Identification of Susceptibility Loci for Cutaneous Squamous Cell Carcinoma

Maryam M. Asgari1,2, Wei Wang3, Nilah M. Ioannidis3,4, Jacqueline Itnyre3, Thomas Hoffmann5, Eric Jorgenson5 and Alice S. Whittemore3

We report a genome-wide association study of cutaneous squamous cell carcinoma conducted among non-Hispanic white members of the Kaiser Permanente Northern California health care system. The study includes a genome-wide screen of 61,457 members (6,891 cases and 54,566 controls) genotyped on the Affymetrix Axiom European array and a replication phase involving an independent set of 6,410 additional members (810 cases and 5,600 controls). Combined analysis of screening and replication phases identified 10 loci containing single-nucleotide polymorphisms (SNPs) with \( P \)-values \(< 5 \times 10^{-8} \). Six loci contain genes in the pigmentation pathway; SNPs at these loci appear to modulate squamous cell carcinoma risk independently of the pigmentation phenotypes. Another locus contains HLA class II genes studied in relation to elevated squamous cell carcinoma risk following immunosuppression. SNPs at the remaining three loci include an intronic SNP in FOXP1 at locus 3p13, an intergenic SNP at 3q28 near TP63, and an intergenic SNP at 9p22 near BNC2. These findings provide insights into the genetic factors accounting for inherited squamous cell carcinoma susceptibility.


INTRODUCTION

Squamous cell carcinoma (SCC) is among the most common and costly malignancies in populations of European ancestry (Housman et al., 2003). Its primary cause is ultraviolet radiation exposure, which causes DNA damage in keratinocytes (Dessinioti et al., 2011; Roewert-Huber et al., 2007), for which melanin provides a protective filter. An inherited basis for SCC risk is supported by the increased risk among first-degree relatives of SCC cases (Hemminki et al., 2003; Hussain et al., 2009), but the specific genetic factors that determine susceptibility are not well understood. Although genome-wide association studies (GWASs) have identified susceptibility loci for other skin cancers, for example, cutaneous malignant melanoma (CMM) (Barrett et al., 2011; Bishop et al., 2009), basal cell carcinoma (BCC) (Nan et al., 2011; Stacey et al., 2008), and BCC and SCC combined as nonmelanoma skin cancer (NMSC) (Stacey et al., 2009), we are unaware of previous GWASs focused solely on SCC. Although the GWAS of NMSC investigated the single-nucleotide polymorphisms (SNPs) identified in the combined analysis for their individual effects on BCC and SCC, the power to detect SNPs specifically related to SCC was limited by the number of SCC cases. We describe an internally validated GWAS based on data from 67,867 non-Hispanic white (NHW) individuals (7,701 SCC cases and 60,166 controls) enrolled in the Kaiser Permanente Research Program on Genes, Environment, and Health (RPGEH).

RESULTS

We identified 10 loci containing SNPs with combined \( P \)-values meeting the genome-wide threshold of \( 5 \times 10^{-8} \) (Pe’er et al., 2008) (Figure 1). Table 1 shows per-allele odds ratios (ORs) and \( P \)-values for the screening phase, replication phase, and combined data for the most significant SNPs at these 10 loci. Six of the 10 loci encompass genes that play established roles in the pigmentation pathway (Scherer and Kumar, 2010); SNPs at these loci have been associated with skin cancers and/or pigment-related phenotypes such as eye, hair or skin color, tanning, burning, or freckles (Gerstenblith et al., 2010).

Pigmentation loci

The first three pigment-related SNPs in Table 1 (section A) have been associated previously with SCC, BCC, and CMM. SNP rs16891982 at locus 5p13, a nonsynonymous SNP (Phe374Leu) in SLC45A2, has been associated with SCC (Stacey et al., 2009), BCC (Stacey et al., 2009), and CMM (Barrett et al., 2011; Duffy et al., 2010b; Fernandez et al., 2008; Guedj et al., 2008; Ibarrola-Villava et al., 2012; Kosiniak-Kamysz et al., 2014; Stacey et al., 2009) as well as eye, hair, and skin color (Branicki et al., 2009; Duffy et al., 2010b; Eriksson et al., 2010; Liu et al., 2015; Soejima and Koda, 2007; Stokowski et al.). SLC45A2 encodes a
membrane-associated transporter enzyme, and loss of SLC45A2 activity has been found to disrupt post-Golgi-level trafficking of tyrosinase to the melanosomes (Newton et al., 2007) where melanin is synthesized and stored. Mutations in SLC45A2 cause type four oculocutaneous albinism (OCA) syndrome, which is characterized by failure to synthesize melanin (Simeonov et al., 2013). SNP rs12203592 at locus 6p25, an intronic SNP in IRF4 (Praetorius et al., 2013), has been associated with SCC, BCC, and CMM (Bishop 2009; Duffy et al., 2010b; Han et al., 2011; Stefanaki et al., 2013; Zhang et al., 2013), and with pigmentation phenotypes (Duffy et al., 2010b; Han et al., 2008; Jacobs et al., 2015; Nan et al., 2009a; Soejima and Koda, 2007; Sulem et al., 2007; Visser et al., 2015). The T allele of this SNP reduces expression of IRF4 (Praetorius et al., 2013), which encodes a transcription factor that, in cooperation with MITF, activates expression of TYR and is used by melanocytes to produce and store melanin. SNP rs1126809 at locus 11q14 is a nonsynonymous SNP (Arg402Gln) in TYR, which encodes the enzyme tyrosinase that catalyzes multiple steps in the conversion of tyrosine to melanin. SNP rs1126809 has been associated with SCC, BCC (Nan et al., 2011), and CMM (Bishop 2009; Duffy et al., 2010b; Gudbjartsson et al., 2008; Hu et al., 2011; Ibarrola-Villava et al., 2012; Nan et al., 2011).

Table 1. Genome-wide association and replication for 10 SCC loci

<table>
<thead>
<tr>
<th>Locus</th>
<th>SNP</th>
<th>Gene</th>
<th>Minor allele</th>
<th>MAF 3</th>
<th>Info 4</th>
<th>OR5 CI5 P-value6</th>
<th>OR CI P-value</th>
<th>OR CI P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chr 11q14</td>
<td>rs1126809</td>
<td>TYR</td>
<td>A</td>
<td>0.28</td>
<td>1.00 1.19 1.16–1.25</td>
<td>0.08</td>
<td>0.96–1.22 2.09×10−4</td>
<td>1.19</td>
</tr>
<tr>
<td>Chr 15q13</td>
<td>rs12916300</td>
<td>HERC2</td>
<td>C</td>
<td>0.26</td>
<td>1.00 0.89 0.85–0.93</td>
<td>0.82</td>
<td>0.72–0.93 2.40×10−3</td>
<td>0.88</td>
</tr>
<tr>
<td>Chr 16q24</td>
<td>rs4268748</td>
<td>DEF8</td>
<td>C</td>
<td>0.26</td>
<td>0.85 1.34 1.28–1.40</td>
<td>3.24×10−4</td>
<td>1.28</td>
<td>1.13–1.45 7.85×10−3</td>
</tr>
<tr>
<td>Chr 20q11</td>
<td>rs6039655</td>
<td>RALY</td>
<td>A</td>
<td>0.08</td>
<td>0.99 1.30 1.22–1.38</td>
<td>5.18×10−5</td>
<td>1.49</td>
<td>1.25–1.78 1.14×10−4</td>
</tr>
</tbody>
</table>

Abbreviations: SCC, squamous cell carcinoma; SNP, single-nucleotide polymorphisms.

1Regions containing SNPs with combined P-values < 5 × 10−8.
2SNP with strongest combined P-value in the region.
3MAF = minor allele frequency among control subjects.
4For imputed data, the info is the IMPUTE-2 information measure for imputation accuracy (Marchini and Howie, 2010).
5OR = odds ratio per minor allele, CI = 95% confidence interval.
6Based on the squared Cochran-Armitage trend test.

Two flanking genes of an intergenic SNP.
The remaining three SNPs in Table 1 (section A) lie near genes that play established roles in the pigmentation pathway and that have been associated with pigmentation traits and/or skin cancer risk. However, the relationships of these SNPs to SCC risk are uncertain. For example, SNP rs12916300 at 15q13 (Supplementary Figure S1 online) lies within intron 68 of HERC2, which encodes a ubiquitin-protein ligase recruited to sites of DNA damage from ionizing radiation, and which has been associated with BCC (Han et al., 2011) and pigmentation phenotypes (Eiberg et al., 2008; Han et al., 2011). This SNP is in high linkage disequilibrium ($R^2 = 0.85$) with rs12913832, another intronic SNP in HERC2 that has been associated with pigmentation phenotypes (Branicki et al., 2009; Eiberg et al., 2008; Han et al., 2008; Nan et al., 2009b; Sturm et al., 2008) and that modifies expression of OCA2, a nearby gene encoding a melanosomal enzyme needed for melanin synthesis (Visser et al., 2012). However, SNP rs12916300 remained associated with SCC risk with genome-wide significance after adjustment for rs12913832, suggesting that its effects are not merely due to its correlation with rs12913832. Currently, there is no specific evidence that rs12916300 disrupts any transcription factor binding sites, although it is located two base pairs from the binding motif of the transcription factor FOXI1 (Pique-Regi et al., 2011, and http://regulomedb.org/snp). In data from the GTEx Consortium (2015), rs12916300 was not associated with OCA2 expression in sun-exposed skin ($P = 0.056$ compared with a $P$-value of 0.013 for rs12913832) or in non—sun-exposed skin ($P = 0.44$), even without adjustment for linkage disequilibrium with rs12913832. None of the established pigment-related SNPs in HERC2/OCA2 met genome-wide significance for association in these data, despite their imputation with high accuracy (Supplementary Table S1a online). However, nine additional SNPs (three in HERC2 and six in OCA2) achieved genome-wide significance after adjustment for all 10 SNPs in Table 1 (Supplementary Table S1b). Further research is needed to clarify the roles in SCC risk of these nine SNPs, whose pairwise correlations range from weak to strong (Supplementary Figure S2 online).

SNP rs4268748 at 16q24 is intronic in DEF8, which encodes an activator of intracellular signal transduction. This SNP is associated with expression levels of CDK10 (a gene critical for cell cycle progression) in sun-exposed skin ($P = 2.3 \times 10^{-11}$) and in nonexposed (supra-pubic) skin ($P = 1.4 \times 10^{-8}$) (GTEx Consortium, 2015). The role of rs4268748 in SCC is complicated by its proximity (39 kb) to MC1R, a gene that encodes a receptor for melanocyte stimulating hormone and that contains several nonsynonymous
SNPs associated with skin pigmentation (Liu et al., 2015; Valverde et al., 1995) and skin cancer (Barrett et al., 2011; Bishop 2009; Box et al., 2001; Duffy et al., 2010b; Ibarrola-Villava et al., 2012; Palmer et al., 2000; Sulem et al., 2007) (Supplementary Table S2 online). In particular, SNPs in MC1R meeting genome-wide significance in this study include rs258322, which has been associated with CMM (Barrett et al., 2011; Bishop 2009; Stefanaki et al., 2013) and pigment-related traits (Valverde, 1995), and the nonsynonymous SNPs rs1805007 (Arg151Cys) and rs1805008 (Arg160Trp), which have been associated with CMM (Duffy et al., 2010b; Ibarrola-Villava et al., 2012; Palmer et al., 2000; Sulem et al., 2007), NMSC (Box et al., 2014), and, among renal transplant patients, SCC (Andresen et al., 2013). However, none of these three SNPs maintained genome-wide significance after adjusting for genotypes at rs4268748, which has been associated with MC1R expression with borderline statistical significance ($P = 0.012$) in sun-exposed skin (GTEx Consortium, 2015). Furthermore, two additional SNPs at this locus (rs35063026 in an untranslated region of SPATA33, and rs78703231 in an intron in SPIRE2) were independently associated with SCC risk after adjustment for rs4268748 (Figure 2 and Supplementary Materials online). These three SNPs (rs4268748, rs35063026, and rs78703231) lie at most 249 kb downstream of MC1R, and their SCC associations could reflect correlation with one or more of the nonsynonymous MC1R SNPs previously associated with pigmentation and/or skin cancer risk. Indeed, after adjustment for the subset of these SNPs with minor allele frequencies greater than 1% in this study (Supplementary Materials and Table S3 online), rs78703231 lost statistical significance, with pre- and post-adjustment risk ratios of 1.21 ($P = 1.77 \times 10^{-17}$) and 1.05 ($P = 8.21 \times 10^{-8}$). However, strong associations persisted for both rs4268748 (postadjustment risk ratio $= 2.12$, $P = 1.96 \times 10^{-15}$) and rs35063026 (postadjustment risk ratio $= 1.33$, $P = 5.38 \times 10^{-11}$). These findings suggest the existence of multiple independent SCC susceptibility loci in the 16q24 region whose precise roles in SCC pathogenesis remain to be elucidated.

Finally, SNP rs6059655 at 20q11, which is intronic in RALY, lies 182 kb upstream of ASIP, a gene that encompasses SNPs associated with pigmentation traits (Jacobs et al., 2015; Liu et al., 2015), CMM (Amos et al., 2011; Barrett et al., 2011; Brown et al., 2008; Duffy et al., 2010b; Maccioni et al., 2013), and NMSC (Binstock et al., 2014; Lin et al 2011; Nan et al., 2009a; Palmer et al., 2000) (Supplementary Figure S3 online). Only three of these SNPs (rs4911442, rs910873, and rs1885120) met the genome-wide significance threshold for association with SCC in the present data, and none maintained genome-wide significance after adjustment for rs6059655. A haplotype in ASIP containing the T allele of rs4911444 and the G allele of rs1015362 has been associated with pigment-related traits and with CMM and BCC (Gudbjartsson et al., 2008; Sulem et al., 2007). However, adjustment for this haplotype did not eliminate the observed association between SCC and rs6059655 (OR: 1.40, $P = 1.31 \times 10^{-10}$ for rs6059655 and 1.30, $P = 2.0 \times 10^{-10}$ for the haplotype). This finding, coupled with the observed association between rs6059655 and ASIP expression levels in sun-exposed skin ($P = 5.3 \times 10^{-4}$) and non–sun-exposed skin ($P = 4.4 \times 10^{-4}$) (GTEx Consortium, 2015), lends support to the possibility that rs6059655 (or a nearby functional SNP) affects ASIP activity independently of the ASIP haplotype.

In summary, the observed associations between SCC risk and SNPs at 15q13, 16q24, and 20q11 could not be explained by their proximity to previously studied SNPs in the pigmentation and skin cancer genes HERC2/OCA2, MC1R, and ASIP, respectively. Further work on the functions of these SNPs is needed to elucidate their potential causal roles and mechanisms in SCC.

We checked whether subjects’ total counts of risk alleles of the six pigment-related SNPs in Table 1 (section A) (called their risk indices) influence their SCC risk independently of their pigment-related phenotypes (skin color, tanning, and sun sensitivity). Because we lacked self-reported or clinical data on these pigment-related phenotypes, we estimated them using subjects’ total allele counts for SNPs associated with fair skin, poor tanning ability, and tendency to sunburn (see Supplementary Materials and Table S4 online for details on how we assembled these pigmentation scores). Supplementary Table S5 online shows the joint distribution of subjects by tertiles of risk index and of skin phenotype score, with tertiles determined by the distributions in controls. The table also shows joint and marginal ORs relating SCC risk to tertiles of risk index and pigmentation score. Because some SNPs were used to calculate both the risk indices and the pigmentation scores, the ORs may be attenuated toward unity. Nevertheless, consistently with the findings of others (Andresen et al., 2013, Bastiaens et al., 2001; Gerstenblith et al., 2010; Kennedy et al., 2001; Nan et al., 2009a; Palmer et al., 2000), we observed increased SCC risk with increasing risk index, independently of the estimated pigmentation phenotypes. Specifically, Supplementary Table S5 shows strong trends of increased SCC risk with increasing risk index within each tertile of the skin pigmentation scores, as well as overall, after adjustment for pigmentation score. Moreover, the data suggest trends of increasing risk index ORs with increasingly sun-sensitive pigmentation phenotypes, with significance levels of $P = 0.85$, $P = 1.15 \times 10^{-2}$, and $P = 0.002$ for skin color, sun sensitivity, and tanning ability, respectively. We also found similar trends of higher SCC risk with increasing pigmentation score within each tertile of risk index, and overall, after adjustment for risk index, suggesting independent contributions of genetic risk index and pigmentation phenotypes in determining SCC susceptibility.

Other loci

Four additional SCC loci met the genome-wide significance threshold (Table 1, section B), which, to our knowledge, have limited or no relation to the pigmentation pathway. Among the four, the HLA locus at 6p21 was most strongly associated with SCC risk, with a per-allele risk ratio of 1.17 ($P = 1.86 \times 10^{-8}$) for the most significant SNP, rs4455710 (Figure 1). This SNP is intronic in the class II HLA gene HLADQA1 (Supplementary Figure S4 online), which binds peptides derived from antigens that access the endocytic route of antigen-presenting cells and presents them on the cell surface for recognition by CD4 T cells. No other SNP in the region maintained genome-wide significance after adjusting for
rs4455710. This SNP is also strongly associated with expression levels of nearby genes, including HLA-DQA1 ($P = 3.3 \times 10^{-7}$ and $P = 4.2 \times 10^{-6}$ in sun-exposed and nonexposed skin, respectively), HLA-DQB1 ($P = 1.1 \times 10^{-6}$ in sun-exposed skin), and HLA-DRB5 ($P = 4.3 \times 10^{-12}$ and $P = 2.6 \times 10^{-9}$ in sun-exposed and nonexposed skin, respectively) (GTEx Consortium, 2015). A role for HLA antigens and immune response in SCC is supported by findings of elevated SCC risk in immunocompromised individuals (Bouwes Bavinck et al., 1991), by data relating NMSCs to specific HLA alleles (Bouwes Bavinck and Claas, 1994), and by elevated SCC risk in cigarette smokers (Leonardi-Bee et al., 2012), because smoking is associated with changes in circulating markers for immune response (Shiels et al., 2014). However, we did not find strong evidence that the risk ratio per allele of rs4455710 varies by the history of smoking or immunosuppression (Supplementary Table S6 online).

The remaining three SNPs in Table 1 (section B) open avenues for further study. SNP rs62246017 at 3p13 is intronic in FOXP1, which encodes a transcription repressor with important roles in immune response, organ development, and the pathogenesis of epithelial malignancies. Alterations in FOXP1 expression have been implicated in the pathogenesis of squamous cell cancers of the lung and of the head and neck (Feng et al., 2012; Yang et al., 2012).

The intergenic SNP rs6791479 at 3q28 lies 144 kb upstream of TP63, which encodes a member of the p53 family of transcription factors (Poligone et al., 2015). One of its two major isoforms (ΔNp63) is involved in multiple functions during skin development and in adult stem/progenitor cell regulation (Crum and McKeon, 2010; Leonard et al., 2011). TP63 immunostaining has utility for differentiating subtypes of head and neck SCCs. In other tissues, TP63 is helpful in distinguishing poorly differentiated SCC (strong staining) from small cell carcinoma or adenocarcinoma (no staining). SNP rs6791479 is associated with expression of TP63 in non-sun-exposed skin ($P = 0.0036$) (GTEx Consortium, 2015).

Finally, the intergenic SNP rs74664507 at 9p22 lies 144 kb upstream of BNC2, which encodes a zinc-finger protein thought to be involved in mRNA processing (Vanhoutteghem and Djan, 2006) and to function as a transcription factor (Romano et al., 2004). The role of BNC2 in human pigmentation is unclear, despite experimental evidence suggesting that the gene influences pigmentation in mice (Smyth et al., 2006) and zebrafish (Patterson and Parichy, 2013), and despite its elevated expression levels in human epidermal skin samples and melanocytic cell lines from dark-skinned donors (Visser et al., 2014). Functional data indicate that BNC2 expression is influenced by the intergenic SNP rs12350739 located in a BNC2 enhancer (Visser et al., 2014). This finding, coupled with the high linkage disequilibrium at this locus, has suggested that rs12350739 is responsible for the observed associations of intronic SNP rs10756819 with skin color (Jacobs et al., 2013) and of intronic SNP rs215327 with freckling (Eriksson et al., 2010). Interestingly, SNPs rs74664507 and rs12350739 are also in high linkage disequilibrium ($R^2 = 0.84$). Thus not surprisingly, univariate regression of SCC risk against genotypes of rs12350739 showed borderline genome-wide significance ($P = 7.62 \times 10^{-6}$), whereas regression of SCC risk against both SNPs showed attenuated association ($P = 0.04$) for rs74664507 and no association ($P = 0.92$) for rs12350739. Additional replication and functional work are needed to elucidate the role of BNC2 in SCC risk.

Allelic effects of SNPs in Table 1

For each of the SNPs at the 10 susceptibility loci in Table 1, we used a likelihood ratio statistic to evaluate how well allelic effects on SCC risk were captured by a one-degree-of-freedom log-additive logistic model relative to a codominant (two-degree-of-freedom) model. The likelihood ratio statistic has a chi-square distribution with one degree of freedom under the null hypothesis that the log-additive allelic model fits well. We found significant improvement in fit for the codominant model relative to the log-additive model only for SNP rs1126809 in TYR at 11q24 ($P = 3.6 \times 10^{-15}$), SCC odds ratios for heterozygote and homozygote carriers of the A-allele of this SNP were 1.13 (95% confidence interval: 1.07–1.19) and 1.51 (1.39–1.64), respectively.

Genetic and nongenetic interactions

We found no evidence that ORs for the SNPs in Table 1 varied by gender, but some varied by age at RPGEH enrollment, as classified into three categories based on the control tertiles and coded ordinarily. Specifically, we found significant interaction $P$-values for SNPs rs12203592 in IRF4 ($P = 3.52 \times 10^{-4}$), rs1126809 in TYR ($P = 0.04$), and SNP rs4268748 in DEF8 ($P = 0.01$). In all three analyses, the ORs decreases with decreasing age. We also evaluated potential epistatic effects on SCC risk among the 45 pairs of the 10 SNPs in Table 1. We observed evidence for superadditivity (on the log odds scale) of effects at pairs of loci 20q11 and 11q14 ($P = 0.017$) and for subadditivity for five other pairs: 5p13 and 6p25 ($P = 0.013$), 6p21 and 15q13 ($P = 0.023$), 20q11 and 5p13 ($P = 0.032$), and 20q11 and 16q24 ($P = 0.041$) (Supplementary Table S7 online). In addition, we evaluated the effects of SNP rs4455710 at 6p21 by a history of smoking and immunosuppression. Specifically, because immunosuppressed recipients of major organ transplants have elevated SCC risk (Andresen 2013), we examined whether the log odds of SCC varied between subjects with and without a history of immunosuppression, as indicated by a history of organ transplantation, chronic lymphocytic, or HIV infection. Similarly, because cigarette smokers have altered immune status compared with nonsmokers (Shiels et al 2014), we compared SNP ORs in ever- versus never-smokers. Supplementary Table S6 shows no evidence of effect modification at this SNP by either smoking status or history of immunosuppression.

**DISCUSSION**

This large SCC GWAS has confirmed existing associations between SCC risk and three SNPs in genes known to be related to the pigmentation pathway. These include non-synonymous SNPs in SLC45A2 on chromosome 5p13 and in TYR on chromosome 11q14, and a functional intronic SNP in IRF4 on chromosome 6p25. It also has identified previously unreported SCC-associated SNPs at three other loci (15q13, 16q24, and 20q11) containing pigmentation genes associated with skin cancer risk, and we were unable to attribute these associations to correlation with putative causal variants.
in these genes. Finally, we have identified four additional loci previously unrelated to SCC risk.

Strengths of this study include its large sample size and its use of a demographically diverse NHW cohort characterized by comprehensive clinical data. The findings need replication in an NHW population having demographic characteristics and ancestries that differ from those of the Northern California population studied here. If the associations are replicated, functional research will be needed to better understand the roles of the associated SNPs in the pigment-related regions 15q13, 16q24, and 20q11, and in the four nonpigment-related regions. In summary, the present findings can facilitate the development of SCC prediction models that identify subgroups in need of more intensive screening. They also can motivate future investigations of genetic factors that modify ultraviolet radiation exposures and other lifestyle characteristics in conferring SCC risk.

MATERIALS AND METHODS

Study population
Potentially eligible study subjects were RPGEH members who at cohort entry reported NHW race/ethnicity, were at least 18 years of age, and had no diagnostic codes for rare genetic disorders associated with increased SCC risk (e.g., xeroderma pigmentosum, oculocutaneous albinism, dystrophic epidermolysis bullosa, epidermodysplasia verruciformis, Fanconi anemia, dyskeratosis congenita, Rothmund-Thomson syndrome, Bloom syndrome, Werner syndrome, and Ferguson Smith syndrome).

Potentially eligible cases were subjects whose pathology records were consistent with incident SCC (invasive or in situ, excluding anogenital and mucosal SCCs) during the period from survey completion until either 31 December 2012 or departure from the Kaiser Permanente Northern California health care system, whichever occurred first. (See Supplementary Materials for details on SCC case verification by pathology review.) Potentially eligible controls (n = 64,218) were subjects with no RPGEH-survey-reported history of skin cancer and no pathology records consistent with skin cancer in the Kaiser Permanente Northern California database. After quality control was applied to all potentially eligible subjects (as described in Supplementary Materials and Table S8 online), a total of 67,867 subjects were eligible for analysis.

Genotyping
Saliva samples from eligible subjects were genotyped on four custom Affymetrix Axiom arrays optimized for individuals of European, East Asian, Latino, and African American race/ethnicity (Hoffmann et al., 2011a, 2011b, for the other arrays [http://www.ncbi.nlm.nih.gov/pubmed/21903159]). Genotyping and SNP quality control have been described previously (dbGaP phs000674.v1.p1) (Bandera et al., 2015; Hoffmann et al., 2011a, 2011b; Kvale et al., 2015). SNP imputation to the 1000 Genomes Project is described in the Supplementary Materials.

Internal validation
For the screening phase we used the genotypes of the 61,457 subjects (6,891 cases and 54,566 controls) who had been typed on the European array, and for the replication phase we used the remaining 6,410 NHW subjects (810 cases and 5,600 controls) who had been typed erroneously on a non-European array. Supplementary Table S9 online shows the classification of all subjects by case-control status, gender, and study phase (screening or replication). Subjects’ genotypes for untyped SNPs were imputed to the 1000 Genomes Project.

eQTL evaluation
We also examined the public GTEx database for evidence that SNPs of interest regulate the expression of nearby genes in skin tissue (see Supplementary Materials).

Statistical analysis
To adjust for population stratification, we determined ancestry principal components for the combined data from screening and replication phases using the smartpca program in the EIGENSOFT4.2 software package (Hemminki and Czene, 2003), as has been described elsewhere (dbGaP phs000674.v1.p1). For each of the screening and replication phases, we then evaluated SCC association with each SNP’s allelic count using the Armitage-Cochran trend statistic, adjusted for gender, the first 10 principal components of ancestry, and for the genotyping array and reagent kit. We combined regression coefficients and P-values obtained from screening and replication phases using Cochran’s method (Cochran 1954; Zhou et al., 2011), and identified as significant those SNPs with combined P-values < 5 × 10⁻⁸ (Pe’er et al., 2008). We checked for residual population stratification by examining the inflation factor λC (Devlin and Roeder 1999), and Q-Q plots of percentiles of the observed distribution of test statistics versus those expected under the global null hypothesis (see Supplementary Materials). Further adjustment for age and immunosuppressive status resulted in negligible changes in the ORs for the top 10 SNPs in Table 1.

Data accessibility
Genotypes used in this GWAS have been registered with dbGaP (Study Accession: phs000674.v1.p1).

All study procedures were approved by the Institutional Review Board of the Kaiser Foundation Research Institute.

CONFLICT OF INTEREST
The authors state no conflicts of interest.

ACKNOWLEDGMENTS
This research was supported by NIH grant R01CA166672, by the GO grant RC2 AG036607, and by the Wayne and Gladys Valley Foundation, the Ellison Medical Foundation, the Robert Wood Johnson Foundation, and Kaiser Permanente Community Benefit Program. We thank Neil Risch, University of California at San Francisco, and Cathy Schaefer, Kaiser Permanente Northern California Division of Research, for establishing the RPGEH cohort and processing and providing the genetic data. We also thank Julia L. Kay and Dilrini K. Ranatunga, Kaiser Permanente Northern California Division of Research, for their help in documenting and transferring the genetic and nongenetic data and providing the background information needed for the analysis.

SUPPLEMENTARY MATERIAL
Supplementary material is linked to the online version of the paper at www.jidonline.org, and at http://dx.doi.org/10.1016/jclid.2016.01.013.

REFERENCES


Vanhouetteghem A, Dijan P. Basonuclins 1 and 2, whose genes share a common origin, are proteins with widely different properties and functions. Proc Natl Acad Sci USA 2006; 103:12421–8.


