Multiple Pigmentation Gene Polymorphisms Account for a Substantial Proportion of Risk of Cutaneous Malignant Melanoma

David L. Duffy[1], Zhen Z. Zhao[1], Richard A. Sturm[2,1], Nicholas K. Hayward[1], Nicholas G. Martin[1] and Grant W. Montgomery[1]

We have previously described the role of red hair (melanocortin-1 receptor, *MC1R*) and blue eye (oculocutaneous albinism type II, *OCA2*) gene polymorphisms in modulating the risk of cutaneous malignant melanoma (CMM) in a highly sun-exposed population of European descent. A number of recent studies, including genome-wide association studies, have identified numerous polymorphisms controlling human hair, eye, and skin color. In this paper, we test a selected set of polymorphisms in pigmentation loci (*ASIP* (Agouti signalling protein, nonagouti homolog (mouse) gene), *TYR* (tyrosinase), *TYRP1* (tyrosinase-related protein 1), *MC1R*, *OCA2*, *IRF4* (interferon regulatory factor 4), *SLC24A4* (solute carrier family 24, member 4), and *SLC45A2* (solute carrier family 45, member 2)) for association with CMM risk in a large Australian population-based case–control study. Variants in *IRF4* and *SLC24A4*, despite being strongly associated with pigmentation in our sample, did not modify CMM risk, but the other six did. Three single nucleotide polymorphisms (rs28777, rs35391, and rs16891982) in the MATP gene (*SLC45A2*) exhibited the strongest crude association with risk, but this was attenuated to approximately the same effect size as that of a *MC1R* red hair color allele by controlling for ancestry of cases and controls. We also detected significant epistatic interactions between *SLC45A2* and *OCA2* alleles, and *MC1R* and *ASIP* alleles. Overall, these measured variants account for 12% of the familial risk of CMM in our population.

Journal of Investigative Dermatology (2010) **130,** 520–528; doi:10.1038/jid.2009.258; published online 27 August 2009

INTRODUCTION

Cutaneous malignant melanoma (CMM) is a common cancer in the fair-skinned population of tropical and subtropical Queensland, Australia, with an estimated lifetime risk of 5% (Queensland Cancer Registry, 2008). Within the fairer skinned populations, risk varies with the degree of paleness as well as with other pigmentation phenotypes, such as hair and eye color, all of which are strongly genetically determined.

We have previously described association between variants in the pigmentation genes encoding the melanocortin-1 receptor (*MC1R*) (Palmer *et al.*, 2000; Box *et al.*, 2001) and the human melanocyte-specific P-protein (oculocutaneous albinism type II, *OCA2*) with CMM risk (Sturm *et al.*, 2008). A number of recent genome-wide association studies have characterized variants in both known pigmentation pathways (*TYR* (tyrosinase), *TYRP1* (tyrosinase-related protein 1), and *SLC45A2* (solute carrier family 45, member 2)) (Sulem *et al.*, 2007; Sulem *et al.*, 2008) and previously unknown (*IRF4* (interferon regulatory factor 4) and *SLC24A4* (solute carrier family 24, member 4)) (Sulem *et al.*, 2007; Han *et al.*, 2008) genes that influence pigmentation, and one study has reported on associations with CMM (Gudbjartsson *et al.*, 2008). Association with both pigmentation (Sulem *et al.*, 2008) and CMM risk (Brown *et al.*, 2008; Gudbjartsson *et al.*, 2008) has also been reported for variants in a region of chromosome 20 fairly close to the human homologue of the mouse Agouti gene (ASIP).

Variants in many of these genes exhibit significant differences in frequency, and underlie the pigmentation differences, between different ethnic groups. It is important, therefore, to differentiate between disease associations due to a direct genotype–phenotype relationship, and those due to confounding by ancestry.

Here, we replicate and extend association results with CMM to *MC1R*, *OCA2*, *TYR* (*OCA1*), *TYRP1* (*OCA3*), *IRF4*, *SLC24A4*, *SLC45A2* (*MATP*, *OCA4*), and *ASIP*, and examine

[1]*Genetic Epidemiology Laboratory, Queensland Institute of Medical Research, Brisbane, Australia and* [2]*Institute of Molecular Bioscience, The University of Queensland, Melanogenix Group, Brisbane, Australia*

Correspondence: Dr David L. Duffy, Genetic Epidemiology Laboratory, Queensland Institute of Medical Research, 300 Herston Road, Brisbane 4029, Australia. E-mail: david.duffy@qimr.edu.au

Abbreviations: ASIP, *Agouti signalling protein, nonagouti homolog (mouse) gene; CMM, cutaneous malignant melanoma;* IRF4, *interferon regulatory factor 4 gene;* MC1R, *melanocortin-1 receptor gene;* OCA2, *oculocutaneous albinism type II gene;* SLC24A4, *solute carrier family 24 (sodium/potassium/calcium exchanger), member 4 gene;* SLC45A2, *solute carrier family 45, member 2 gene (alias* OCA4, MATP, AIM-1*); SNP, single nucleotide polymorphism;* TYR, *tyrosinase gene (alias* OCA1*);* TYRP1, *tyrosinase-related protein 1 gene (alias* OCA3*)*

Received 9 December 2008; revised 27 May 2009; accepted 19 June 2009; published online 27 August 2009

the nature of interaction between these loci on disease risk in a large sample from a population at high environmental and genetic risk of melanoma.

RESULTS

There were 1,738 CMM cases and 4,517 controls. Ancestry and coloring for the sample are shown in Table 1. Differences between cases and controls are in the expected direction, as reported previously for this sample (Palmer *et al.*, 2000), with cases tending to have lighter colored eyes and hair. Participants were genotyped at up to 161 single nucleotide polymorphisms (SNPs) across the gene regions of interest (Supplementary Table 1). Table 2 shows case and control allele frequencies and Hardy–Weinberg equilibrium tests for selected SNPs. A number of SNPs exhibited extreme Hardy–Weinberg disequilibrium, but as the chosen SNPs are located in pigmentation loci that are known to be highly informative for coloring and ancestry, it could be shown that the disequilibrium was due to ethnic stratification of the sample rather than to genotyping error (see Supplementary Table 2).

One or more of the SNPs selected in each gene region could be shown to be associated with skin, hair, or eye color (Table 3), and in the subsequent analyses, we only present results for association with CMM for the polymorphisms most penetrant for pigmentary characteristics from each gene. In keeping with the selection criteria, a number of these SNPs were also strongly associated with the reported ancestry (see Supplementary Table 2). Among the CMM cases with ancestry information, only 2.7% reported other than 100% Northern European ancestry, as compared with 13.5% of the control sample. To control these effects on CMM association analyses given below, we report analyses on the subset of 1,438 cases and 3,098 controls reporting four grandparents of Northern European ancestry. We found that further inclusion of an ancestry indicator for English, Scottish/Welsh/Irish, and continental Europe as a covariate did not detect finer levels of ethnic confounding on CMM association. Interestingly, rs12203592 (in *IRF4*) did exhibit significant allele frequency differences between controls from continental Europe, England and Scotland, Ireland and Wales, but these could be safely ignored given that this locus does not exhibit significant association with CMM (see below).

Table 1. Descriptive statistics for genotyped CMM cases and controls (percentages are of total with data for that phenotype, except for ancestry)

	CMM Cases	Controls
All four grandparents of Northern European ancestry	1,483 (93.0%)	3,098 (83.3%)
Blue/grey eyed	758 (43.7%)	1,838 (44.7%)
Green/hazel eyed	692 (39.9%)	1,212 (29.5%)
Brown/black eyed	283 (16.3%)	1,059 (25.8%)
Blond hair	419 (24.1%)	569 (23.7%)
Light brown hair	653 (37.6%)	1,391 (33.9%)
Red hair	193 (11.1%)	217 (5.3%)
Dark brown hair	402 (23.1%)	1,705 (41.5%)
Black hair	71 (4.1%)	226 (5.5%)
Male	967 (47.8%)	2,129 (47.2%)
Total genotyped individuals	1,738	4,517

CMM, cutaneous malignant melanoma.

In univariate analyses, *MC1R*, *ASIP*, *TYRP1*, (*IRF4*), and *SLC45A2* SNPs were the markers most strongly associated with CMM risk in terms of crude risk (Table 4). Given the star-shaped genealogy of *MC1R* red-hair-associated haplotypes (that is, all these coding variants are in complete negative linkage disequilibrium; in our population we have observed only three haplotypes with more than one coding variant present), we have recoded *MC1R* haplotypes as ''*R*'' (high red hair penetrance: rs1805006 (*D84E*), rs11547464 (*R142H*), rs1805007 (*R151C*), rs1805008 (*R160W*), and rs1805009 (D294H)) and ''*r*'' (low red hair penetrance: rs1805005, rs2228479, and rs885479) (Sturm *et al.*, 2003) for this and subsequent analyses. On restricting the case–control analysis only to individuals reporting all four grandparents to be of Northern European ancestry, the significant predictors of CMM risk were four ''R'' SNPs in *MC1R* (rs1805007, rs1805008, rs1805006, and rs11547464); the three *SLC45A2* SNPs; rs1126809 in *TYR*; and rs4911442 in the region of *ASIP*. In the case of *SLC45A2*, the three SNPs were in very strong disequilibrium, and the noncoding SNPs, rs35391 and rs28777, were not significant melanoma risk predictors in multivariate logistic regression models after adjusting for rs16891982.

When we further adjusted for hair, eye, and skin color, several SNPs exhibited a persistent association, suggesting that they act as mediating variables (Table 4) between ancestry and melanoma risk. The *TYR* coding SNP, rs1126809, remained important, but rs12913832 in *OCA2* (a regulatory SNP actually located in an intron of the adjacent *HERC2* gene, but the main determinant of blue eye color, Sturm *et al.*, 2008) and the *TYRP1* SNPs were less impressive as CMM risk factors. The effect of *OCA2**R419Q was relatively small in its own right. When we combine the strongest associated SNPs from *TYR*, *SLC45A2*, *OCA2*, *MC1R*, and *ASIP* in a multivariate logistic regression (Table 5), this gives a Nagelkerke R^2 index = 0.059. For comparison, a regression including eye, skin, and hair color obtained $R^2 = 0.048$, and supplementing this with the genotype data increased the R^2 to 0.08. As logistic regression R^2-values are difficult to interpret (Nagelkerke 1991), we also present locus-specific sibling recurrence risk ratios (λ_S). Under the usual multiplicative model assumption, these can be multiplied to give a total contribution of 1.14 to the sibling recurrence risk. We can calculate the equivalent contribution of measured pigmentation phenotype (skin, hair, and eye color) at 1.37, so that the current SNPs explain ~43% of the effects of pigmentation on familial risk (log(1.15)/log(1.37)) and ~15% of the total observed sibling risk $\lambda_S = 2.49$ estimated previously in our population (Do *et al*, 2004).

Table 2. Allele frequencies and tests of Hardy–Weinberg equilibrium for pigmention-associated polymorphisms in Australian CMM cases and controls

				Controls			CMM Cases		
Chr	Locus/variant	rsID	Alleles	N	MAF	HWE-P	N	MAF[1]	HWE-P[2]
5	SLC45A2 *i3*[1]	rs35391	A (G)	4,107	0.0321	0.1527	1,708	0.0088	1.0000
5	SLC45A2 *i3*	rs28777	G (T)	4,111	0.0429	0.0001	1,716	0.0118	1.0000
5	SLC45A2*F374L	rs16891982	G (C)	4,032	0.0518	0.0000	1,731	0.0133	0.0346
6	IRF4 *i4*	rs12203592	T (C)	4,113	0.2259	0.8696	1,727	0.2107	0.0264
9	TYRP1	rs1408799	T (C)	4,110	0.3275	0.0766	1,703	0.2877	0.0980
9	TYRP1 *i6*	rs2733832	C (T)	4,098	0.4119	0.0492	1,732	0.3753	0.1514
11	TYR *i3*	rs10765198	C (T)	4,103	0.3138	0.1739	1,721	0.3405	0.2930
11	TYR*S192Y	rs1042602	A (C)	4,112	0.3486	0.7407	1,728	0.3390	0.0150
11	TYR*R402Q	rs1126809	A (G)	4,102	0.3061	0.7688	1,731	0.35	0.08
14	SLC24A4	rs12896399	T (G)	4,102	0.4493	0.2857	1,705	0.4564	0.1512
15	OCA2*R419Q	rs1800407	A (G)	3,996	0.0798	0.6061	1,052	0.1042	1.0000
15	HERC2 *i86*	rs12913832	T (C)	4,102	0.2315	0.0190	1,747	0.2116	0.1133
16	MC1R*V60L	rs1805005	T (G)	2,849	0.1212	0.6061	1,945	0.1181	0.2041
16	MC1R*D84E	rs1805006	T (G)	2,851	0.0111	0.4651	1,947	0.0194	0.0008
16	*MC1R**V92M	rs2228479	A (G)	1,882	0.0961	0.0794	1,286	0.0995	0.5023
16	*MC1R**R142H	rs11547464	A (G)	1,805	0.0044	0.4878	1,114	0.0198	1.0000
16	*MC1R**R151C	rs1805007	T (C)	1,876	0.1096	0.9091	1,126	0.1567	0.5223
16	*MC1R**R160W	rs1805008	T (C)	2,848	0.0697	0.6061	1,947	0.1122	0.8365
16	*MC1R**R163Q	rs885479	A (G)	1,896	0.0466	0.7407	1,291	0.0476	0.6452
16	*MC1R**D294H	rs1805009	C (G)	2,851	0.0266	0.9091	1,947	0.0328	1.0000
20	*ASIP* region[2]	rs4911442	C (T)	4,089	0.1306	0.3003	1,745	0.1848	0.2003

ASIP, Agouti signaling protein; CMM, cutaneous malignant melanoma; HWE-P, Hardy–Weinberg Equilibrium test *P*-value; IRF4, interferon regulatory factor 4; MAF, minor allele frequency; MC1R, melanocortin-1 receptor; OCA2, oculocutaneous albinism type II; SLC45A2, solute carrier family 45, member 2; SLC24A4, solute carrier family 24, member 4; TYR, tyrosinase; TYRP1, tyrosinase-related protein 1.
[1]The letter *i* denotes intron, so HERC2 *i*86 denotes the 86th intron of the HERC2 gene.
[2]Actually present in intron of *NCOA6*, 500 kbp distal to *ASIP*.

In analyses including pairwise interactions, we observed an interaction only between *OCA2**R419Q and rs1126809 ($P=0.01$, Supplementary Table 3), although there was a suggestive pattern of interaction between *MC1R* and *ASIP* genotypes and CMM, which mirrored a highly significant interaction between these loci in their effects on the red hair color (Table 6).

As we noted elsewhere (Brown *et al.*, 2008), a number of SNPs spread over the region around ASIP exhibit association with CMM, and the analyses above have concentrated on rs4911442 only because it is genotyped in more controls than the other peak SNPs. The pattern of linkage disequilibrium around *ASIP* is quite complex, with ''hotspots'' just distal to the gene and 70 kbp proximal (see Figure 1), and the pattern of individual SNP association with CMM and with red hair color is bimodal, dipping rather than rising over *ASIP*.

Examining haplotypes that are comprised of 31 SNPs spanning from rs17305657 to rs1885120 (a distance of 1.77 Mbp) revealed that the commonest haplotype (present in 4% of controls) is significantly increased in CMM cases (7.3%, odds ratio = 1.89; see Supplementary Table 4). This haplotype is tagged by the SNPs that are most strongly individually associated with CMM (including rs4911442). When analysis is restricted just to six tagging SNPs, there are eight haplotypes present at greater than 1% study frequency, one (78% of control samples) being significantly increased in controls and three (3.8%, 3.9%, 0.6%) increased in cases (global $\chi^2=47.2$, d.f. = 14, $P=2\times10^{-5}$). The latter three share only the distal three markers in common; hence the most likely location for the susceptibility locus lies downstream to *ASIP*.

Using a publicly available US population genome-wide association studies data set (Simon-Sanchez *et al.*, 2007), we were able to show that the peak CMM SNP set described by Gudbjartsson *et al.* (2008) was in strong linkage disequilibrium with rs4911442, despite being 600 kbp distant ($P=4\times10^{-13}$). Specifically, an individual carrying the G-A risk haplotype defined by Gudbjartsson *et al.* (2008) had a 50% probability of carrying the rs4911442*C on that haplotype, whereas the other three haplotypes were associated with rs4911442*T in 99% of cases.

Table 3. Results of association analyses of selected SNPs with skin, eye, and hair color in individuals with all four grandparents of Northern European ancestry

Locus/variant	SNP	Allelic odds ratio for pale skin (*P*-value)	Allelic odds ratios for blue eye color (*P*-value)	Allelic odds ratios for light hair (*P*-value)[1]
SLC45A2 *i3*	rs35391	3.42 (<1.0e-5)	2.60 (0.001)	5.34 (<1.0e-5)
SLC45A2 *i3*	rs28777	4.31 (<1.0e-5)	2.56 (0.017)	6.30 (<1.0e-5)
SLC45A2*F374L	rs16891982	3.80 (<1.0e-5)	2.50 (<1.0e-5)	6.72 (<1.0e-5)
IRF4 *i4*	rs12203592	1.63 (<1.0e-5)	1.28 (0.0017)	2.22 (<1.0e-5)
TYRP1	rs1408799	1.06 (0.2985)	1.22 (0.018)	1.22 (<1.0e-5)
TYRP1 *i6*	rs2733832	1.05 (0.3333)	1.22 (0.007)	1.22 (<1.0e-5)
TYR *i3*	rs10765198	1.25 (0.0015)	1.33 (<1.0e-5)	1.01 (0.77)
TYR*S192Y	rs1042602	1.06 (0.11)	1.04 (0.57)	1.13 (0.13)
TYR*R402Q	rs1126809	1.27 (<1.0e-5)	1.43 (<1.0e-5)	
SLC24A4	rs12896399	1.17 (0.061)	1.85 (<1.0e-5)	1.45 (<1.0e-5)
OCA2*R419Q	rs1800407	1.10 (0.59)	2.37 (<1.0e-5)	1.21 (0.14)
HERC2 *i86*	rs12913832	1.74 (<1.0e-5)	30.77 (<1.0e-5)	2.95 (<1.0e-5)
MC1R*V60L	rs1805005	1.01 (0.85)	1.00 (0.80)	1.13 (0.008)
MC1R*D84E	rs1805006	5.19 (<1.0e-5)	1.23 (0.61)	1.74 (0.041)
MC1R*V92M	rs2228479	1.38 (0.024)	1.06 (0.61)	1.17 (0.016)
MC1R*R142H	rs11547464	1.31 (0.54)	1.03 (0.71)	1.28 (0.74)
MC1R*R151C	rs1805007	2.57 (<1.0e-5)	1.09 (0.94)	2.13 (<1.0e-5)
MC1R*R160W	rs1805008	1.72 (<1.0e-5)	1.12 (0.17)	1.82 (<1.0e-5)
MC1R*R163Q	rs885479	1.10 (0.53)	1.21 (0.15)	1.12 (0.014)
MC1R*D294H	rs1805009	2.88 (<1.0e-5)	1.16 (0.74)	1.68 (<1.0e-5)
MC1R "R"[2]		3.70 (<1.0e-5)	1.09 (0.43)	2.30 (<1.0e-5)
MC1R "r"[3]		1.84	1.10	1.26
ASIP region	rs4911442	1.43 (<1.0e-5)	1.01 (0.87)	1.28 (<1.0e-5)

ASIP, Agouti signaling protein; CMM, cutaneous malignant melanoma; IRF4, interferon regulatory factor 4; MC1R, melanocortin-1 receptor; OCA2, oculocutaneous albinism type II; SLC45A2, solute carrier family 45, member 2; SLC24A4, solute carrier family 24, member 4; SNP, small nucleotide polymorphisms; TYR, tyrosinase; TYRP1, tyrosinase-related protein 1.
[1]Light hair=blond, light brown, red.
[2]High-penetrance red hair color MC1R haplotype: variant at D84E, R142H, R151C, R160W, or D294H.
[3]Low-penetrance red hair color MC1R haplotype: variant at V60L, V92M, or R163Q.

DISCUSSION

This paper attempts to address two topics. The first is to confirm and quantify the overall contribution of currently recognized pigmentation loci to variation in risk of cutaneous melanoma. Second, we expand the evidence of the recently recognized association between SNPs near *ASIP* and CMM, and attempt to reconcile results from different studies (Gudbjartsson *et al.* 2008; Brown *et al.*, 2008).

We have replicated association of SNPs at our chosen loci *MC1R*, *OCA2*, *ASIP*, *SLC45A2*, *SLC24A4*, *TYR*, *TYRP1*, and *IRF4* with skin, hair, and eye color, and shown that most of these SNPs are also significant predictors of CMM risk in our population. As some SNPs are not genotyped in all subjects, the data set suitable for multivariate analysis is reduced; hence the confidence intervals around simultaneously adjusted effect sizes are broader than that for (most of) the univariate analyses. The pattern is broadly consistent with multiplicative epistasis, with only weak evidence for non-multiplicative interactions. The combination of SNPs in the five associated loci (*MC1R*, *SLC45A2*, *OCA2*, *TYR*, and *ASIP*) explains roughly one-third to one-half of the variation in risk due to observed pigmentation phenotype. Pigmentation phenotype itself explains approximately one-third of the familial aggregation of melanoma. By contrast, the high-penetrance CDKN2A mutations segregating in families with multiple cases of melanoma explain only 0.1% of recurrence risk in the total population. This is because they are extremely rare and involved in 0.2% of all Queensland melanoma cases (Aitken *et al.*, 1999), with an allele frequency of $\sim 5 \times 10^{-6}$. Attributable risk is another measure of the importance of a risk locus, but is not completely satisfactory, in that genotypes that are rare in our population may offer the lowest absolute reduction in risk (for example, the rare protective allele in *SLC45A2*).

Table 4. Results of association analyses of individual pigmentation-associated SNPs and risk of CMM

		Allelic odds ratio for CMM (*P*-value)			
Locus/variant	SNP	All cases and controls	Subset: 100% Northern European ancestry	Same subset, adjusted for hair, eye and skin color[1]	Same subset, adjusted for MC1R genotype
SLC45A2 *i3*	rs35391	**3.71** (<1.0e-4)[2]	2.49 (0.0014)	1.53 (0.11)	1.68 (0.16)
SLC45A2 *i3*	rs28777	**3.75** (<1.0e-4)	2.37 (0.0012)	1.68 (0.045)	2.28 (0.0028)
SLC45A2*F374L	rs16891982	**3.44** (<1.0e-4)	**2.86** (<1.0e-5)	**1.68** (0.0001)	**1.68** (0.0004)
IRF4 *i4*	rs12203592	1.04 (0.32)	1.18 (0.0075)	1.05 (0.45)	1.03 (0.77)
TYRP1	rs1408799	**1.20** (0.0007)	1.14 (0.0117)	1.166 (0.014)	1.06 (0.56)
TYRP1 *i6*	rs2733832	1.16 (0.0018)	1.12 (0.0243)	1.121 (0.025)	1.06 (0.33)
TYR*R402Q	rs1126809	1.20 (0.002)	**1.13** (0.001)	1.19 (0.0059)	**1.37** (0.0004)
TYR*S192Y	rs1042602	1.05 (0.30)	1.10 (0.10)	1.14 (0.040)	1.17 (0.041)
SLC24A4	rs12896399	1.11 (0.24)	1.12 (0.37)	1.07 (0.30)	1.09 (0.24)
OCA2*R419Q	rs1800407	1.31 (0.004)	1.18 (0.20)	1.23 (0.096)	1.35 (0.022)
HERC2 *i86*	rs12913832	1.15 (0.042)	1.06 (0.51)	1.20 (0.078)	1.21 (0.048)
MC1R*V60L	rs1805005	1.01 (0.80)	1.02 (0.87)	1.02 (0.83)	—
MC1R*D84E	rs1805006	1.61 (0.013)	2.15 (0.004)	1.58 (0.10)	—
MC1R*V92M	rs2228479	1.07 (0.53)	1.12 (0.42)	1.08 (0.62)	—
MC1R*R142H	rs11547464	2.25 (0.055)	3.33 (0.014)	3.85 (0.011)	—
MC1R*R151C	rs1805007	**1.58** (7.0e-4)	1.39 (0.0042)	1.00 (0.99)	—
MC1R*R160W	rs1805008	**1.68** (<1.0e-4)	**1.53** (<1e-4)	1.29 (0.092)	—
MC1R*R163Q	rs885479	1.15 (0.54)	1.04 (0.99)	1.04 (0.83)	—
MC1R*D294H	rs1805009	1.42 (0.019)	1.27 (0.87)	1.01 (0.87)	—
MC1R "R"[3]		**1.65** (<1e-5)	**1.56** (<1e-5)	1.12 (0.1481)	
MC1R "r" [4]		**1.10**	**1.08**	0.95	
ASIP region	rs4911442	**1.49** (<1e-5)	**1.46** (<1e-5)	**1.32** (8.0e-4)	1.39 (0.0037)

ASIP, Agouti signaling protein; CMM, cutaneous malignant melanoma; IRF4, interferon regulatory factor 4; MC1R, melanocortin-1 receptor; OCA2, oculocutaneous albinism type II; SLC45A2, solute carrier family 45, member 2; SLC24A4, solute carrier family 24, member 4; SNP, small nucleotide polymorphisms; TYR, tyrosinase; TYRP1, tyrosinase-related protein 1.
[1]Hair color encoded as four-category categorical variable (fair, light brown, red, and dark); eye color as three-category ordinal trait; and skin color as three-category ordinal trait.
[2]Bold values denote *P* value $< 10^{-3}$ (equivalent to P<0.04 after a Bonferroni's correction for 40 tests).
[3]High-penetrance red hair color MC1R haplotype: variant at D84E, R142H, R151C, R160W, or D294H.
[4]Low-penetrance red hair color MC1R haplotype: variant at V60L, V92M, or R163Q.

This is only the third study reporting an association between *SLC45A2* and CMM. Fernandez *et al.* (2008) genotyped the coding polymorphism rs16891982 (F374L) in a collection of Spanish CMM cases and controls, whereas Guedj *et al.* (2008) report results from a French case–control study. In keeping with our findings, the minor (*L374*) allele was protective (7% frequency in Spanish cases and 16% frequency in Spanish controls; 4% and 10% in the French sample). This protective allele is associated with olive and dark skin, and hence is least common in Northern European populations (1.7% in the Utah-derived CEU HapMap sample), more common in Southern Europeans, but is fixed in African and most Asian populations (Yuasa *et al.*, 2006). In our Australian sample, examination of the three SNP haplotypes and logistic regression analysis showed that rs16891982 is the SNP most strongly associated with melanoma risk in our population.

The variant in the *TYR* most strongly associated with CMM risk by Gudbjartsson *et al.* (2008) was rs1126809 (*R402Q*). We found a significant association with rs1126809, and any association with the intronic SNP rs10765198 disappeared in multivariate models in which rs1126809 was included as a covariate. The *Q402* variant form of *TYR* has 25% of the catalytic activity of the *R402* form at 37 °C (Tripathi *et al.*, 1991), and causes autosomal recessive ocular albinism when co-occurring with more penetrant OCA mutations in compound heterozygote form (Fukai *et al.*, 1995; Hutton and Spritz, 2008). It is therefore a strong candidate for the causative risk variant in this gene. Similar to Gudbjartsson *et al.* (2008), we did not detect any association between rs1042602 (*TYR**S192Y) and CMM in our Australian sample, even though those authors did detect an association with freckling in their Scandinavian sample.

Table 5. Multivariate analysis of CMM versus five polymorphic pigmentation loci, and pigmentation phenotype in individuals of 100% Northern European ancestry (1,062 cases, 1,262 controls)

Variant	Allelic odds ratio (no covariates)	Allelic odds ratios (pigmentation as covariate)	Locus-specific sibling recurrence risk ratio[1] (attributable risk[2])
MC1R ''r''	1.10 (0.94–1.27)	1.09 (0.94–1.27)	1.047 (19.4%)
MC1R ''R''	1.74 (1.49–2.01)	1.46 (1.24–1.72)	
	1.047 (19.4%)		
rs16891982*C SLC45A2	2.47 (1.51–4.07)	2.04 (1.27–3.40)	1.058 (90.6%)
rs1800407*A OCA2*R419Q	1.32 (1.09–1.61)	1.34 (1.10–1.64)	1.008 (5.1%)
rs1126809*A TYR*Q402R	1.24 (1.09–1.41)	1.18 (1.04–1.34)	1.010 (8.0%)
rs4911442*C ASIP	1.44 (1.23–1.69)	1.40 (1.19–1.65)	1.016 (10.8%)
Hair color[3]	—	0.88, 1.50, 0.72, 0.59 (P=1.8e-5)[4]	1.374
Skin color[5]	—	0.98, 0.63 (P=0.007)[4]	
Eye color[6]	—	1.22, 1.05 (P=0.10)[4]	
Total model R^2	5.9%	8.3%	

ASIP, Agouti signaling protein; CMM, cutaneous malignant melanoma; MC1R, melanocortin-1 receptor; OCA2, oculocutaneous albinism type II; SLC45A2, solute carrier family 45, member 2; TYR, tyrosinase.
[1]Assuming a lifetime melanoma risk of 5.2% in the Queensland population.
[2]Proportional reduction in disease rate if entire population carried lowest risk genotype.
[3]Five categories (fair, light brown, red, dark brown, and black); reference category, fair.
[4]*P*-value from sequential likelihood ratio test.
[5]Three categories (light, medium, and dark); reference category, light.
[6]Three categories (blue/grey, green/hazel, and brown/black); reference category, blue/grey.

In the case of *OCA2*, we found that the SNP (rs12913832), most strongly associated with blue eye color in our sample (Sturm *et al.*, 2008), was not associated with CMM, especially after adjusting for ancestry. Rather, the less common coding mutation *R419Q* (rs1800407) was a significant risk factor, notably in the multivariate analysis adjusted for other pigmentation loci. This association has also been seen in a Spanish data set (Fernandez *et al.*, 2009) with similar effect sizes. In the French case–control study of Jannot *et al.* (2005), *R419Q* was not a significant risk factor for melanoma, but the nearby *R305W* and neighboring intronic SNPs were. The *R305W* and *R419Q* variants are in complete linkage disequilibrium in our population. The *OCA2* variants screened by Gudbjartsson *et al.* (2008) were not associated with CMM.

The *TYRP1* (*OCA3*) variants we genotyped have been previously shown to be associated with eye and hair color (Sulem *et al.* 2008), and with CMM risk (Gudbjartsson *et al.*, 2008). We confirmed these associations and effect size estimates. These variants differ significantly in frequency between the main HapMap ethnic samples, and we could show a gradient with a proportion of Northern European ancestry paralleling hair and eye color. We did not detect specific interactions between *TYR* and *TYRP1* in their effects on melanoma risk or on pigmentation phenotypes.

The two recent reports of association of SNPs located in the region of *ASIP* with CMM risk (Gudbjartsson *et al.*, 2008; Brown *et al.*, 2008) implicated widely separated SNPs. We observed that the association of CMM with the *ASIP* region closely paralleled that with red hair color, and seems to be largely due to a single long (~1.8 Mb) haplotype that straddles a number of haplotype block boundaries. This is likely to represent the effects of selection along with a relatively recent origin, as has been observed around other pigmentation loci (Barreiro *et al.*, 2008), given that this particular haplotype is also associated with lighter hair and skin color.

Given the role of *ASIP* as an antagonist of the melanocortin pathway, it seems biologically plausible that there would be an interaction between *MC1R* genotype and *ASIP* genotype. For red hair color, this was markedly so, with the rs4911442*C allele being roughly equivalent to an ''r'' MC1R haplotype. This is most easily seen by comparing the penetrances of the joint ''R/w; C/T'' genotype with the ''R/r; T/T'' genotype, and in the linear dose–response relationship between number of rs4911442*C alleles on an ''R/r'' back-

Table 6. Interaction between rs4911442 (near ASIP) and *MC1R* effects on red hair color, and CMM risk

Joint genotype		Cutaneous Melanoma			Red hair		
MC1R	rs4911442	Cases	*N*	Odds ratio	Red	Total	Penetrance
R/R	C/C	1	2	5.00	1	2	0.500
	C/T	17	34	5.00	30	38	0.789
	T/T	45	94	4.59	73	108	0.676
R/r	C/C	3	11	1.88	5	12	0.417
	C/T	27	60	4.09	12	74	0.162
	T/T	73	226	2.55	13	250	0.052
R/w	C/C	2	5	3.33	0	6	0.000
	C/T	43	112	3.12	8	140	0.057
	T/T	99	343	2.03	13	412	0.032
r/r	C/C	3	5	7.50	0	5	0.000
	C/T	20	35	6.67	2	39	0.051
	T/T	19	83	1.48	0	100	0.000
r/w	C/C	2	16	0.71	0	19	0.000
	C/T	33	127	1.76	3	148	0.020
	T/T	74	323	1.49	2	417	0.005
w/w	C/C	2	2	24.81	0	3	0.000
	C/T	24	101	1.56	0	135	0.000
	T/T	52	312	1.00	0	412	0.000

ASIP, Agouti signaling protein; CMM, cutaneous malignant melanoma; MC1R, melanocortin-1 receptor

ground (Table 6). In the same fashion, MC1R ''R'' haplotypes were dominant to rs4911442*C. Therefore, *ASIP* is a red hair color locus, in the same way as *MC1R*. In keeping with this, rs4911442 genotype was a significant predictor of facial freckling ($P=0.0001$), as noted by others (Gudbjartsson *et al.*, 2008). A similar pattern was observed for CMM, but did not reach formal statistical significance. The association signal for red hair color and CMM across chromosome 20 did track quite strongly, so we expect that the epistatic pattern seen for these two loci with respect to hair color would also be demonstrable for CMM given a larger sample size.

An interesting negative finding in this study is the absence of an association between *IRF4* genotype and melanoma risk, despite the very strong evidence for association of this locus with skin coloring in this and other studies (Han *et al.*, 2008; Sulem *et al.*, 2008). This was also noted by Gudbjartsson *et al.* (2008), and is in some ways the obverse of our previous observation (Palmer *et al.*, 2000) that some pigmentation loci, such as MC1R, seem to affect CMM risk by pathways other than that mediated by measurable changes in skin coloring or tanning ability. Here, we observe an effect on coloring, which one would expect to significantly influence response of the organism to exposure to ultraviolet light, but appears neutral with respect to CMM risk. In the case of *OCA2*, we and others have previously described how a single SNP rs12913832 strongly predicts eye color (Sturm *et al.*, 2008; Kayser *et al.*, 2008; Eiberg *et al.*, 2008), but has little or no effect on CMM risk. We do see an effect of another *OCA2* polymorphism rs1800407 on CMM risk, but this is less impressive in terms of statistical significance.

A peculiarity of these analyses is that most of these pigmentation loci markedly differ in allele frequencies between different ethnic groups, in a way that parallels the differences in CMM rates between these groups. The simplest hypothesis, given that we have shown that all the selected variants are associated with pigmentation, is that these variants underlie ethnic variation in CMM risk. To definitively exclude ethnic confounding as a cause of the observed association between ancestry informative variants and disease risk, we would need either a cohort design, in which we could show that the rate of CMM in carriers of all low-risk variants is comparable with that seen in ethnic groups of known low ''intrinsic'' risk, or a family-based design, in which we could also show linkage of these variants. Unfortunately, the age of onset of CMM in our sample and sampling design means that we cannot pursue the latter strategy, although we can show significant transmission-disequilibrium test results for pigmentation phenotypes, such as skin coloring.

In conclusion, the variants in pigmentation genes that we describe explain ~40% of the variation in familial risk of melanoma ascribable to phenotypes, such as skin, hair, and eye color. In the cases of *TYR*, *SLC45A2*, *OCA2*, and *MC1R*, it is likely that the major causative risk variants have been identified, but this is still open for *ASIP*, and other minor risk variants at all loci await characterization.

MATERIALS AND METHODS

Study participants

Cutaneous malignant melanoma cases were a stratified sample of all cases of CMM diagnosed in the state of Queensland between the period 1982 and 1990, as described in detail elsewhere (Baxter *et al.*, 2008). These individuals were originally studied from 1991 to 1994, but were recontacted and interviewed between 2002 and 2004.

The controls for the present analysis came from the Brisbane Twin Nevus Study. As described in detail elsewhere (Sturm *et al.*, 2008), adolescent twins, their siblings, and parents have been recruited over 16 years into an ongoing study of genetic and environmental factors contributing to the development of pigmented nevi and other risk factors for skin cancer. The proband twins are recruited at the age of 12 years through schools around Brisbane, Australia, and followed up at the age of 14 years. All controls are screened to be unaffected by CMM. The sample is overwhelmingly (>95%) of Northern European origin (mainly Anglo-Celtic). All cases and controls gave informed consent to participate in this study,

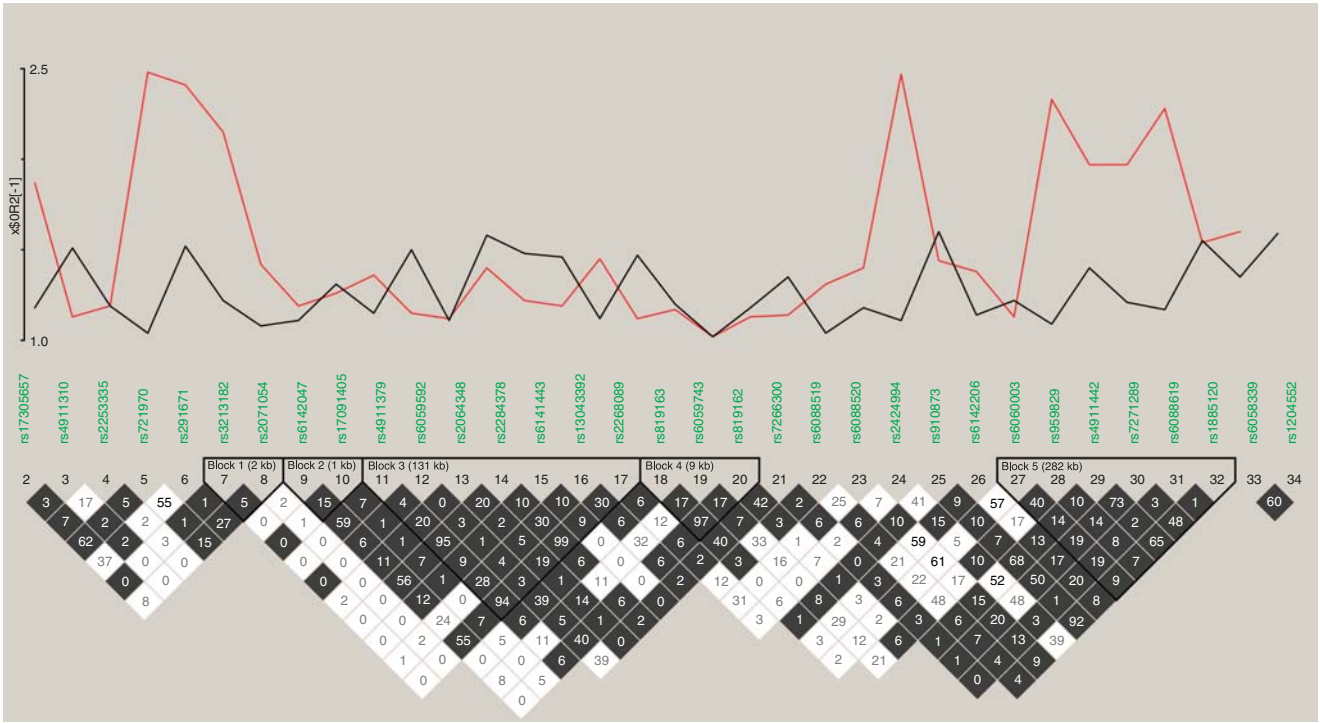

Figure 1. Odds ratios for red hair color (red line) and cutaneous malignant melanoma (black line), and pairwise linkage disequilibrium (r2, with colors generated using the ''4 gamete rule'') versus SNPs around Agouti signalling protein (ASIP). CMM, cutaneous malignant melanoma.

and the study protocol was approved by appropriate institutional review boards and adhered to the Helsinki Guidelines.

Ancestry was measured by questioning either the individual or a parent about the country of birth and ancestry of each of the grandparents of the individual. As most subjects were of European ancestry, we have constructed three ancestry scores on the basis of the proportion of grandparents of Northern European (British, Scandinavian, Danish, Dutch, German, and French), Southern European (Spanish, Italian, and Greek), and Eastern European ancestry. In addition, the Northern indicator was supplemented by ''Celtic'' (Ireland, Scotland, and Wales), English, and Continental indicators, to test whether there were further substructure effects. Ancestry data is available for all cases and controls included in these analyses.

Genotyping and statistical analysis

We selected SNPs for association testing based on association with pigmentation phenotypes in our own and other studies (Palmer *et al.*, 2000; Sturm *et al.*, 2008; Sulem *et al.*, 2007; Han *et al.*, 2008). In the case of SNPs near *ASIP* (Brown *et al.*, 2008), we had observed maximal association with CMM at SNPs away from *ASIP* itself, hence the two SNPs reported to be most strongly associated to CMM in the study of Gudbjartsson *et al.* (2008) were not genotyped. We did not genotype within *SLC24A5* (''Golden''), given that the variant associated with human skin color is almost monomorphic in whites. SNPs were typed using iPLEX Gold chemistry on a MALDI-TOF Mass Spectrometer (Sequenom, San Diego, CA). PCR was carried out using a 2.5-μl reaction mixture in standard 384-well plates. PCR was performed using 10 ng genomic DNA, 0.5 unit of *Taq* polymerase (HotStarTaq, Qiagen, Valencia, CA), 500 μmol of each dNTP, and 100 nmol of each PCR primer**.** PCR thermal cycling in an ABI-9700 instrument was 15 minutes at 94 °C, followed by 45 cycles of 20 sec at 94 °C, 30 seconds at 56 °C, and 60 seconds at 72 °C. To the completed PCR mixture, 1 μl containing 0.15 units of Shrimp alkaline phosphatase was added and the mixture incubated for 30 minutes at 37 °C followed by inactivation for 5 minutes at 85 °C. After adjusting the concentrations of extension primers to equilibrate signal-to-noise ratios, the post-PCR primer extension reaction of the iPLEX assay was performed in a final 5 μl volume extension reaction containing 0.1 μl of termination mix, 0.02 μl of DNA polymerase (Sequenom), and 600–1,200 nM extension primers. A two-step 200 short cycles program was used for the iPLEX reaction: initial denaturation was 30 seconds at 94 °C followed by five cycles of 5 seconds at 52 °C and 5 seconds at 80 °C. An additional 40 annealing and extension cycles were then looped back to 5 seconds at 94 °C, 5 seconds at 52 °C, and 5 seconds at 80 °C. The final extension was carried out at 72 °C for 3 minutes and the sample was cooled to 20 °C. The iPLEX reaction products were desalted by diluting samples with 15 μl of water and adding 3 μl of resin, and centrifuged to remove resin. The products were spotted on a SpectroChip (Sequenom), processed, and analysed in a Compact Mass Spectrometer using MassARRAY Workstation (version 3.3) software (Sequenom). As genotyping has been carried out in various subsets of the data, the number of genotyped cases and controls varies from marker to marker. Specifically, some markers are genotyped in the entire collection, whereas others have been genotyped only in a nested sample of unrelated cases and controls.

As some cases (105 individuals) are from 48 multiplex families and all our controls come from twin families, the association analyses were carried out using Sib-pair 1.0 (Duffy, 2008). The association analysis in Sib-pair implements a logistic regression penetrance analysis of measured genotypes within pedigrees that may include monozygotic twins, and uses Monte Carlo (gene-dropping) simulation

to obtain *P*-values adjusted for the relationships of cases and controls. Other analyses have been carried out in the *R* statistical analysis environment (R Core Development Team, 2008), using especially the *haplo.stats* package (Sinnwell *et al.*, 2007), and also using Haploview (Barrett *et al.*, 2005) to generate plots and long-range haplotypes. We express the overall contribution of the measured loci to risk of melanoma in logistic regression analyses as the Nagelkerke R^2 (Nagelkerke 1991), and have also calculated the locus-specific sibling recurrence risk (James, 1971; Risch, 1990) predicted for that locus under an assumed lifetime risk of 5.2% (Queensland Cancer Registry, 2008). This further assumes that a proportional hazards model is appropriate, but is quite robust to the assumed lifetime risk.

CONFLICT OF INTEREST

The authors state no conflict of interest.

ACKNOWLEDGMENTS

This work was supported by grants from the National Health and Medical Research Council of Australia (NHMRC 241944, 389875, 389891) and the National Cancer Institute (CA88363). RAS, DLD, NKH, and GWM are NHMRC Senior Research Fellows. We thank Anjali Henders, Megan Campbell, Shane Thomas, and Mitchell Stark for assistance with sample preparation and genotyping. This study used data from the SNP Database at the NINDS Human Genetics Resource Center DNA and Cell Line Repository (http://ccr.coriell.org/ninds/).

SUPPLEMENTARY MATERIAL

Supplementary material is linked to the online version of the paper at http://www.nature.com/jid

REFERENCES

Aitken J, Welch J, Duffy DL, Milligan A, Green A, Martin NG *et al.* (1999) CDKN2A variants in a population-based sample of Queensland families with melanoma. *Journal of the National Cancer Institute* 91:446–52

Barreiro LB, Laval G, Quach H, Patin E, Quintana-Murci L (2008) Natural selection has driven population differentiation in modern humans. *Nat Genet* 40:340–5

Barrett JC, Fry B, Maller J, Daly MJ (2005) Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 21:263–5

Baxter AJ, Hughes MC, Kvaskoff M, Siskind V, Shekar S, Aitken JF *et al.* (2008) The Queensland Study of Melanoma: environmental and genetic associations (Q-MEGA); study design, baseline characteristics, and repeatability of phenotype and sun exposure measures. *Twin Res Hum Genet* 11:183–96

Box NF, Duffy DL, Chen W, Stark M, Martin NG, Sturm RA *et al.* (2001) MC1R genotype modifies risk of melanoma in families segregating CDKN2A mutations. *Am J Hum Genet* 69:765–73

Brown KM, Macgregor S, Montgomery GW, Craig DW, Zhao ZZ, Iyadurai K *et al.* (2008) Common sequence variants on 20q11.22 confer melanoma susceptibility. *Nat Genet* 40:838–40

Do K-A, Aitken JF, Green AC, Martin NG (2004) Analysis of melanoma onset: assessing familial aggregation by using estimating equations and fitting variance components via Bayesian random effects models. *Twin Res* 7:98–113

Duffy DL (2008) Sib-pair (computer program). Version 1.0b. URL: http://www.qimr.edu.au/davidD#sib-pair (accessed on 7 July 2008)

Eiberg H, Troelsen J, Nielsen M, Mikkelsen A, Mengel-From J, Kjaer KW *et al.* (2008) Blue eye color in humans may be caused by a perfectly associated founder mutation in a regulatory element located within the HERC2 gene inhibiting OCA2 expression. *Hum Genet* 123:177–87

Fernandez LP, Milne RL, Pita G, Avilés JA, Lázaro P, Benítez J *et al.* (2008) SLC45A2: a novel malignant melanoma-associated gene. *Hum Mutat* 29:1161–7

Fernandez LP, Milne RL, Pita G (2009) Pigmentation-related genes and their implication in malignant melanoma susceptibility. *Exp Dermatol* 18:634–42

Fukai K, Holmes SA, Lucchese NJ, Siu VM, Weleber RG, Schnur RE *et al.* (1995) Autosomal recessive ocular albinism associated with a functionally significant tyrosinase gene polymorphism. *Nat Genet* 9:92–5

Guedj M, Bourillon A, Combadières C, Rodero M, Dieudé P, Descamps V *et al.* (2008) Variants of the MATP/SLC45A2 gene are protective for melanoma in the French population. *Hum Mutat* 29:1154–60

Gudbjartsson DF, Sulem P, Stacey SN, Goldstein AM, Rafnar T, Sigurgeirsson B *et al.* (2008) ASIP and TYR pigmentation variants associate with cutaneous melanoma and basal cell carcinoma. *Nat Genet* 40:886–91. (doi:10.1038/ng.161)

Han J, Kraft P, Nan H, Guo Q, Chen C, Qureshi A *et al.* (2008) A genome-wide association study identifies novel alleles associated with hair color and skin pigmentation. *PLoS Genet* 4:e1000074

Hutton SM, Spritz RA (2008) A comprehensive genetic study of autosomal recessive ocular albinism in caucasian patients. *Invest Ophthalmol Vis Sci* 49:868–72

Jannot AS, Meziani R, Bertrand G, Gérard B, Descamps V, Archimbaud A *et al.* (2005) Allele variations in the OCA2 gene (pink-eyed-dilution locus) are associated with genetic susceptibility to melanoma. *Eur J Hum Genet* 13:913–20

James JW (1971) Frequency in relatives for an all-or-none trait. *Ann Hum Genet* 35:47–9

Kayser M, Liu F, Janssens AC, Rivadeneira F, Lao O, van Duijn K *et al.* (2008) Three genome-wide association studies and a linkage analysis identify HERC2 as a human iris color gene. *Am J Hum Genet* 82:411–23

Nagelkerke NJD (1991) A note on a general definition of the coefficient of determination. *Biometrika* 78:691–2

Palmer JS, Duffy DL, Box NF, Aitken JF, O'Gorman LE, Green AC. *et al.* (2000) MC1R polymorphisms and risk of melanoma: is the association explained solely by pigmentation phenotype? *Am J Hum Genet* 66:176–86

Queensland Cancer Registry (2008) *Cancer in Queensland: Incidence and Mortality 1982–2005*. Queensland: The Cancer Council Queensland

R Development Core Team (2008) R: A language and environment for statistical computing. Version 2.7.1. Vienna, Austria: R Foundation for Statistical Computing. ISBN 3-900051-07-0, URL http://www.R-project.org

Risch N (1990) Linkage strategies for genetically complex traits. I. Multilocus models. *Am J Hum Genet* 46:222–8

Simon-Sanchez J, Scholz S, Fung HC, Matarin M, Hernandez D, Gibbs JR *et al.* (2007) Genome-wide SNP assay reveals structural genomic variation, extended homozygosity and cell-line induced alterations in normal individuals. *Hum Mol Genet* 16:1–14

Sinnwell JP, Schaid DJ and with contributions from Zhaoxia Yu (2007) haplo.stats: Statistical analysis of haplotypes with traits and covariates when linkage phase is ambiguous. R package version 1.3.0. http://mayoresearch.mayo.edu/mayo/research/schaid_lab/software.cfm

Sturm RA, Duffy DL, Box NF, Newton RA, Shepherd AG, Chen W *et al.* (2003) Genetic association and cellular function of MC1R variant alleles in human pigmentation. *Ann N Y Acad Sci* 994:348–58

Sturm RA, Duffy DL, Zhao ZZ, Leite FP, Stark MS, Hayward NK *et al.* (2008) A single SNP in an evolutionary conserved region within intron 86 of the HERC2 gene determines human blue-brown eye color. *Am J Hum Genet* 82:424–31

Sulem P, Gudbjartsson DF, Stacey SN, Helgason A, Rafnar T, Jakobsdottir M. *et al.* (2008) Two newly identified genetic determinants of pigmentation in Europeans. *Nat Genet* 40:835–7

Sulem P, Gudbjartsson DF, Stacey SN, Helgason A, Rafnar T, Magnusson KP. *et al.* (2007) Genetic determinants of hair, eye and skin pigmentation in Europeans. *Nat Genet* 39:1443–52

Tripathi RK, Giebel LB, Strunk KM, Spritz RA (1991) A polymorphism of the human tyrosinase gene is associated with temperature-sensitive enzymatic activity. *Gene Expr* 1:103–10

Yuasa I, Umetsu K, Harihara S, Kido A, Miyoshi A, Saitou N *et al.* (2006) Distribution of the F374 allele of the SLC45A2 (MATP) gene and founder-haplotype analysis. *Ann Hum Genet* 70:802–11