Published in final edited form as:

Pigment Cell Melanoma Res. 2013 May ; 26(3): . doi:10.1111/pcmr.12070.

Variants in *melanocortin 1 receptor* gene contribute to risk of melanoma – a direct sequencing analysis in a Texas population

Xiaoxiang Guan^{1,2,*}, Jiangong Niu^{2,*}, Zhengshen Liu^{2,*}, Li-E Wang², Christopher I. Amos², Jeffrey E. Lee³, Jeffrey E. Gershenwald³, Elizabeth A. Grimm⁴, and Qingyi Wei^{2,**}

¹Department of Medical Oncology, Jinling Hospital, Nanjing University School of Medicine, Nanjing, 210002, China

²Department of Epidemiology, The University of Texas M. D. Anderson Cancer Center, Houston, TX 77030, USA

³Department of Surgical Oncology, The University of Texas M. D. Anderson Cancer Center, Houston, TX 77030, USA

⁴Department of Experimental Therapeutics, The University of Texas M. D. Anderson Cancer Center, Houston, TX 77030, USA

Abstract

Summary—In this study, we directly sequenced the *Melanocortin 1 Receptor (MC1R)* gene in 2,212 individuals to detect all variants and assessed their associations with cutaneous melanoma (CM) risk in a hospital-based study of 1,106 CM patients and 1,106 control subjects. Of 61 *MC1R* variants identified, 16 rare variants have not been previously reported by others; three *MC1R* variants were associated with a significant CM risk [c.451C>T (OR = 1.78, 95% CI = 1.44–2.20), c.478C>T (OR = 1.31, 95% CI = 1.05–1.63), and c.880G>C (OR = 1.69, 95% CI = 1.15–2.48)]; and two with borderline CM risk [c.942A>G (OR = 1.23, 95% CI = 1.00–1.51, and c.274G>A (OR = 1.23, 95% CI = 0.99–1.53)] under a dominant model. When combined these five *MC1R* variants for cumulative effect analysis, we found that subjects with an increased number of variant genotypes from any of these five variants had significantly increased risk of CM with ORs of 1.68 (95% CI = 1.39–2.04), 1.61 (95% CI = 1.27–2.04), and 2.64 (95% CI = 1.72–4.05) for one, two, and three or more variant genotypes, respectively (trend test: $P < 0.001$). Further haplotype and diplotype analyses based on the above-mentioned five SNPs suggested that the c.451T allele contributed to the high risk of CM and that the five variants may have joint effects on the risk of CM. Additional analysis suggests that the three most significant SNPs may be the molecular mechanisms underlying the known risk factors of the colors of the eyes, skin and hair in this study population. In conclusion, our study provided confirmatory evidence that both common and rare variants in the *MC1R* coding region may be biomarkers for susceptibility to CM in US populations.

Keywords

melanocortin 1 receptor gene; direct sequencing; interaction; melanoma; case-control

^{**}To whom correspondence should be addressed: Qingyi Wei, Department of Epidemiology, Box 189, The University of Texas MD Anderson Cancer Center, 1515 Holcombe Boulevard, Houston, TX 77030; Phone: 713-792-3020; Fax: 713-563-0999; gwei@mdanderson.org..

^{*}These authors contributed equally to this work.

Supplemental Materials

Supplemental Data contains one table (Table S1).

Table S1. Associations of genetic variant allele with skin, hair and eye phenotypes in control subjects.

None of the authors have any current conflicts of interest.

Introduction

Cutaneous melanoma (CM) is one of the most common cancer types in humans. It is estimated that there are nearly one million melanoma survivors living in the United States and that an additional 76,250 individuals will be diagnosed in 2012 (Siegel et al., 2012). Melanoma incidence rates have been increasing for at least 30 years (Siegel et al., 2012). Although solar UV exposure is a well-known risk factor for CM, most UV-exposed individuals do not develop CM (Marks, 1995; Moan and Dahlback, 1992), suggesting that genetic susceptibility modifies CM risk. For example, a family history of CM is strongly associated with CM risk (Kefford et al., 1999), which supports the hypothesis that genetic factors play a role in the etiology of CM.

In a recent genome-wide association (GWA) study and the subsequent replications, three loci located in the *MC1R*, *TYR* genes and the region adjacent to *MTAP* and *CDKN2A* were identified to be risk factors for CM (Bishop et al., 2009). However, this study did not cover all the potential genetic risk regions of susceptibility genes. More recently, three additional GWASs identified additional risk loci for CM, including a few SNPs in *MC1R* (Amos et al., 2011; Barrett et al., 2011; Macgregor et al., 2011). Because there are a number of SNPs in *MC1R* that contains only one exon, only a few of which were included in the published GWASs, it is necessary to sequence the whole gene to provide a holistic view of the contribution of all *MC1R* SNPs to risk of CM (Meng et al., 2012).

MC1R relays intracellular signals to regulate the melanin biosynthetic pathway, therefore determining pigmentation of the skin and hair (Abdel-Malek et al., 1999), which can be further stimulated by UV irradiation (Abdel-Malek et al., 2008). *MC1R* is highly polymorphic in white populations. In one study, over 80 variants were identified from the 954-bp coding nucleotides of *MC1R* in 3,301 CM cases (Kanetsky et al., 2006). However, their conclusions on risk assessment of primary CM were relatively undetermined, because the case-series study did not include CM-free controls; therefore, the risk predictions of primary CM in carriers of *MC1R* variants may have been attenuated. To date, there is still lack of large and thorough case-control studies of *MC1R* variants in CM risk in US populations (Kanetsky et al., 2010). In the present study, we reported a detailed profile of newly identified *MC1R* variants by direct sequencing of its entire coding region and investigated their associations with CM risk in a Texas non-Hispanic white population of 1,106 CM patients and 1,106 cancer-free control subjects frequency matched by age and sex.

Results

Distributions of Demographic Characteristics and Risk Factors

As shown in Table 1, the 1,106 CM patients and 1,106 control subjects of non-Hispanic whites appeared to be adequately matched by age and sex, with a mean age of 51.2 ± 13.3 years (range 18–92) for patients and 51.8 ± 13.0 years (range 18–87) for control subjects ($P = 0.297$); although no difference statistically, there were more young (age ≤ 55) and male subjects among both patients and controls (59.4% and 59.1% for age ≤ 55 years, respectively, and 60.7% and 59.9% for males, respectively) ($P = 0.696$ for sex). The known risk factors in this study population included eye color (blue, hazel or green vs. brown or gray), hair color (red or blonde vs. dark), skin color (fair vs. dark), lifetime sunburn with blistering, freckling in the sun as a child, low tanning ability, prolonged exposure in the sun, and the presence of moles or dysplastic nevi (Table 1); however, family history of cancer was not a risk factor ($P = 0.401$, age and sex adjusted OR = 1.10, 95% CI = 0.92–1.31) in

this study population, which is likely due to a hospital-based study with possible over matching on the family history.

Identification of Variants in the Coding Regions of *MC1R*

By using direct sequencing of the entire coding region of *MC1R*, we identified a total of 61 variants or SNPs in this study population, including 42 non-synonymous variants, one insertion, and 14 synonymous variants (Table 2). Because the sequencing primers covered parts of the 5' and 3' un-translated regions (UTRs), we also identified four variants in the 5' and 3' UTRs (Table 2); of the 61 variants, 16 rare (i.e., identified in only one or two subjects) variants had not been previously reported by others, to the best of our knowledge; six of these 16 newly identified variants appeared only in control subjects, nine only in case patients, and one variant in both case patients and control subjects. Five non-synonymous SNPs [c.178G>T (V60L, rs1805005), c.274G>A (V92M, rs2228479), c.451C>T (R151C, rs1805007), c.478C>T (R160W, rs1805008) and c.488G>A (R163Q, rs885479)] and one synonymous SNP [c.942A>G (T314T, rs2228478)] were considered common, because their minor allele frequencies (MAF) were >5.0%; among others, one non-synonymous SNP [c.880G>C (D294H, rs1805009)] had a MAF of 2.5%, and one [c.252C>A (D84E, rs1805006)] had a MAF of 1.0%, with all other variants having a MAF of <1%. As shown in Table 2, only three SNPs (i.e., c.451C>T, c.880G>C, and c.464T>C) had an allele frequency distribution that was significant different between cases and controls ($P = 0.001, 0.027$, and 0.044 , respectively). Also, we identified an insertion c.86_87insA (p.Asn29Lys) in 31 of the 2,212 individuals (1.46%) (Table 2), which was similar to the previously reported frequency (42 of 3,301 individuals, 1.27%) (Kanetsky et al., 2006).

Association between Common *MC1R* Variants and CM Risk

Because all *MC1R* variants we identified are located in a narrow region of chromosome 16, we first performed a linkage disequilibrium (LD) analysis for the 16 SNPs that had a MAF of >0.001 in control subjects (Figure 1A). We found that c.274G>A and c.942A>G were in high LD ($D' = 0.988$, $LOD = 187.41$, $r^2 = 0.819$, $P < 0.001$). Figure 1B shows the LD among the eight SNPs (c.178G>T, c.252C>A, c.274G>A, c.451C>T, c.478C>T, c.488G>A, c.880G>C, and c.942A>G) that had a MAF >1%, which were further evaluated for their associations with CM risk.

As shown in Table 3, the c.451C>T SNP had the highest risk for CM (OR = 1.65, 95% CI = 1.33–2.05 for the CT genotype and OR = 5.06, 95% CI = 2.21–11.6 for the TT genotype). The combined c.451 CT+TT, c.478 CT+TT, and c.880 CG+CC genotypes were associated with significantly increased risk for CM (OR = 1.78, 95% CI = 1.44–2.20; OR = 1.31, 95% CI = 1.05–1.63; and OR = 1.69, 95% CI = 1.15–2.48, respectively) in a dominant model, whereas the c.478 CT and c.880 CG genotypes were associated with significantly increased risk of CM (OR = 1.32, 95% CI = 1.05–1.65; OR = 1.70, 95% CI = 1.15–2.51, respectively) in a co-dominant model, compared with the corresponding common genotypes. In addition, a borderline risk of CM was observed with the c.274G>A SNP (OR = 1.26, 95% CI = 1.01–1.57 for the GA genotype and OR = 1.23, 95% CI = 0.99–1.53 for GA+AA genotypes, respectively) and the c.942A>G SNP (OR = 1.23, 95% CI = 1.00–1.52 for the GA genotype and OR = 1.23, 95% CI = 1.00–1.51 for GA+GG genotypes) (Table 3), but c.942A>G was in LD with c.274G>A. However, the other three (c.178G>T, c.252C>A, and c.488G>A) SNPs were not associated with CM risk (Table 3). Taken together, of the eight *MC1R* SNPs with a MAF >1%, three (c.451C>T, c.478C>T, and c.880G>C) were significantly associated with increased CM risk, and two (c.274G>A and c.942A>G) in LD showed a borderline increased risk of CM. All risk estimates were adjusted for age and sex. We did not adjust for the known risk factors, because it is known that *MC1R* variants are associated with some of these risk factors, which may result in over-adjustment.

We further investigated cumulative effects of the five *MC1R* SNPs (c.274G>A, c.451C>T, c.478C>T, c.880G>C, and c.942A>G) that were individually found to be associated with either a significant or a borderline increased risk of CM. For each subject, the number of such combined variant genotypes was added up to create one single variable. As a result, subjects with an increased number of variant genotypes from any of these SNPs had significantly increased risk of CM with ORs of 1.68 (95% CI = 1.39–2.04), 1.61 (95% CI = 1.27–2.04), and 2.64 (95% CI = 1.72–4.05) for one, two, and three or more variant genotypes, respectively (trend test: $P < 0.001$), suggesting a cumulative effect of *MC1R* variant genotypes on CM risk (Table 4).

Association between *MC1R* Haplotypes/Diplotypes and CM Risk

We also performed haplotype and diplotype analyses based on the above-mentioned five SNPs (c.274G>A, c.451C>T, c.478C>T, c.880G>C, and c.942A>G), again assuming a dominant model (Table 5). Among all 19 possible haplotypes we inferred from the observed genotypes, four had frequencies of >5%, and the overall distribution of *MC1R* haplotypes between cases and controls was statistically significant (global test: $P < 0.001$). The most frequent haplotype G-C-C-G-A (in the order of c.274-c.451-c.478-c.880-c.942) (60.0% in the cases and 70.8% in the controls) was used as the reference, whereas all haplotypes <5% were combined as one group of rare haplotypes. All variant haplotypes, including the group of rare haplotypes, were associated with significantly increased risk for CM (OR = 1.94, 95% CI = 1.60–2.36 for G-T-C-G-A; OR = 1.36, 95% CI = 1.11–1.67 for A-C-C-G-G; OR = 1.43, 95% CI = 1.16–1.76 for G-C-T-G-A; and OR = 1.70, 95% CI = 1.27–2.28 for the group of rare haplotypes), compared with the reference (Table 5). Two of the common haplotypes (frequency >5%) had only one variant allele (T of c.451 in G-T-C-G-A and T of c.478 in G-C-T-G-A), and another haplotype had two variant alleles (A of c.274 and G of c.942 in A-C-C-G-G), representing the variant alleles from two SNPs with high LD. Consistent with the individual association analysis (Table 3), the haplotype G-T-C-G-A was associated with the highest risk, as were the c.451T variant genotypes, whereas the haplotype A-C-C-G-G was associated with the lowest increased risk, as were either c.274A or c.942G variant genotypes, suggesting only those haplotypes that contain these variant alleles contributed to the CM risk independently.

According to the haplotype data, we inferred 24 possible diplotypes, four of which had a frequency >5%. The overall distribution of *MC1R* diplotypes between cases and controls was statistically significant ($P < 0.001$). The most frequent diplotype G-C-C-G-A/G-C-C-G-A (in the order of c.274-c.451-c.478-c.880-c.942) (37.4% in the cases and 50.6% in the controls) was used as the reference, and all the diplotypes with frequencies <5% were grouped as “others”, representing the group of rare diplotypes. As shown in Table 5, all the variant diplotypes, including the group of rare diplotypes, showed significantly increased risk for CM (OR = 1.69, 95% CI = 1.30–2.19 for G-C-C-G-A/G-T-C-G-A; OR = 1.32, 95% CI = 1.01–1.74 for A-C-C-G-G/G-C-C-G-A; OR = 1.56, 95% CI = 1.18–2.06 for G-C-C-G-A/G-C-T-G-A; and OR = 2.17, 95% CI = 1.72–2.74 for the group of rare diplotypes), compared with the reference. Again, the diplotype G-C-C-G-A/G-T-C-G-A containing c.451T variant allele was associated with the highest risk; the diplotype A-C-C-G-G/G-C-C-G-A containing both c.274A and c.942G variant alleles had the lowest increased risk of CM, compared with the reference after adjustment for age and sex (Table 5). Taken together, these data consistently suggested that the c.451T allele was the major allele that contributed to the high risk of CM and that the five variants may have joint effects on the risk of CM.

To further evaluate the relationship between the observed *MC1R* variants and constitutional risk factors (i.e., colors of the eyes, skin and hair) for CM, the associations between the variants genotypes of the most three significant, relatively common SNPs (i.e., c.478C>T, c.

451C>T and c.880G>C) and the risk factors are presented in Supplemental Table S1. Indeed, c.478C>T was likely to predict colors of the eyes, skin and hair, whereas c.451C>T and c.880G>C were likely to predict colors of the skin and hair. These data indicate that these variants are genetic factors underlying these risk factors, or host phenotypes, for CM. However, we did not have enough statistical power to detect any gene-environment and gene-gene interactions (data not shown).

Discussion

In the present case-control analysis, we found that c.451T, c.478T, and c.880C variant genotypes were associated with increased risk of CM and that c.274A and c.942A variant genotypes were associated with a borderline risk of CM in this study population. Among all *MC1R* variants we analyzed, c.451C>T showed the most significant association with the risk of CM, not only in its individual association but also in its contribution to the haplotypes/diplotypes, because the haplotypes and diplotypes containing the c.451 T variant were associated with a higher risk of CM than did other haplotypes and diplotypes. Considering the five variants that were associated with significant or borderline risks of CM, there was a significant trend that individuals carrying more variant genotypes had a higher risk of CM. Moreover, all common haplotypes and diplotypes we inferred were associated with a significantly increased risk of CM. Taken together, these results suggest that there may be a joint effect among these *MC1R* variants. These results were consistent with the findings from previous studies in Australian and European populations (Matichard et al., 2004; Palmer et al., 2000; Sturm, 2002), although slightly different from the results of a recent study, in which there was a weak risk association between higher frequency of high degree primary CM and c.451T variant genotypes and a borderline association for c.478T and c.880C variant genotypes (Kanetsky et al., 2006). Furthermore, no significant association with the c.252A variant genotypes was found in our present and others studies (Ichii-Jones et al., 1998; Palmer et al., 2000), which is inconsistent with two other publications (R et al., 2008; Valverde et al., 1996). There may be ethnic and geographic differences in the etiology of melanoma in addition to CM risk associated with this variant.

Considering that c.451C>T, c.478C>T, and c.880G>C represented a group of red hair color (RHC) variants, we examined associations of these three variants with eye color, skin color, hair color, and three variants were strongly associated with fair skin and red hair (Table S1). Given *MC1R* is highly polymorphic and both RHC alleles and related phenotypes are known to be associated with CM risk, gene-gene and gene-environment interactions between *MC1R* variants and risk factors may exist. In the present study, although we did not have enough study power to detect gene-gene and gene-environment interactions, we did find that the effects of each single SNP can accumulate, which suggests that the future risk prediction model should take into account all these variants.

The strength of the present study is that we evaluated the role of *MC1R* variants in CM risk in a Texas population by a direct sequencing analysis that is supplementary to our recent GWAS of the same study population (Amos et al., 2011); however, there are several limitations in the present study. First, because of limited study power, gene-gene and gene-environment interactions could not be addressed. To better address the interaction between risk factors and *MC1R* variants, larger population-based, preferably prospective, studies and rigorous mechanistic studies are needed. Second, in our data analyses, we calculated ORs with the adjustments only for age and sex but not for other CM risk factors. This is because that some known risk factors for CM were in correlations with multiple *MC1R* variants that were also related to CM risk. Adjustment for such correlations among these risk factors and *MC1R* variant genotypes would lead to over-adjustment. Because the subjects were born with the *MC1R* variants, these variants may be more objective for us to predict CM risk.

Furthermore, cancer is a disease involving multiple genes/SNPs, and therefore the combined analysis of multiple gene or haplotype-based approach might be more powerful than the analysis of single gene. It is also likely that our results could be biased owing to possible selection biases in hospital-based case-control studies. Therefore larger and well-design prospective studies are warranted to confirm our findings.

Materials and Methods

Study Subjects

The recruitment of this study population was described previously (Li et al., 2008). Briefly, the final analysis included 1,106 non-Hispanic white patients with newly diagnosed CM seen at The University of Texas M. D. Anderson Cancer Center between April 1994 and August 2008 and 1,106 age (± 5 years), sex and ethnicity frequency-matched, cancer-free controls recruited during the same period. The cases had histopathologically confirmed primary melanoma, and the majority of patients had early-stage melanoma (stages 0 [melanoma in situ], I, II, III, and IV in 9.7%, 59.1%, 14.6%, 16.1%, and 0.5%, respectively). The control subjects were recruited from among cancer-free visitors accompanying patients on outpatient clinic visits to M. D. Anderson Cancer Center and were not genetically related to the patients. Having provided a written informed consent, all subjects were interviewed in person to provide information about age, sex, ethnicity, and known risk factors, including eye color, hair color, skin color, freckling in the sun as a child, presence of moles or dysplastic nevi, tanning ability, lifetime number of sunburns with blistering, and Fitzpatrick classification of sun-reactive skin type (Fitzpatrick, 1988). Each subject donated a one-time 30-mL sample of venous blood that was used for biomarker assays, including DNA extraction for genotyping. The research protocol was approved by the institutional review board at M. D. Anderson.

Genotyping of *MC1R*

The genomic DNA from study subjects were amplified by polymerase chain reaction (PCR) with two pairs of previously reported primers: pair 1: (forward) 5'-CAGCACCATGAACCTAAGCAGGACACCTG-3', (reverse) 5'-CCAGCATAGCCAGGAAGAAGACCACGA-3' and pair 2: (forward) 5'-TGGGTGGCCAGTGTCTCTTCAGC-3', (reverse) 5'-AAGGGTCCGCGCTTCAACACTTTCAGAG-3' (Kanetsky et al., 2006). The two PCR products cover the region from the 93 bp upstream of the start codon to the 161 bp downstream of the stop codon of *MC1R*, with a 97-bp overlap. The PCR reactions were performed by 35 cycles with an annealing temperature of 64°C. Following PCR, the products were cleaned with the ExoSAP-IT® reagent, according to the manufacturer's protocol (USB Corporation, Cleveland, OH) and sequenced using an ABI 3730 sequencer (Applied Biosystems, Foster City, CA) at the DNA Core Laboratory at M. D. Anderson. The sequencing chromatograms were analyzed using the Variant Reporter software (Applied Biosystems) that helped in determining the genotypes. We also visually evaluated all chromatograms that could not be read by the Variant Reporter (<1%) to identify the genotypes. All genotypes were incorporated into a Microsoft Excel data file used for further statistical analysis.

Statistical Analysis and Data Coding

All data were analyzed using the SAS statistical software program (SAS/STAT version 9.1.3; SAS Institute Inc., Cary, NC). The differences in the distributions of demographic variables and known risk factors for CM obtained from the self-assessed screening questionnaire between case patients and control subjects were evaluated with chi-square tests. Skin color was self-reported on a scale from 1 (light) to 10 (dark) instead of using

descriptive categories and was re-coded for analysis into three evenly distributed groups by the tertile of the total subjects' distribution as fair, brown, and dark skin, respectively. Crude and age and sex-adjusted odds ratios (ORs) and 95% confidence intervals (CIs) for each of the known risk factors and *MC1R* variants were obtained from unconditional univariate and multivariable logistic regression analyses, respectively, to evaluate their associations with CM. risk. For genotype analyses, the most common homozygous genotype was used as the reference group, and the heterozygous and rare homozygous genotypes were compared with the reference separately, under the assumption of an additive model. Likewise, heterozygous and rare homozygous genotypes were combined together and compared with the reference group, under the assumption of a dominant model. To determine the effect of combined variant genotypes on the risk of CM, we assigned each individual with a number representing the total number of variant genotypes the individual had from the five *MC1R* variants that had either an increased or a borderline increased risk for CM. Standard unconditional logistic regression analyses and the trend test were used to test for the associations and the trend in risk associations accordingly.

The associations with minor variant genotypes were also stratified by age, sex, and the known risk factors (i.e., skin color, eye color, hair color, tanning ability, lifetime number of sunburns with blistering, freckling in the sun as a child, presence of moles and dysplastic nevi, and family history of cancer). Homogeneity of ORs between different strata was tested using a Cochran-Mantel-Haenszel test, followed by analyses of gene-environment interactions as previously described (Hong et al., 2005). Potential multiplicative and additive interactions between gene and environment were analyzed with multivariable logistic regression models with and without adjustment for age and sex, as previously described (Brennan, 2002). For these analyses, the main effects of each variable and an interaction effect were jointly modeled.

Haplotypes for the selected variants were reconstructed using the PHASE program (Stephens and Scheet, 2005; Stephens et al., 2001). For each individual, the probability of having a particular haplotype pair was estimated, and the haplotype pair with the highest estimated probability was assigned to the individual. Those haplotypes having a frequency less than 5% were combined into one group of rare haplotypes, and the most frequent haplotype in control subjects was used as the reference group for comparisons, and diplotype analysis was performed in a similar manner, with the assignment of a pair of haplotypes as a diplotype for each individual (Sprague et al., 2007). Pearson's chi-square test or global test was used to test for the case-control differences in the frequencies of all haplotypes and diplotypes (Jiao et al., 2007). All tests were two-sided, with a P value of <0.05 considered statistically significant.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by National Cancer Institute R01 CA100264 (to Q.W.), P50 CA093459 (to E.A.G.), R25T CA57730 (to Robert M. Chamberlain and Shine Chang), and the McCarthy Fellowship in Skin Cancer Research (to J.E.L.). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Cancer Institute or the National Institutes of Health. We thank Margaret Lung and Kathryn L. Tipton for their assistance in recruiting the subjects and gathering the questionnaire information; Yawei Qiao, Jianzhong He, Min Zhao for laboratory assistance, and Michael Worley for scientific editing.

Abbreviations

CI	confidence intervals
CM	cutaneous melanoma
MC1R	Melanocortin-1 receptor
OR	odds ratio
PCR	polymerase chain reaction
SNP	single nucleotide polymorphism

References

- Abdel-Malek Z, Suzuki I, Tada A, Im S, Akcali C. The melanocortin-1 receptor and human pigmentation. *Ann N Y Acad Sci.* 1999; 885:117–133. [PubMed: 10816645]
- Abdel-Malek ZA, Knittel J, Kadekaro AL, Swope VB, Starner R. The melanocortin 1 receptor and the UV response of human melanocytes--a shift in paradigm. *Photochem Photobiol.* 2008; 84:501–508. [PubMed: 18282187]
- Amos CI, Wang LE, Lee JE, Gershenwald JE, Chen WV, Fang S, Kosoy R, Zhang M, Qureshi AA, Vattathil S, et al. Genome-wide association study identifies novel loci predisposing to cutaneous melanoma. *Hum Mol Genet.* 2011; 20:5012–5023. [PubMed: 21926416]
- Barrett JH, Iles MM, Harland M, Taylor JC, Aitken JF, Andresen PA, Akslen LA, Armstrong BK, Avril MF, Azizi E, et al. Genome-wide association study identifies three new melanoma susceptibility loci. *Nat Genet.* 2011; 43:1108–1113. [PubMed: 21983787]
- Bishop DT, Demenais F, Iles MM, Harland M, Taylor JC, Corda E, Randerson-Moor J, Aitken JF, Avril MF, Azizi E, et al. Genome-wide association study identifies three loci associated with melanoma risk. *Nat Genet.* 2009; 41:920–925. [PubMed: 19578364]
- Brennan P. Gene-environment interaction and aetiology of cancer: what does it mean and how can we measure it? *Carcinogenesis.* 2002; 23:381–387. [PubMed: 11895852]
- Fitzpatrick TB. The validity and practicality of sun-reactive skin types I through VI. *Arch Dermatol.* 1988; 124:869–871. [PubMed: 3377516]
- Hong Y, Miao X, Zhang X, Ding F, Luo A, Guo Y, Tan W, Liu Z, Lin D. The role of P53 and MDM2 polymorphisms in the risk of esophageal squamous cell carcinoma. *Cancer Res.* 2005; 65:9582–9587. [PubMed: 16230424]
- Ichii-Jones F, Lear JT, Heagerty AH, Smith AG, Hutchinson PE, Osborne J, Bowers B, Jones PW, Davies E, Ollier WE, et al. Susceptibility to melanoma: influence of skin type and polymorphism in the melanocyte stimulating hormone receptor gene. *J Invest Dermatol.* 1998; 111:218–221. [PubMed: 9699720]
- Jiao L, Hassan MM, Bondy ML, Abbruzzese JL, Evans DB, Li D. The XPD Asp312Asn and Lys751Gln polymorphisms, corresponding haplotype, and pancreatic cancer risk. *Cancer Lett.* 2007; 245:61–68. [PubMed: 16458430]
- Kanetsky PA, Panossian S, Elder DE, Guerry D, Ming ME, Schuchter L, Rebbeck TR. Does MC1R genotype convey information about melanoma risk beyond risk phenotypes? *Cancer.* 2010; 116:2416–2428. [PubMed: 20301115]
- Kanetsky PA, Rebbeck TR, Hummer AJ, Panossian S, Armstrong BK, Krickler A, Marrett LD, Millikan RC, Gruber SB, Culver HA, et al. Population-based study of natural variation in the melanocortin-1 receptor gene and melanoma. *Cancer Res.* 2006; 66:9330–9337. [PubMed: 16982779]
- Kefford RF, Newton Bishop JA, Bergman W, Tucker MA. Counseling and DNA testing for individuals perceived to be genetically predisposed to melanoma: A consensus statement of the Melanoma Genetics Consortium. *J Clin Oncol.* 1999; 17:3245–3251. [PubMed: 10506626]
- Li C, Liu Z, Wang LE, Gershenwald JE, Lee JE, Prieto VG, Duvic M, Grimm EA, Wei Q. Haplotype and genotypes of the VDR gene and cutaneous melanoma risk in non-Hispanic whites in Texas: a case-control study. *Int J Cancer.* 2008; 122:2077–2084. [PubMed: 18183598]

- Macgregor S, Montgomery GW, Liu JZ, Zhao ZZ, Henders AK, Stark M, Schmid H, Holland EA, Duffy DL, Zhang M, et al. Genome-wide association study identifies a new melanoma susceptibility locus at 1q21.3. *Nat Genet.* 2011; 43:1114–1118. [PubMed: 21983785]
- Marks R. An overview of skin cancers. Incidence and causation. *Cancer.* 1995; 75:607–612. [PubMed: 7804986]
- Matichard E, Verpillat P, Meziani R, Gerard B, Descamps V, Legroux E, Burnouf M, Bertrand G, Bouscarat F, Archimbaud A, et al. Melanocortin 1 receptor (MC1R) gene variants may increase the risk of melanoma in France independently of clinical risk factors and UV exposure. *J Med Genet.* 2004; 41:e13. [PubMed: 14757863]
- Meng S, Zhang M, Liang L, Han J. Current opportunities and challenges: genome-wide association studies on pigmentation and skin cancer. *Pigment Cell Melanoma Res.* 2012; 25:612–617. [PubMed: 22762390]
- Moan J, Dahlback A. The relationship between skin cancers, solar radiation and ozone depletion. *Br J Cancer.* 1992; 65:916–921. [PubMed: 1616864]
- 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? *Am J Hum Genet.* 2000; 66:176–186. [PubMed: 10631149]
- R FD-M, Hernandez-Jimenez JG, Carretero Hernandez G, Perez-Mendez L, Aguirre-Jaime A, Flores C, Suarez Hernandez J, Perera Molinero A, Claverie-Martin F. The D84E variant of the alpha-MSH receptor 1 gene is associated with cutaneous malignant melanoma early onset. *J Dermatol Sci.* 2008; 52:186–192. [PubMed: 18657399]
- Siegel R, Desantis C, Virgo K, Stein K, Mariotto A, Smith T, Cooper D, Gansler T, Lerro C, Fedewa S, et al. Cancer treatment and survivorship statistics, 2012. *CA Cancer J Clin.* 2012; 62:220–241. [PubMed: 22700443]
- Sprague BL, Trentham-Dietz A, Garcia-Closas M, Newcomb PA, Titus-Ernstoff L, Hampton JM, Chanock SJ, Haines JL, Egan KM. Genetic variation in TP53 and risk of breast cancer in a population-based case control study. *Carcinogenesis.* 2007; 28:1680–1686. [PubMed: 17449902]
- Stephens M, Scheet P. Accounting for decay of linkage disequilibrium in haplotype inference and missing-data imputation. *Am J Hum Genet.* 2005; 76:449–462. [PubMed: 15700229]
- Stephens M, Smith NJ, Donnelly P. A new statistical method for haplotype reconstruction from population data. *Am J Hum Genet.* 2001; 68:978–989. [PubMed: 11254454]
- Sturm RA. Skin colour and skin cancer - MC1R, the genetic link. *Melanoma Res.* 2002; 12:405–416. [PubMed: 12394181]
- Valverde P, Healy E, Sikkink S, Haldane F, Thody AJ, Carothers A, Jackson IJ, Rees JL. The Asp84Glu variant of the melanocortin 1 receptor (MC1R) is associated with melanoma. *Hum Mol Genet.* 1996; 5:1663–1666. [PubMed: 8894704]

Significance

To date, there is still lack of large and thorough case-control studies of *MC1R* variants in CM risk in US populations. In this study, we reported a detailed profile of newly identified *MC1R* variants by direct sequencing of its entire coding region and investigated their associations with CM risk in a Texas non-Hispanic white population of 1,106 CM case patients and 1,106 cancer-free control subjects. This study provided confirmatory evidence that both common and rare variants in the *MC1R* coding region may be biomarkers for susceptibility to CM in US populations.

(A).

(B).

Figure 1. LD display for *MC1R* SNPs

(A). LD display for *MC1R* SNPs that have a minor allele frequency (MAF) of >0.001 in cancer-free control subjects. (B). LD display for common *MC1R* SNPs that have MAF of >0.01 in cancer-free control subjects. Color were based on D' and logarithm (base 10) of odds (LOD) values. The red area represented $D' = 1$ and $\text{LOD} \geq 2$, the grey area represented $D' < 1$ and $\text{LOD} \geq 2$, and the white area represented $D' < 1$ and $\text{LOD} < 2$. The numbers displayed in the diamonds are $r^2 \times 100$ for pairwise SNPs.

Table 1
Demographic characteristics and risk factors in CM case patients and cancer-free control subjects

Variables	Case patients No. (%)	Control subjects No. (%)	P value	Crude OR (95% CI)	Adjusted OR (95% CI) *
All subjects	1,106 (100.0)	1,106 (100.0)			
Age group (y)			0.297		
Range	18–92	18–87		--	--
Mean	51.2±13.3	51.8±13.0		--	--
45	361 (32.6)	326 (29.5)		--	--
46 – 55	296 (26.8)	327 (29.6)		--	--
56 – 65	288 (26.1)	282 (25.5)		--	--
>65	161 (14.6)	171 (15.5)		--	--
Sex			0.696		
Female	435 (39.3)	444 (40.1)		--	--
Males	671 (60.7)	662 (59.9)		--	--
Ethnicity			1.000		
Non-Hispanic whites	1,106 (100)	1,106 (100)		--	--
Eye color			<.001		
Brown or grey	243 (22.5)	356 (32.4)		1.00	1.00
Hazel or green	390 (36.0)	394 (35.9)		1.35 (1.09–1.66)	1.34 (1.09–1.66)
Blue	449 (41.5)	349 (31.8)		1.75 (1.42–2.16)	1.75 (1.41–2.16)
Hair Color			<.0001		
Brown, black or others	696 (64.3)	851 (77.5)		1.00	1.00
Blonde	280 (25.9)	203 (18.5)		1.65 (1.34–2.03)	1.65 (1.35–2.03)
Red	107 (9.9)	44 (4.0)		2.90 (2.02–4.18)	2.92 (2.03–4.21)
Skin color			<.0001		
Dark	445 (41.1)	611 (55.6)		1.00	1.00
Mid-dark	335 (30.9)	293 (26.7)		1.51 (1.24–1.84)	1.53 (1.26–1.87)
Fair	304 (28.0)	195 (17.7)		2.06 (1.66–2.56)	2.10 (1.69–2.61)
Sunburn			<.0001		
No	298 (27.5)	464 (42.2)		1.00	1.00
Yes	785 (72.5)	635 (57.8)		1.93 (1.61–2.30)	1.92 (1.60–2.30)

Variables	Case patients No. (%)	Control subjects No. (%)	P value	Crude OR (95% CI)	Adjusted OR (95% CI) *
Freckling in the sun as a child			<0.001		
No	453 (41.9)	641 (58.3)		1.00	1.00
Yes	629 (58.1)	458 (41.7)		1.94 (1.64–2.30)	2.01 (1.69–2.40)
Tanning ability			<0.001		
Easily tan	248 (23.1)	472 (43.0)		1.00	1.00
Some degree of tan	411 (38.2)	340 (30.9)		2.08 (1.69–2.55)	2.08 (1.69–2.55)
Minimal or no tan	416 (38.7)	287 (26.1)		2.49 (2.02–3.07)	2.51 (2.03–3.10)
Prolonged exposure			<0.001		
Deeply tanned	197 (18.4)	324 (29.7)		1.00	1.00
Moderately tanned	525 (49.0)	526 (48.2)		1.45 (1.18–1.79)	1.47 (1.19–1.80)
Mildly tanned	282 (26.3)	207 (19.0)		1.99 (1.55–2.54)	2.00 (1.56–2.56)
No suntan	67 (6.3)	35 (3.2)		2.79 (1.79–4.34)	2.90 (1.86–4.53)
Moles			<0.001		
No	262 (24.2)	544 (50.3)		1.00	1.00
Yes	821 (75.8)	538 (49.7)		3.17 (2.64–3.81)	3.24 (2.69–3.90)
Dysplastic nevi			<0.001		
No	962 (89.2)	1052 (98.2)		1.00	1.00
Yes	116 (10.8)	19 (1.8)		6.67 (4.08–10.92)	6.72 (4.10–11.01)
Family history of cancer			0.401		
No	404 (37.2)	428 (39.0)		1.00	1.00
Yes	681 (62.8)	670 (61.0)		1.08 (0.91–1.28)	1.10 (0.92–1.31)

Some strata did not add to the total number of 1,106 case patients and 1,106 control subjects, because of missing data in the questionnaire.

* Adjusted for age and sex in the logistic regression model.

Table 2
Variants for *MC1R* identified in 1,106 CM case patients and 1,106 cancer-free control subjects

SNP*	Amino acid change	rs no.	MAF in cases	MAF in controls	P [†]
a. Non synonymous variants					
c.451C>T	p.Arg151Cys	rs1805007	0.141	0.084	<0.001
c.178G>T	p.Val60Leu	rs1805005	0.124	0.124	0.320
c.274G>A	p.Val92Met	rs2228479	0.104	0.089	0.109
c.478C>T	p.Arg160Trp	rs1805008	0.103	0.083	0.055
c.488G>A	p.Arg163Gln	rs885479	0.052	0.049	0.597
c.880G>C	p.Asp294His	rs1805009	0.033	0.020	0.027
c.252C>A	p.Asp84Glu	rs1805006	0.011	0.009	0.451
c.464T>C	p.Ile155Thr	rs1110400	0.011	0.005	0.044
c.425G>A	p.Arg142His	rs11547464	0.010	0.006	0.167
c.456C>A	p.Tyr152Stop [‡]		0.002	0.001	0.414
c.247T>C	p.Ser83Pro	rs34474212	0.001	<0.001	0.563
c.284C>T	p.Thr95Met	rs34158934	0.001	0.001	1.000
c.392G>A	p.Ser131Asn		0.001	0	0.157
c.424C>T	p.Arg142Cys [‡]		0.001	0	0.223
c.652G>A	p.Ala218Thr		0.001	0	0.157
c.832A>G	p.Lys278Glu		0.001	0	0.157
c.917G>A	p.Arg306His		0.001	<0.001	0.563
c.56C>T	p.Thr191Ile [‡]		<0.001	0	0.317
c.88C>T	p.Gln303Stop ^{‡,§}		<0.001	0	0.317
c.112G>A	p.Val38Met		<0.001	0.001	0.317
c.133T>C	p.Phe451Leu	rs61996344	<0.001	<0.001	1.000
c.221A>G	p.Tyr74Cys [‡]		v0.001	0	0.317
c.265G>C	p.Gly89Arg	rs34540312	<0.001	<0.001	1.000
c.363C>G	p.Asp121Glu [‡]		<0.001	0	0.317
c.364G>A	p.Val122Met		<0.001	0	0.317
c.466G>C	p.Val156Ileu	rs3212365	<0.001	0	0.317

SNP*	Amino acid change	rs no.	MAF in cases	MAF in controls	P†
c.550G>A	p.Asp184Asn		<0.001	0	0.317
c.586T>C	p.Phe196Leu	rs3212366	<0.001	0	0.317
c.637C>T	p.Arg213Trp		<0.001	<0.001	1.000
c.667C>T	p.Arg223Trp		<0.001	0	0.317
c.743G>T	p.Gly248Val		<0.001	0	0.317
c.766C>T	p.Pro256Ser		<0.001	0	0.317
c.779A>T	p.His260Leu‡		<0.001	0	0.317
c.836A>G	p.Asn279Ser‡		<0.001	<0.001	1.000
c.837C>A	p.Asn279Lys		<0.001	0	0.317
c.860T>G	p.Ile287Ser		<0.001	<0.001	1.000
c.107T>C	p.Leu36Pro‡		0	<0.001	0.317
c.140G>C	p.Ser47Thr‡		0	<0.001	0.317
c.328G>A	p.Ala110Thr‡		0	<0.001	0.317
c.424C>A	p.Arg142Ser‡		0	<0.001	0.223
c.467T>C	p.Val156Ala		0	<0.001	0.317
c.689C>T	p.Pro230Leu		0	<0.001	0.317
b. Insertions					
c.86_87insA	p.Asn29Lys#		<0.001	0	0.317
c. Synonymous SNP					
c.942A>G	p.Thr314Thr	rs2228478	0.123	0.104	0.128
c.699G>A	p.Gln233Gln		0.004	0.003	0.796
c.948C>T	p.Ser316Ser		0.003	0.002	0.365
c.717C>T	p.Gly239Gly	rs344490506	0.001	0.001	0.654
c.792C>T	p.Ile264Ile		0.001	<0.001	0.317
c.318G>A	p.Leu106Leu	rs3212364	0.001	<0.001	0.563
c.399C>T	p.Cys133Cys		0.001	<0.001	0.563
c.453C>G	p.Arg151Arg		0.001	<0.001	0.564
c.498G>A	p.Ala156Ala		<0.001	<0.001	1.000
c.555C>T	p.His185His		<0.001	<0.001	1.000

SNP*	Amino acid change	rs no.	MAF in cases	MAF in controls	P [†]
c.576C>G	p.Leu192Leu [‡]		<0.001	0	0.317
c.615G>A	p.Val205Val [‡]		<0.001	0	0.317
c.651C>T	p.His217His		<0.001	0	0.317
c.900C>T	p.Phe300Phe	rs3212367	0	<0.001	0.317
d.5',3'UTR					
c.-28C>T [‡]			0	<0.001	0.317
c.-13C>T [‡]			0	<0.001	0.317
c. [§] 2G>A		rs3212368	0.001	<0.001	0.317
c. [*] 107G>C [‡]			<0.001	0	0.317

MAF, Minor Allele Frequency

*The SNPs position were determined according to the *MC1R* mRNA sequence NM_002386.2 from NCBI.

[†] P value is from the chi-square analysis for difference between case patients and control subjects.

[‡]These SNPs were either never reported or not found in the following SNPs database: dbSNP database, Ensembl Genome Browser, and SNP500 Cancer.

[§] Changed to stop codon at codon 30;

[‡] Changed to stop codon at codon 152;

[#] Frameshift, stop codon at 42.

Table 3
Genotyping distribution of selected SNPs in CM case patients and cancer-free control subjects for *MC1R*

SNP position	Case patients No. (%)	Control subjects No. (%)	P value	Crude OR (95% CI)	Adjusted OR (95% CI)*
All subjects	1,106 (100.0)	1,106 (100.0)			
c.451C>T (rs1805007)			<0.001 [‡]		
CC	826 (74.7)	928 (83.9)		1.00	1.00
CT	249 (22.5)	171 (15.5)		1.64 (1.32–2.03)	1.65 (1.33–2.05)
TT	31 (2.8)	7 (0.6)		4.98 (2.18–11.4)	5.06 (2.21–11.6)
CT + TT	280 (25.3)	178 (16.1)	<0.001 [‡]	1.77 (1.43–2.18)	1.78 (1.44–2.20)
c.478C>T (rs1805008)			0.055 [‡]		
CC	891 (80.5)	934 (84.5)		1.00	1.00
CT	202 (18.3)	161 (14.5)		1.32 (1.05–1.65)	1.32 (1.05–1.65)
TT	13 (1.2)	11 (1.0)		1.24 (0.55–2.78)	1.22 (0.54–2.74)
CT + TT	215 (19.5)	172 (15.5)	0.016 [‡]	1.31 (1.05–1.63)	1.31 (1.05–1.63)
c.880G>C (rs1805009)			0.027 [‡]		
GG	1034 (93.5)	1062 (95.9)		1.00	1.00
CG	71 (6.4)	43 (3.9)		1.70 (1.15–2.50)	1.70 (1.15–2.51)
CC	1 (0.1)	1 (0.1)		1.03 (0.06–16.4)	1.05 (0.06–16.9)
CG + CC	72 (6.5)	44 (4.1)	0.008 [‡]	1.68 (1.14–2.47)	1.69 (1.15–2.48)
c.942A>G (rs2228478)			0.128 [‡]		
AA	851 (76.9)	890 (80.5)		1.00	1.00
GA	237 (21.4)	201 (18.2)		1.23 (1.00–1.52)	1.23 (1.00–1.52)
GG	18 (1.6)	15 (1.4)		1.26 (0.63–2.51)	1.24 (0.62–2.49)
GA + GG	255 (23.1)	216 (19.5)	0.043 [‡]	1.24 (1.01–1.51)	1.23 (1.00–1.51)
c.274G>A (rs2228479)			0.109 [‡]		
GG	886 (80.1)	921 (83.3)		1.00	1.00
GA	210 (19.0)	173 (15.6)		1.26 (1.01–1.57)	1.26 (1.01–1.57)
AA	10 (0.9)	12 (1.1)		0.87 (0.37–2.02)	0.86 (0.37–2.00)
GA + AA	220 (19.9)	185 (16.7)	0.054 [‡]	1.24 (1.00–1.53)	1.23 (0.99–1.53)

* Adjusted for age and sex in the logistic regression model.

[†] Distribution of three genotypes.

[‡] Distribution of combined genotypes.

Table 4

Cumulative effects of selected variant genotypes of *MC1R* in CM *

No. of variant genotypes, assuming a dominant model	Case patients (n = 1,106) No. (%)	Control subjects (n = 1,106) No. (%)	P	Crude OR (95% CI)	Adjusted OR (95% CI) [†]
0	414 (37.4)	560 (50.6)	<0.001	1.00	1.00
1	410 (37.1)	332 (30.0)		1.67 (1.38–2.03)	1.68 (1.39–2.04)
2	214 (19.4)	179 (16.2)		1.62 (1.28–2.05)	1.61 (1.27–2.04)
3	68 (6.2)	35 (3.2)		2.63 (1.71–4.03)	2.64 (1.72–4.05)

* The variant genotypes were based on the five SNPs c.451C>T (rs1805007), c.478C>T (rs1805008), c.880G>C (rs1805009), c.942A>G (rs2228478) and c.274G>A (rs2228479) and based on the dominant model.

[†] Adjusted for age and sex in the logistic regression model.

Table 5

Haplotype and diplotype distribution in CM case patients and cancer-free control subjects*

Haplotype	No. of variant genotypes	Case patients (chr. No. = 2,212) No. (%)	Control subjects (chr. No. = 2,212) No. (%)	Crude OR (95% CI)	Adjusted OR (95% CI) [†]
G-C-C-G-A	0	1,327 (60.0)	1,566 (70.8)	1.00	1.00
G-T-C-G-A	1	310 (14.0)	185 (8.4)	1.93 (1.59–2.34)	1.94 (1.60–2.36)
A-C-C-G-G	2	228 (10.3)	195 (8.8)	1.36 (1.11–1.67)	1.36 (1.11–1.67)
G-C-T-G-A	1	227 (10.3)	183 (8.3)	1.43 (1.16–1.76)	1.43 (1.16–1.76)
Others [‡]		120 (5.4)	83 (3.8)	1.70 (1.27–2.27)	1.70 (1.27–2.28)
Diplotype	No. of variant genotypes	Case patients (n = 1,106) No. (%)	Control subjects (n = 1,106) No. (%)	Crude OR (95% CI)	Adjusted OR (95% CI) [†]
G-C-C-G-A/G-C-C-G-A	0/0	414 (37.4)	560 (50.6)	1.00	1.00
G-C-C-G-A/G-T-C-G-A	0/1	166 (15.0)	134 (12.1)	1.67 (1.29–2.17)	1.69 (1.30–2.19)
A-C-C-G-G/G-C-C-G-A	2/0	132 (11.9)	134 (12.1)	1.33 (1.01–1.75)	1.32 (1.01–1.74)
G-C-C-G-A/G-C-T-G-A	0/1	132 (11.9)	115 (10.4)	1.55 (1.17–2.05)	1.56 (1.18–2.06)
Others [§]		262 (23.7)	163 (14.7)	2.17 (1.72–2.74)	2.17 (1.72–2.74)

*The P value for χ^2 tests were $P < 0.001$ for both haplotype and diplotype frequencies distributions.

The haplotypes and diplotypes were based on the order of 5 SNPs c.451C>T (rs1805007), c.478C>T (rs1805008), c.880G>C (rs1805009), c.942A>G (rs2228478) and c.274G>A (rs2228479).

[†] Adjusted for age and sex in the logistic regression model.

[‡] Included 15 minor haplotypes, each with a frequency < 0.05 .

[§] Included 20 minor diplotypes, each with a frequency < 0.05 .