity as well as CNAs present at low frequencies in tumor subclones [29]. We also conducted additional tests evaluating copy number gain and loss separately. The analysis identified no significant associations between carrier status and somatic mutational events for any of the four genes (Supplemental Table S2). In general, somatic mutational events were proportionally represented among carriers and non-carriers in each gene.

To test for associations with UV mutational signatures, we considered a total of 258 samples with available UV signature, tumor subtype information, and germline data [14]. We tested the fractions of germline damaging variants carrier between samples with and without UV signatures for each subtype. We observed that 69% of triple WT samples with UV signatures were *MC1R* carriers, compared to only 33% of samples without UV signatures ($p = 0.040$, one-sided Fisher's exact test, see Figure 3A). The other three tumor subtypes were

characterized by a high rate of UV mutational signatures among both *MC1R* carriers and non-carriers, with no significant differences.

We also tested for potential relationships between germline susceptibility variants and age at disease onset using a one-sided Wilcoxon rank sum test. Both *MTAP* and *MC1R* carriers were diagnosed with melanoma at significantly younger ages, with $p = 0.035$ and 0.005 , respectively (see Figure 3B and 3C). The mean age at diagnosis was 35.5 among *MTAP* carriers, 53.9 among *MC1R* carriers, and 57.7 among non-carriers. We observed no significant association with age at diagnosis among *BRCA2* or *MITF* carriers (Figure 3D and 3E).

4. Discussion

Continued advances in next-generation sequencing technologies present growing opportunities to conduct large-scale sequence-based association studies by pooling data from multiple original sources. These studies can potentially have a major impact on precision medicine through more accurate effect size estimates for rare variants in known susceptibility genes as well as improved statistical power to detect new disease associations. However, technological stratification biases resulting from the heterogeneous nature of next-generation sequencing technologies have been a major barrier to pooled analysis efforts [35]. In this study, we initially employed a set of QC metrics typically applied in sequence-based association studies and observed substantial inflation in Type I error. However, after applying the cross-platform QC metrics in XPAT, we observed no Type I error inflation and thus were able to successfully control for technological stratification biases (Supplemental Figure S4). Our gene-based association analysis replicated four established melanoma susceptibility genes: *MC1R*, *MITF*, *BRCA2*, and *MTAP*. For variants in these genes with sufficient allele counts for individual effect size estimates, the odds ratios were broadly consistent with previous estimates (Figure 1B). We observed no evidence supporting a two-hit model of carcinogenesis in any of the four genes, indicating that only one copy of an inherited predisposition allele at the cellular level was sufficient to help drive tumor initiation.

Gene-based association tests increase statistical power to identify disease associations by aggregating information from multiple variants. After a gene-level association has been identified, these tests generally provide limited information regarding the increase in risk conferred by any given variant. To overcome this problem, effect size estimates are often obtained by collapsing variants into specific functional categories based on *a priori in silico* predictions. Given the inherent imprecision of *in silico* functional variant prediction tools [24], such estimates are necessarily biased by misclassification errors, but nonetheless provide substantial insights into the degree of risk conferred by rare functional variants in a gene of interest. In this study, we considered three functional variant categories for gene-based rare variant effect size estimates: LGD and known pathogenic variants, non-LGD variants predicted to be damaging by both CASM and PP2, and non-LGD variants predicted to be benign by both CASM and PP2. Although the confidence intervals were wide, LGD variants in *BRCA2* and predicted damaging non-LGD variants in *MITF* all exhibited odds

ratios of greater than 1. In contrast, the odds ratio estimates for variants predicted to be benign ranged from 0 to 1.4 (Figure 1A, Table 3).

Missense variants in *MC1R* are associated with phenotypical features such as light skin pigmentation and red hair, in addition to an increased risk of cutaneous melanoma [36]. We identified four *MC1R* variants (rs1805006, rs1805007, rs1805008, rs1805009) with ORs significantly greater than 1 in this study (Table 2). All four variants are known melanoma risk factors and are strongly associated with red hair [37]. The OR estimates of these variants ranged from 1.4 to 3.0 and were consistent with previous reports (Table 2) [38]. The association between *MC1R* carrier status and increased UV mutational signatures in triple-WT tumors is likely due to an interaction between lighter skin pigmentation and UV exposure. The absence of an association with UV signatures and *MC1R* carrier status among *BRAF*, *RAS*, and *NF1* tumor subtypes may result from the high rates of UV signatures among these subtypes. The modest but significant relationship between *MC1R* carrier status and younger age at melanoma diagnosis could be explained by age-specific increases in relative risk or by earlier detection due to recognized phenotypic risk factors.

MITF (melanogenesis associated transcription factor) encodes a transcription factor that regulates melanocyte development and is responsible for pigment cell-specific transcription of the melanogenesis enzyme genes. We identified an association between melanoma and a rare functional non-synonymous SNV (rs149617956), with an adjusted OR of 4.48 (95% CI: 1.57 to 12.78). Previous studies have reported somewhat higher OR estimates for this variant [6, 7], although the confidence intervals overlap (Figure 1B, Table 2). Based on a meta-analysis of our estimate and prior work, the estimated OR of this variant is 6.01 (95% CI: 2.81 to 12.89). This variant is also associated with a higher incidence of renal cancer, increased nevus count, and non-blue eye color [39]. Although the gene-based OR estimate for rare, predicted pathogenic non-LGD variants was significantly larger than 1.0, this association was driven solely by rs149617956.

MTAP (methylthioadenosine phosphorylase) encodes an enzyme that plays a major role in polyamine metabolism. Previous studies have reported multiple variants associated with occurrence of cutaneous melanoma [13, 40]. In this study, VAAST identified two additional potential risk variants in *MTAP*, one LGD and one non-LGD. Although only two *MTAP* carriers were identified in this study, the age at diagnosis was significantly different between carriers and non-carriers. Both *MTAP* carriers were diagnosed with melanoma before the age of 45, compared to an average age at diagnosis among all study participants of 56.7. To our knowledge, associations with *MTAP* susceptibility variants and early-onset melanoma have not been previously reported.

BRCA2 is a cancer susceptibility gene with a well-established pattern of dominant inheritance in several cancers, including melanoma. *BRCA2* is also one of the thirteen genes involved in Fanconi anemia, a recessive disorder characterized by genomic instability and increased susceptibility to leukemia and cancer [41, 42]. The estimated effect size for LGD variants in *BRCA2* was somewhat higher than previously reported, with a crude OR of 5.6 and an adjusted OR of 3.4. However, given the wide confidence intervals, these results are consistent with previous relative risk estimate of 2.6 (95% CI 1.3 to 5.2) for pathogenic

BRCA2 variants [43]. We also replicated the association with the *BRCA2* coding variant rs1799944, which had a MAF of 0.03 among controls. Our estimated OR for this variant was 1.87 (95% CI 1.47 to 2.38) based on a meta-analysis of our results and a previous estimate. Note that rs766173 and rs1799944 are in near complete linkage disequilibrium ($r^2 = .89$), which complicates interpretations of causality. Although *BRCA2* has been reported to follow a two hit model in some cancers [34], this does not appear to be an important mechanism for tumorigenesis in melanoma.

Notably, a second Fanconi anemia pathway gene, *FANCA*, exhibited a nominally significant association with melanoma risk ($p = 0.0020$). Coding variants in both *FANCA* and *BRCA2* have been previously associated with overall survival of melanoma patients [44]. Our study also identified two additional candidate melanoma susceptibility genes with nominal p less than 0.05 that have been identified as susceptibility genes in other cancers (Supplementary Table S1): *RAD50* ($p = 0.0387$) and *HNFI1A* ($p = 0.0332$). *HNFI1A* was recently reported as a pancreatic cancer susceptibility locus in a genome-wide pleiotropy scan [45]. One of the targets of *HNFI1A* is the melanoma inhibitory activity 2 (*MIA2*) gene [46], the overexpression of which promotes metastatic behavior of malignant melanoma [47]. *RAD50* is an established intermediate breast cancer risk gene [48–50], but has not been previously implicated in melanoma susceptibility. Note that *FANCA*, *HNFI1A*, and *RAD50* were not significant after multiple testing correction and are highlighted here based on *post hoc* exploratory analyses.

Taken together, our findings highlight the potential importance of damaging protein-coding variation in melanoma susceptibility. However, because this study was restricted to a relatively small sample of individuals of European ancestry, the statistical power to identify new melanoma-gene associations was limited and the relevance of our findings in other populations is uncertain. Additional sequencing studies with larger sample sizes involving multiple ancestry backgrounds are needed to more fully characterize the contribution of rare, protein-coding variation to melanoma development and progression.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

An allocation of computer time on the University of Texas MD Anderson Research Computing High Performance Computing (HPC) facility is gratefully acknowledged.

Funding

This work was supported by US National Institutes of Health grants R01 CA195614, R01 GM104390, and R01 HG005859.

Abbreviations

CASM	conservation-controlled amino acid substitution matrix
CI	confidence interval

CNA	copy number alteration
ExAC	Exome Aggregation Consortium
GATK	Genome Analysis Toolkit
LGD	likely gene-disrupting
NDAR	National Database for Autism Research
OR	odds ratio
PCA	principle component analysis
QC	quality control
SGI	somatic-germline interaction
SKCM	skin cutaneous Melanoma
SNP	single nucleotide polymorphism
SNV	single nucleotide variant
SSC	Simons Simplex Collection
TCGA	The Cancer Genome Atlas
VAAST2	The Variant Annotation, Analysis and Search Tool
XPAT	The cross-Platform Association Toolkit

References

- [1]. Howlader N, Noone AM, Krapcho M, Miller D, Bishop K, Kosary C, Yu M, Ruhl J, Tatalovich Z, Mariotto A, Lewis DR, Chen HS, Feuer EJ, Cronin KA, SEER Cancer Statistics Review, 1975–2014, National Cancer Institute Bethesda, MD, https://seer.cancer.gov/csr/1975_2014/, based on November 2016 SEER data submission, posted to the SEER web site, (4 2017).
- [2]. Hussussian CJ, Struewing JP, Goldstein AM, Higgins PA, Ally DS, Sheahan MD, Clark WH Jr., Tucker MA, Dracopoli NC, Germline p16 mutations in familial melanoma, *Nature genetics*, 8 (1994) 15–21. [PubMed: 7987387]
- [3]. Zuo L, Weger J, Yang Q, Goldstein AM, Tucker MA, Walker GJ, Hayward N, Dracopoli NC, Germline mutations in the p16INK4a binding domain of CDK4 in familial melanoma, *Nature genetics*, 12 (1996) 97–99. [PubMed: 8528263]
- [4]. Palmer JS, Duffy DL, Box NF, Aitken JF, O’Gorman LE, Green AC, Hayward NK, Martin NG, Sturm RA, Melanocortin-1 receptor polymorphisms and risk of melanoma: is the association explained solely by pigmentation phenotype?, *American journal of human genetics*, 66 (2000) 176–186. [PubMed: 10631149]
- [5]. Macgregor S, Montgomery GW, Liu JZ, Zhao ZZ, Henders AK, Stark M, Schmid H, Holland EA, Duffy DL, Zhang M, Painter JN, Nyholt DR, Maskiell JA, Jetann J, Ferguson M, Cust AE, Jenkins MA, Whiteman DC, Olsson H, Puig S, Bianchi-Scarra G, Hansson J, Demenais F, Landi MT, Debniak T, Mackie R, Azizi E, Bressac-de Paillerets B, Goldstein AM, Kanetsky PA, Gruis NA, Elder DE, Newton-Bishop JA, Bishop DT, Iles MM, Helsing P, Amos CI, Wei Q, Wang LE, Lee JE, Qureshi AA, Kefford RF, Giles GG, Armstrong BK, Aitken JF, Han J, Hopper JL, Trent JM, Brown KM, Martin NG, Mann GJ, Hayward NK, Genome-wide association study identifies

a new melanoma susceptibility locus at 1q21.3, *Nature genetics*, 43 (2011) 1114–1118. [PubMed: 21983785]

- [6]. Bertolotto C, Lesueur F, Giuliano S, Strub T, de Lichy M, Bille K, Dessen P, d'Hayer B, Mohamdi H, Remenieras A, Maubec E, de la Fouchardiere A, Molinie V, Vabres P, Dalle S, Poulalhon N, Martin-Denavit T, Thomas L, Andry-Benzaquen P, Dupin N, Boitier F, Rossi A, Perrot JL, Labeille B, Robert C, Escudier B, Caron O, Brugieres L, Saule S, Gardie B, Gad S, Richard S, Couturier J, Teh BT, Ghiorzo P, Pastorino L, Puig S, Badenas C, Olsson H, Ingvar C, Rouleau E, Lidereau R, Bahadoran P, Vielh P, Corda E, Blanche H, Zelenika D, Galan P, French G Familial Melanoma Study, Aubin F, Bachollet B, Becuwe C, Berthet P, Bignon YJ, Bonadona V, Bonafe JL, Bonnet-Dupeyron MN, Cambazard F, Chevrant-Breton J, Coupier I, Dalac S, Demange L, d'Incan M, Dugast C, Faivre L, Vincent-Fetita L, Gauthier-Villars M, Gilbert B, Grange F, Grob JJ, Humbert P, Janin N, Joly P, Kerob D, Lasset C, Leroux D, Levang J, Limacher JM, Livideanu C, Longy M, Lortholary A, Stoppa-Lyonnet D, Mansard S, Mansuy L, Marrou K, Mateus C, Maugard C, Meyer N, Nogues C, Souteyrand P, Venat-Bouvet L, Zattara H, Chaudru V, Lenoir GM, Lathrop M, Davidson I, Avril MF, Demenais F, Ballotti R, Bressac-de Paillerets B, A SUMOylation-defective MITF germline mutation predisposes to melanoma and renal carcinoma, *Nature*, 480 (2011) 94–98. [PubMed: 22012259]
- [7]. Yokoyama S, Woods SL, Boyle GM, Aoude LG, MacGregor S, Zismann V, Gartside M, Cust AE, Haq R, Harland M, Taylor JC, Duffy DL, Holohan K, Dutton-Regester K, Palmer JM, Bonazzi V, Stark MS, Symmons J, Law MH, Schmidt C, Lanagan C, O'Connor L, Holland EA, Schmid H, Maskiell JA, Jetann J, Ferguson M, Jenkins MA, Kefford RF, Giles GG, Armstrong BK, Aitken JF, Hopper JL, Whiteman DC, Pharoah PD, Easton DF, Dunning AM, Newton-Bishop JA, Montgomery GW, Martin NG, Mann GJ, Bishop DT, Tsao H, Trent JM, Fisher DE, Hayward NK, Brown KM, A novel recurrent mutation in MITF predisposes to familial and sporadic melanoma, *Nature*, 480 (2011) 99–103. [PubMed: 22080950]
- [8]. Debniak T, Scott RJ, Gorski B, Cybulski C, van de Wetering T, Serrano-Fernandez P, Huzarski T, Byrski T, Nagay L, Debniak B, Kowalska E, Jakubowska A, Gronwald J, Wokolorczyk D, Maleszka R, Kladny J, Lubinski J, Common variants of DNA repair genes and malignant melanoma, *Eur J Cancer*, 44 (2008) 110–114. [PubMed: 18024013]
- [9]. Ginsburg OM, Kim-Sing C, Foulkes WD, Ghadirian P, Lynch HT, Sun P, Narod SA, Hereditary G Breast Cancer Clinical Study, BRCA1 and BRCA2 families and the risk of skin cancer, *Familial cancer*, 9 (2010) 489–493. [PubMed: 20809262]
- [10]. Law MH, Montgomery GW, Brown KM, Martin NG, Mann GJ, Hayward NK, MacGregor S, M. Q, A. Investigators, Meta-analysis combining new and existing data sets confirms that the TERT-CLPTM1L locus influences melanoma risk, *The Journal of investigative dermatology*, 132 (2012) 485–487. [PubMed: 21993562]
- [11]. Carbone M, Ferris LK, Baumann F, Napolitano A, Lum CA, Flores EG, Gaudino G, Powers A, Bryant-Greenwood P, Krausz T, Hyjek E, Tate R, Friedberg J, Weigel T, Pass HI, Yang H, BAP1 cancer syndrome: malignant mesothelioma, uveal and cutaneous melanoma, and MBAITs, *J Transl Med*, 10 (2012) 179. [PubMed: 22935333]
- [12]. Robles-Espinoza CD, Harland M, Ramsay AJ, Aoude LG, Quesada V, Ding Z, Pooley KA, Pritchard AL, Tiffen JC, Petljak M, Palmer JM, Symmons J, Johansson P, Stark MS, Gartside MG, Snowden H, Montgomery GW, Martin NG, Liu JZ, Choi J, Makowski M, Brown KM, Dunning AM, Keane TM, Lopez-Otin C, Gruis NA, Hayward NK, Bishop DT, Newton-Bishop JA, Adams DJ, POT1 loss-of-function variants predispose to familial melanoma, *Nature genetics*, 46 (2014) 478–481. [PubMed: 24686849]
- [13]. Gibbs DC, Orlov I, Kanetsky PA, Luo L, Krickler A, Armstrong BK, Anton-Culver H, Gruber SB, Marrett LD, Gallagher RP, Zanetti R, Rosso S, Dwyer T, Sharma A, La Pilla E, From L, Busam KJ, Cust AE, Ollila DW, Begg CB, Berwick M, Thomas NE, Group GEMS, Inherited genetic variants associated with occurrence of multiple primary melanoma, *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*, 24 (2015) 992–997.
- [14]. Cancer Genome Atlas N, Genomic Classification of Cutaneous Melanoma, *Cell*, 161 (2015) 1681–1696. [PubMed: 26091043]
- [15]. Iossifov I, O'Roak BJ, Sanders SJ, Ronemus M, Krumm N, Levy D, Stessman HA, Witherspoon KT, Vives L, Patterson KE, Smith JD, Paepfer B, Nickerson DA, Dea J, Dong S, Gonzalez LE,

Mandell JD, Mane SM, Murtha MT, Sullivan CA, Walker MF, Waqar Z, Wei L, Willsey AJ, Yamrom B, Lee YH, Grabowska E, Dalkic E, Wang Z, Marks S, Andrews P, Leotta A, Kendall J, Hakker I, Rosenbaum J, Ma B, Rodgers L, Troge J, Narzisi G, Yoon S, Schatz MC, Ye K, McCombie WR, Shendure J, Eichler EE, State MW, Wigler M. The contribution of de novo coding mutations to autism spectrum disorder, *Nature*, 515 (2014) 216–221. [PubMed: 25363768]

- [16]. Fischbach GD, Lord C. The Simons Simplex Collection: a resource for identification of autism genetic risk factors, *Neuron*, 68 (2010) 192–195. [PubMed: 20955926]
- [17]. Colhoun HM, McKeigue PM, Davey Smith G. Problems of reporting genetic associations with complex outcomes, *Lancet*, 361 (2003) 865–872. [PubMed: 12642066]
- [18]. Yu Y, Hu H, Bohlender R, Hu F, Chen J, Holt C, Fowler J, Guthery SL, Scheet P, Hildebrandt MAT, Yandell M, Huff C, XPAT: A toolkit to conduct cross-platform association studies with heterogeneous sequencing datasets, *Nucleic Acids Res*, In press.
- [19]. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform, *Bioinformatics*, 25 (2009) 1754–1760. [PubMed: 19451168]
- [20]. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, Genome S Project Data Processing, The Sequence Alignment/Map format and SAMtools, *Bioinformatics*, 25 (2009) 2078–2079. [PubMed: 19505943]
- [21]. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytzky A, Garimella K, Altshuler D, Gabriel S, Daly M, DePristo MA, The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data, *Genome research*, 20 (2010) 1297–1303. [PubMed: 20644199]
- [22]. Genomes Project C, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, Korbel JO, Marchini JL, McCarthy S, McVean GA, Abecasis GR, A global reference for human genetic variation, *Nature*, 526 (2015) 68–74. [PubMed: 26432245]
- [23]. Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, O'Donnell-Luria AH, Ware JS, Hill AJ, Cummings BB, Tukiainen T, Birnbaum DP, Kosmicki JA, Duncan LE, Estrada K, Zhao F, Zou J, Pierce-Hoffman E, Berghout J, Cooper DN, Deflaux N, DePristo M, Do R, Flannick J, Fromer M, Gauthier L, Goldstein J, Gupta N, Howrigan D, Kiezun A, Kurki MI, Moonshine AL, Natarajan P, Orozco L, Peloso GM, Poplin R, Rivas MA, Ruano-Rubio V, Rose SA, Ruderfer DM, Shakir K, Stenson PD, Stevens C, Thomas BP, Tiao G, Tusie-Luna MT, Weisburd B, Won HH, Yu D, Altshuler DM, Ardissino D, Boehnke M, Danesh J, Donnelly S, Elosua R, Florez JC, Gabriel SB, Getz G, Glatt SJ, Hultman CM, Kathiresan S, Laakso M, McCarroll S, McCarthy MI, McGovern D, McPherson R, Neale BM, Palotie A, Purcell SM, Saleheen D, Scharf JM, Sklar P, Sullivan PF, Tuomilehto J, Tsuang MT, Watkins HC, Wilson JG, Daly MJ, MacArthur DG, Exome Aggregation C, Analysis of protein-coding genetic variation in 60,706 humans, *Nature*, 536 (2016) 285–291. [PubMed: 27535533]
- [24]. Hu H, Huff CD, Moore B, Flygare S, Reese MG, Yandell M, VAAST 2.0: improved variant classification and disease-gene identification using a conservation-controlled amino acid substitution matrix, *Genetic epidemiology*, 37 (2013) 622–634. [PubMed: 23836555]
- [25]. Yandell M, Huff C, Hu H, Singleton M, Moore B, Xing J, Jorde LB, Reese MG, A probabilistic disease-gene finder for personal genomes, *Genome research*, 21 (2011) 1529–1542. [PubMed: 21700766]
- [26]. Epstein MP, Duncan R, Jiang Y, Conneely KN, Allen AS, Satten GA, A permutation procedure to correct for confounders in case-control studies, including tests of rare variation, *American journal of human genetics*, 91 (2012) 215–223. [PubMed: 22818855]
- [27]. Korn JM, Kuruvilla FG, McCarroll SA, Wysoker A, Nemesh J, Cawley S, Hubbell E, Veitch J, Collins PJ, Darvishi K, Lee C, Nizzari MM, Gabriel SB, Purcell S, Daly MJ, Altshuler D, Integrated genotype calling and association analysis of SNPs, common copy number polymorphisms and rare CNVs, *Nature genetics*, 40 (2008) 1253–1260. [PubMed: 18776909]
- [28]. Li Y, Willer CJ, Ding J, Scheet P, Abecasis GR, MaCH: using sequence and genotype data to estimate haplotypes and unobserved genotypes, *Genetic epidemiology*, 34 (2010) 816–834. [PubMed: 21058334]
- [29]. Vattathil S, Scheet P, Haplotype-based profiling of subtle allelic imbalance with SNP arrays, *Genome research*, 23 (2013) 152–158. [PubMed: 23028187]

- [30]. Garraway LA, Widlund HR, Rubin MA, Getz G, Berger AJ, Ramaswamy S, Beroukhir R, Milner DA, Granter SR, Du JY, Lee C, Wagner SN, Li C, Golub TR, Rimm DL, Meyerson ML, Fisher DE, Sellers WR, Integrative genomic analyses identify MITF as a lineage survival oncogene amplified in malignant melanoma, *Nature*, 436 (2005) 117–122. [PubMed: 16001072]
- [31]. Altman DG, *Practical statistics for medical research*, CRC press, Place Published, 1990.
- [32]. Deeks JJ, Higgins JP, *Statistical algorithms in review manager 5*, Statistical Methods Group of The Cochrane Collaboration, (2010) 1–11.
- [33]. Pagano M, Gauvreau K, Pagano M, *Principles of biostatistics*, Duxbury Pacific Grove, CA, Place Published, 2000.
- [34]. Hu H, Huff CD, Detecting statistical interaction between somatic mutational events and germline variation from next-generation sequence data, *Pac Symp Biocomput.*, (2014) 51–62. [PubMed: 24297533]
- [35]. Freedman ML, Reich D, Penney KL, McDonald GJ, Mignault AA, Patterson N, Gabriel SB, Topol EJ, Smoller JW, Pato CN, Pato MT, Petryshen TL, Kolonel LN, Lander ES, Sklar P, Henderson B, Hirschhorn JN, Altshuler D, Assessing the impact of population stratification on genetic association studies, *Nature genetics*, 36 (2004) 388–393. [PubMed: 15052270]
- [36]. Kennedy C, ter Huurne J, Berkhout M, Gruis N, Bastiaens M, Bergman W, Willemze R, Bavinck JN, Melanocortin 1 receptor (MC1R) gene variants are associated with an increased risk for cutaneous melanoma which is largely independent of skin type and hair color, *The Journal of investigative dermatology*, 117 (2001) 294–300. [PubMed: 11511307]
- [37]. Duffy DL, Zhao ZZ, Sturm RA, Hayward NK, Martin NG, Montgomery GW, Multiple pigmentation gene polymorphisms account for a substantial proportion of risk of cutaneous malignant melanoma, *The Journal of investigative dermatology*, 130 (2010) 520–528. [PubMed: 19710684]
- [38]. Raimondi S, Sera F, Gandini S, Iodice S, Caini S, Maisonneuve P, Fargnoli MC, MC1R variants, melanoma and red hair color phenotype: a meta-analysis, *International journal of cancer. Journal international du cancer*, 122 (2008) 2753–2760. [PubMed: 18366057]
- [39]. Law MH, Macgregor S, Hayward NK, Melanoma genetics: recent findings take us beyond well-traveled pathways, *The Journal of investigative dermatology*, 132 (2012) 1763–1774. [PubMed: 22475760]
- [40]. Liede A, Karlan BY, Narod SA, Cancer risks for male carriers of germline mutations in BRCA1 or BRCA2: a review of the literature, *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*, 22 (2004) 735–742. [PubMed: 14966099]
- [41]. Kee Y, D'Andrea AD, Expanded roles of the Fanconi anemia pathway in preserving genomic stability, *Genes Dev*, 24 (2010) 1680–1694. [PubMed: 20713514]
- [42]. Garcia-Higuera I, Taniguchi T, Ganesan S, Meyn MS, Timmers C, Hejna J, Grompe M, D'Andrea AD, Interaction of the Fanconi anemia proteins and BRCA1 in a common pathway, *Molecular cell*, 7 (2001) 249–262. [PubMed: 11239454]
- [43]. Breast Cancer Linkage C, Cancer risks in BRCA2 mutation carriers, *J Natl Cancer Inst*, 91 (1999) 1310–1316. [PubMed: 10433620]
- [44]. Yin J, Liu H, Liu Z, Wang LE, Chen WV, Zhu D, Amos CI, Fang S, Lee JE, Wei Q, Genetic variants in fanconi anemia pathway genes BRCA2 and FANCA predict melanoma survival, *The Journal of investigative dermatology*, 135 (2015) 542–550. [PubMed: 25243787]
- [45]. Pierce BL, Ahsan H, Genome-wide “pleiotropy scan” identifies HNF1A region as a novel pancreatic cancer susceptibility locus, *Cancer research*, 71 (2011) 4352–4358. [PubMed: 21498636]
- [46]. Kong B, Wu W, Valkovska N, Jager C, Hong X, Nitsche U, Friess H, Esposito I, Erkan M, Kleeff J, Michalski CW, A common genetic variation of melanoma inhibitory activity-2 labels a subtype of pancreatic adenocarcinoma with high endoplasmic reticulum stress levels, *Sci Rep*, 5 (2015) 8109. [PubMed: 25657029]
- [47]. El Fitori J, Kleeff J, Giese NA, Guweidhi A, Bosserhoff AK, Buchler MW, Friess H, Melanoma Inhibitory Activity (MIA) increases the invasiveness of pancreatic cancer cells, *Cancer Cell Int*, 5 (2005) 3. [PubMed: 15710044]

- [48]. Heikkinen K, Rapakko K, Karppinen SM, Erkkö H, Knuutila S, Lundan T, Mannermaa A, Borresen-Dale AL, Borg A, Barkardottir RB, Petrini J, Winqvist R, RAD50 and NBS1 are breast cancer susceptibility genes associated with genomic instability, *Carcinogenesis*, 27 (2006) 1593–1599. [PubMed: 16474176]
- [49]. Heikkinen K, Karppinen SM, Soini Y, Mäkinen M, Winqvist R, Mutation screening of Mre11 complex genes: indication of RAD50 involvement in breast and ovarian cancer susceptibility, *Journal of medical genetics*, 40 (2003) e131. [PubMed: 14684699]
- [50]. Damiola F, Pertesi M, Oliver J, Le Calvez-Kelm F, Voegelé C, Young EL, Robinot N, Forey N, Durand G, Vallee MP, Tao K, Roane TC, Williams GJ, Hopper JL, Southey MC, Andrulis IL, John EM, Goldgar DE, Lesueur F, Tavtigian SV, Rare key functional domain missense substitutions in MRE11A, RAD50, and NBN contribute to breast cancer susceptibility: results from a Breast Cancer Family Registry case-control mutation-screening study, *Breast Cancer Res*, 16 (2014) R58. [PubMed: 24894818]

Highlights

- Rare, protein-coding variants in *MC1R*, *MITF*, *BRCA2*, and *MTAP* confer between a 1.4-fold to 6-fold increase in melanoma risk.
- Susceptibility variants within *MC1R* are associated with ultraviolet (UV) radiation mutational signatures in triple-wildtype tumors.
- Susceptibility variants in *MC1R* and *MTAP* are associated with earlier age of melanoma diagnosis.

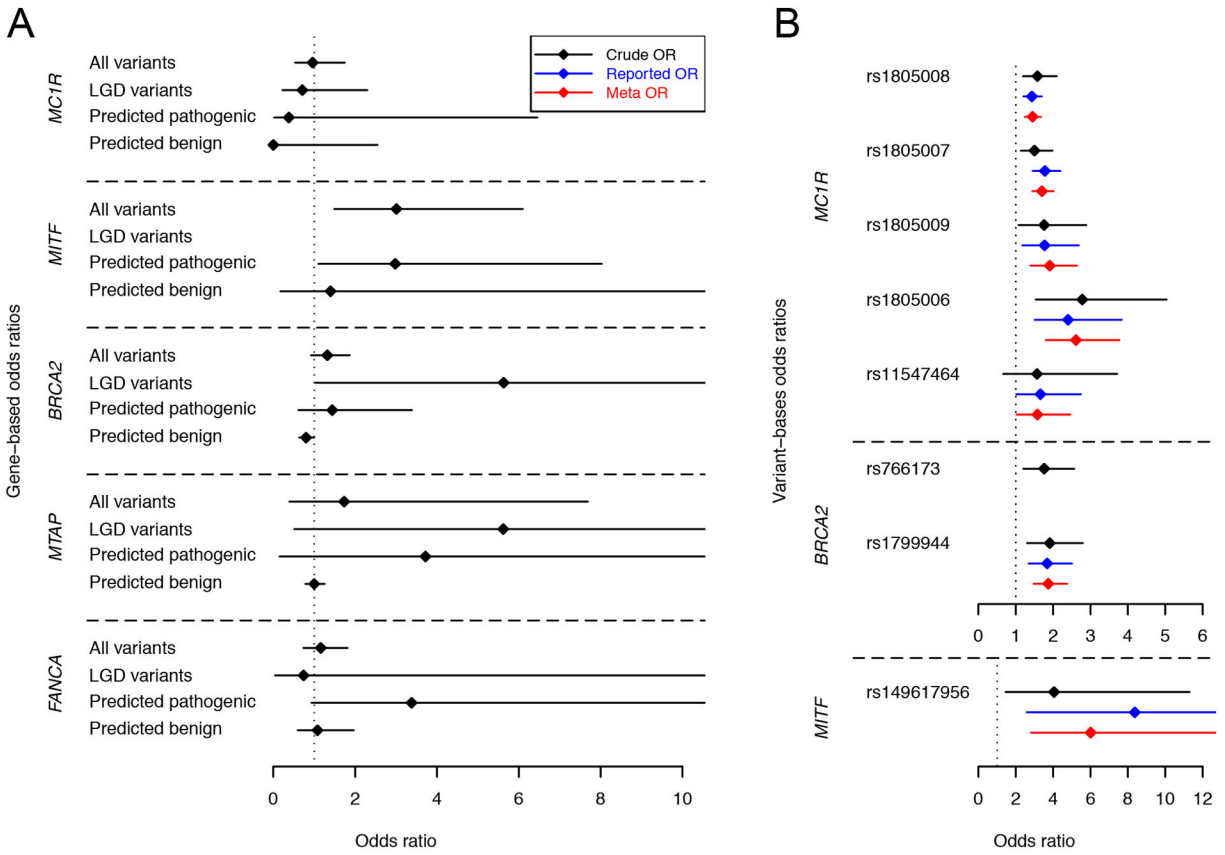


Figure 1. Odds ratio estimates for melanoma susceptibility genes and variants.

We estimated the ORs of four established and replicated melanoma susceptibility genes with TCGA-SKCM cases and NDAR controls. (A) depicts the gene-based ORs for four particular classes of variants: (1) all variants in a given gene, (2) only LGD variants, (3) predicted pathogenic missense variants, and (4) predicted benign missense variants. (B) depicts the variant-based ORs of eight variants in *MC1R*, *BRCA2* and *MITF*. See Table 2 for OR estimate details

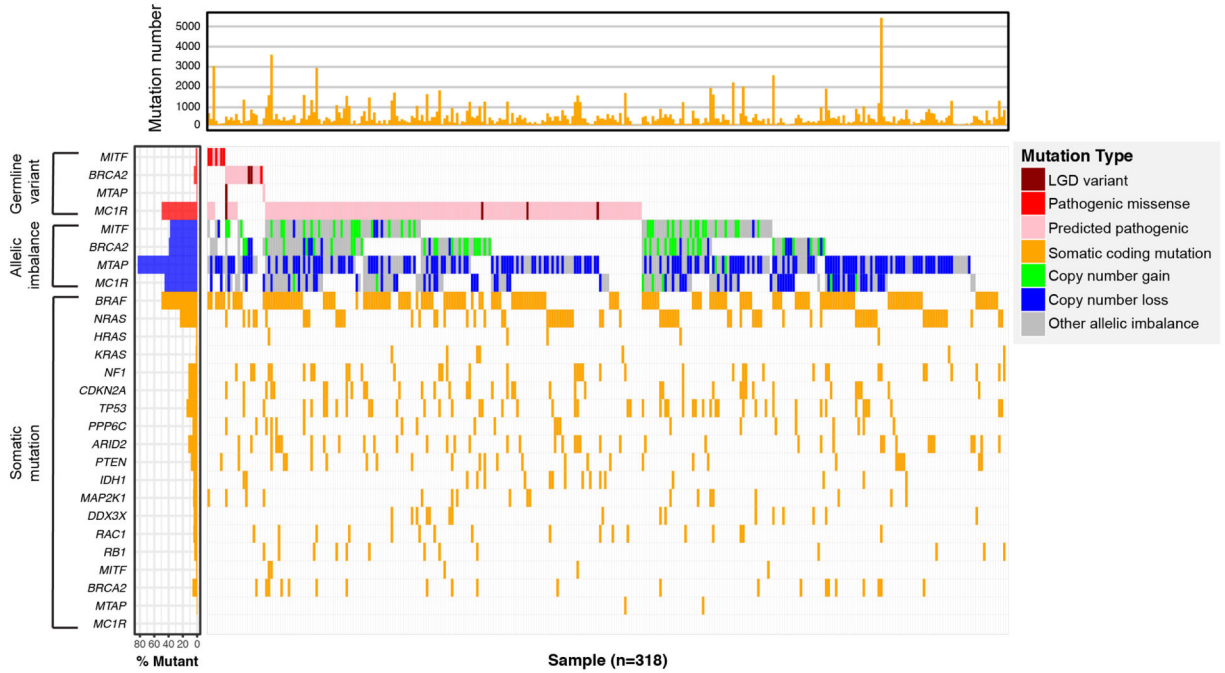


Figure 2. Landscape of germline variants and somatic events in select genes.

The main panel (middle) indicated the germline variants and somatic allelic imbalance events of four genes (i.e. *MC1R*, *MITF*, *BRCA2*, and *MTAP*), and somatic mutations of melanoma driver genes reported in the TCGA paper [14]. The top panel indicated the total number of somatic mutations for each individual. The left panel indicated the mutation burdens of each gene in germline and somatic samples. Germline variants include: likely gene-disrupting (LGD) variants (i.e. stop gained, stop lost, splicing region variants and frame-shift INDELS), and non-LGD variants but known as pathogenic ones according to the ClinVar database, and non-LGD variants with pathogenic potential according to CASM score from VAAST 2. Somatic allelic imbalance events include: copy number gain with segment mean $> \log_2(2.5/2)$, copy number loss with segment mean $< \log_2(1.5/2)$, and other allelic imbalance events provided by hapLOH only.

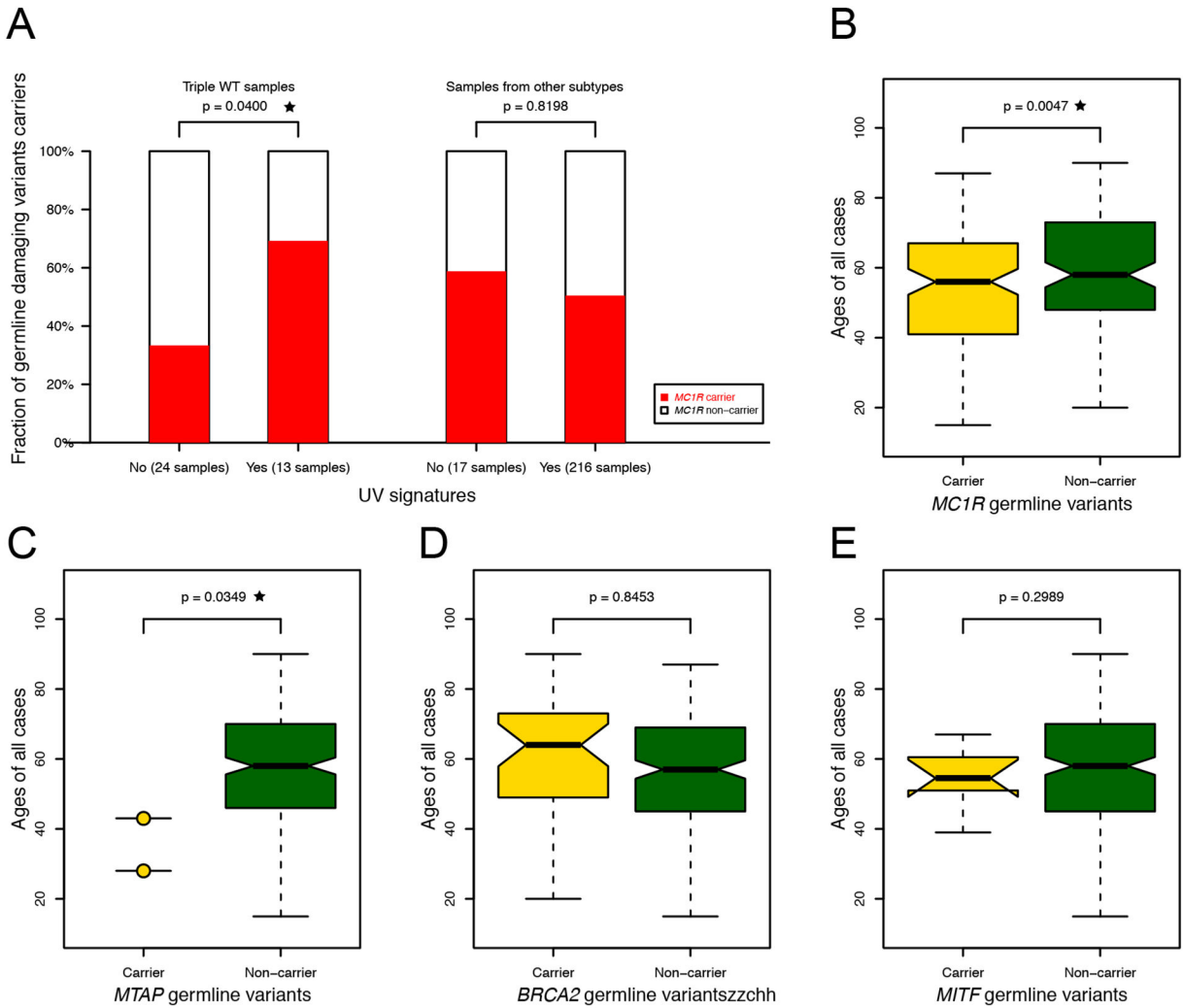


Figure 3. Associations of melanoma susceptibility genes with UV mutational signatures, sample subtypes, and ages at diagnosis.

(A) depicts the proportion of *MC1R* carriers among samples with and without UV signatures among triple WT tumors and other tumor subtypes. Boxplots in (B-E) depict the distribution of age at diagnosis among carriers and non-carriers.

Table 1.

VAAST 2 analysis results of selected genes with reported melanoma risk.

Gene	All coding variants	Excluding LGD variants	Comments
<i>MC1R</i>	0.0014	0.0006	
<i>MITF</i>	0.0165	0.0159	Reported melanoma susceptibility genes
<i>BRCA2</i>	0.0206	0.0448	
<i>MTAP</i>	0.0393	0.2870	
<i>FANCA</i>	0.0020	0.0005	Reported susceptibility genes in other cancers
<i>HNF1A</i>	0.0332	0.0156	

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2.

Odds ratio estimates for coding variants in MC1R, MITF, and BRCA2 with five or more allele counts among cases and controls.

Gene	Variant	Allele frequency		Case		Control		Odds ratio			Meta analysis*
		SKCM	NDAR	Carrier	Non-carrier	Carrier	Non-carrier	Crude	Adjusted	Reported	
	rs1805008	11.02%	7.55%	68	254	511	3024	1.58 (1.19–2.10)	1.49 (1.11–1.99)	1.43 (1.20–1.70)	1.45 (1.24–1.68)
	rs1805007	10.90%	7.70%	66	255	530	3069	1.50 (1.13–1.99)	1.56 (1.16–2.10)	1.78 (1.45–2.20)	1.70 (1.44–2.02)
<i>MC1R</i>	rs1805009	2.95%	1.73%	19	303	124	3482	1.76 (1.07–2.89)	2.15 (1.30–3.57)	1.77 (1.17–2.69)	1.91 (1.39–2.64)
	rs1805006	2.17%	0.80%	14	308	58	3548	2.78 (1.53–5.04)	3.00 (1.63–5.54)	2.40 (1.50–3.84)	2.61 (1.80–3.78)
	rs11547464	0.93%	0.61%	6	316	43	3562	1.57 (0.66–3.72)	1.36 (0.55–3.38)	1.66 (1.01–2.75)	1.58 (1.02–2.46)
<i>BRCA2</i>	rs766173	5.28%	3.22%	34	288	227	3375	1.76 (1.20–2.57)	1.81 (1.22–2.69)	na	na
	rs1799944	5.28%	3.04%	34	288	204	3298	1.91 (1.30–2.80)	2.03 (1.36–3.02)	1.84 (1.34–2.51)	1.87 (1.47–2.38)
<i>MITF</i>	rs149617956	0.78%	0.19%	5	317	14	3593	4.05 (1.45–11.31)	4.48 (1.57–12.78)	8.37 (2.58–23.80)	6.01 (2.81–12.89)

* The OR in meta-analysis was calculated by combining the adjusted OR and reported OR from literature.

Table 3.

Odds ratio estimates and confidence intervals for genes associated with melanoma cancer.

Variant type	Gene	Number of cases		Number of controls		Crude OR	Adjusted OR
		Carrier	Non-carrier	Carrier	Non-carrier		
All variant	<i>MC1R</i>	12	310	140	3467	0.96 (0.53–1.75)	1.18 (0.64–2.18)
	<i>MITF</i>	10	312	38	3569	*3.01 (1.49–6.10)	*3.89 (1.88–8.06)
	<i>BRCA2</i>	39	283	342	3265	1.32 (0.92–1.87)	1.29 (0.90–1.86)
	<i>MTAP</i>	2	320	13	3594	1.73 (0.39–7.69)	2.21 (0.48–10.25)
	<i>FANCA</i>	22	300	215	3392	1.16 (0.73–1.82)	1.33 (0.84–2.13)
	<i>MC1R</i>	3	319	47	3560	0.71 (0.22–2.30)	0.83 (0.25–2.71)
LGD variant	<i>MITF</i>	0	322	0	3607	NA	NA
	<i>BRCA2</i>	2	320	4	3603	*5.63 (1.03–30.86)	3.40 (0.52–22.37)
	<i>MTAP</i>	1	321	2	3605	5.62 (0.51–62.1)	0.00 (0.00-INF)
	<i>FANCA</i>	0	322	7	3600	0.74 (0.04–13.06)	NA
	<i>MC1R</i>	0	322	14	3593	0.38 (0.02–6.46)	NA
	<i>MITF</i>	5	317	19	3588	*2.98 (1.10–8.03)	*3.55 (1.29–9.81)
Non-LGD variants	Predicted pathogenic <i>BRCA2</i>	6	316	47	3560	1.44 (0.61–3.39)	1.55 (0.65–3.73)
	<i>MTAP</i>	0	322	1	3606	3.72 (0.15–91.70)	NA
	<i>FANCA</i>	3	319	10	3597	3.38 (0.93–12.35)	3.69 (0.97–14.00)
	<i>MC1R</i>	0	322	18	3589	0.00 (0.00–2.55)	NA
	<i>MITF</i>	1	321	8	3599	1.40 (0.17–11.24)	0.00 (0.00-INF)
	Predicted benign <i>BRCA2</i>	176	146	2171	1436	0.80 (0.63–1.00)	0.89 (0.70–1.13)
<i>MTAP</i>	210	112	2356	1251	1.00 (0.78–1.26)	0.94 (0.73–1.20)	
<i>FANCA</i>	12	310	125	3482	1.08 (0.59–1.97)	1.28 (0.68–2.38)	

*OR was significantly greater than 1.0 with p value < 0.05.