Permutation tests

Univariate analyses of SNPs and CM

We carried out a permutation test to check the significance of the univariate associations of SNPs with melanoma. We permuted the case/control status for all individuals 999 times and estimated the association P-values for all SNPs by a logistic regression model comprising a single genotype and an intercept (age and sex were also included when a permutation test was carried out for cofactor models). Only individuals without any missing genotypes for all SNPs were considered. The empirical P-value for the variant *i* was determined as follows:

$$p_{i,emp} = (count_k(p_k \leq p_{i,nom}) + 1)/1000$$

where $count_k(p_k \leq p_{i,nom})$ was the count of the smallest permutation P-values p_k among all SNPs in the permuted data set k that were less than or equal to the actual P-value $p_{i,nom}$ for the variant i. Finally, empirical P-values were compared with the significance threshold of 0.05.

Univariate analyses of SNPs and pigmentation traits

We assessed whether the associations between individual SNPs and a pigmentation trait are significant by the following permutation test. For a particular trait, only the individuals having information about this trait were used in permutations. The pigmentation trait information was permuted, while the genotypes of individuals were held constant. For each permutation of a trait and each genotype Fisher's exact test was used to obtain a nominal P-value of the association. The empirical P-value for the variant *i* and a fixed pigmentation trait was estimated as follows:

$$p_{i,emp} = (count_k(p_k \leq p_{i,nom}) + 1)/1000$$

where $count_k(p_k \leq p_{i,nom})$ was the count of the smallest permutation P-values p_k among all SNPs in the permuted data set k that were less than or equal to the nominal P-value $p_{i,nom}$ for the variant i and the chosen pigmentation trait. Empirical P-values were compared to 0.05.

Joint analyses of CM and pigmentation traits

Multivariate models incorporating the CM status and a pigmentation trait were also subjected to a permutation test. For a particular pair of traits, only the individuals having information about the chosen pigmentation trait were used in permutations (the CM status was always present). Both the CM phenotype and the pigmentation trait were permuted together, while the genotypes of individuals were held constant. For each permutation of phenotypes and each genotype an ordinal regression model was fitted with both traits as explanatory variables for the genotype. Another ordinal regression model was fitted where only an intercept modelled the genotype. A likelihood ratio test for a model with two traits against a model without any traits was carried out for each combination of a phenotype permutation and a variant. All P-values corresponded to a Chi-squared test. The empirical P-value for the variant *i* and a fixed pair of the pigmentation trait and CM was estimated as follows:

$$p_{i,emp} = (count_k(p_k \leq p_{i,nom}) + 1)/1000$$

where $count_k(p_k \leq p_{i,nom})$ was the count of the smallest permutation P-values p_k among all SNPs in the permuted data set k that were less than or equal to the nominal P-value $p_{i,nom}$ for the variant i and the same pair of phenotypes. Empirical P-values were compared to 0.05.

Associations between SNPs and histopathological features

We devised a permutation test for the associations between SNPs and histopathological features (ulceration and Breslow thickness). Permutations for ulceration and Breslow thickness were constructed separately-in each case a permutation involved only those individuals that had the information about particular feature. Thus all individuals with a missing feature were excluded from respective analysis. Then each feature was permuted across the CM cases for which it was present 999 times. We estimated the association P-values for all SNPs for each permutation by either the t-test (for Breslow thickness) or the Wald test (for ulceration) applied to the genotype term of a regression model. Finally, the empirical P-value for the feature j and the variant i was determined as follows:

$$p_{i,j,emp} = (count_k(p_{k,j} \leq p_{i,j,nom}) + 1)/1000$$

where $count_k(p_{k,j} \leq p_{i,j,nom})$ was the count of the smallest permutation P-values $p_{k,j}$ among all SNPs in the permuted data set k for the feature j that were less than or equal to the actual P-value $p_{i,j,nom}$ for the variant i and the feature j. All empirical P-values were compared with the threshold of 0.05.

Table SI. MAFs of rare MC1R variants within cases and controls.

Gene	SNP	MAF controls % (n=224)	MAF patients % (n=255)
A, Non- synonymous			
MC1R	rs1805005 p.Val60**Leu** c.178G>T	2.9	5.6
MC1R	rs777024553 p.Ser83**Leu** c.248C>T	0.0	0.3
MC1R	rs1805006 p.Asp84**Glu** c.252C>A	0.0	0.4
MC1R	rs34540312 p.Gly89**Arg** c.265G>C	0.3	0.0
MC1R	rs34158934 p.Thr95**Met** c.284C>T	0.0	0.1
MC1R	rs200616835 p.Asp121**Glu** c.363C>G	0.2	0.2
MC1R	rs11547464 p.Arg142**His** c.425G>A	2.1	1.6
MC1R	rs885479 p.Arg163**Gln** c.488G>A	3.8	3.4
MC1R	rs762096175 p.Val165**Ile** c.493G>A	0.3	0.0
MC1R	rs780875127 c.495_496insGG	0.3	0.0
MC1R	rs530102853 p.Asp184**His** c.550G>C	0.3	0.0
MC1R	rs774680166 p.Val188**Ile** c.562G>A	0.0	0.2
MC1R	rs200000734 p.Arg213**Trp** c.637C>T	0.2	0.0
B, Synonymous			
MC1R	rs201429598 p.Cys133= c.399C>T	0.0	0.4
MC1R	rs201827012 p.Arg151= c.453C>G	0.0	0.4
MC1R	rs374959395 p.Ala166= c.498G>A	0.0	0.2
MC1R	rs146544450 p.Gln233= c.699G>A	0.5	0.9
MC1R	rs375813196 p.Cys273= c.819C>T	0.2	0.0
MC1R	rs2228478 p.Thr314= c.942A>G	14.0	15.7
MC1R	rs151318945 p.Ser316= c.948C>T	1.7	1.7

Risk alleles are indicated in bold. MAF, minor allele frequency; MC1R, melanocortin 1 receptor; SNP, single nuclear polymorphism.