Identification of Susceptibility Loci for Cutaneous Squamous Cell Carcinoma

Maryam M. Asgari[1,2], Wei Wang[3], Nilah M. Ioannidis[3,4], Jacqueline Itnyre[3], Thomas Hoffmann[5], Eric Jorgenson[2] and Alice S. Whittemore[3]

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.

Journal of Investigative Dermatology (2016) **136,** 930–937; doi:10.1016/j.jid.2016.01.013

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).

[1]*Department of Dermatology, Massachusetts General Hospital, Boston, Massachusetts, USA;* [2]*Division of Research, Kaiser Permanente Northern California, Oakland, California, USA;* [3]*Department of Health Research and Policy, Stanford University School of Medicine, Stanford, California, USA;* [4]*Department of Genetics, Stanford University School of Medicine, Stanford, California, USA; and* [5]*Department of Epidemiology and Biostatistics and Institute for Human Genetics, University of California, San Francisco, California, USA*

Correspondence: Alice S. Whittemore, Department of Health Research and Policy, Stanford University School of Medicine, Stanford, California 94305, USA. E-mail: alicesw@stanford.edu

Abbreviations: BCC, basal cell carcinoma; CMM, cutaneous malignant melanoma; GWAS, genome-wide association study; LD, linkage disequilibrium; NHW, non-Hispanic white; NMSC, nonmelanoma skin cancer; OCA, oculocutaneous albinism; ORs, odds ratios; RPGEH, Research Program on Genes, Environment, and Health; SCC, squamous cell carcinoma; SNPs, single-nucleotide polymorphisms

Received 16 November 2015; revised 4 January 2016; accepted 6 January 2016; accepted manuscript published online 29 January 2016; corrected proof published online 9 March 2016

RESULTS

We identified 10 loci containing SNPs with combined *P*-values meeting the genome-wide threshold of 5×10^{-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

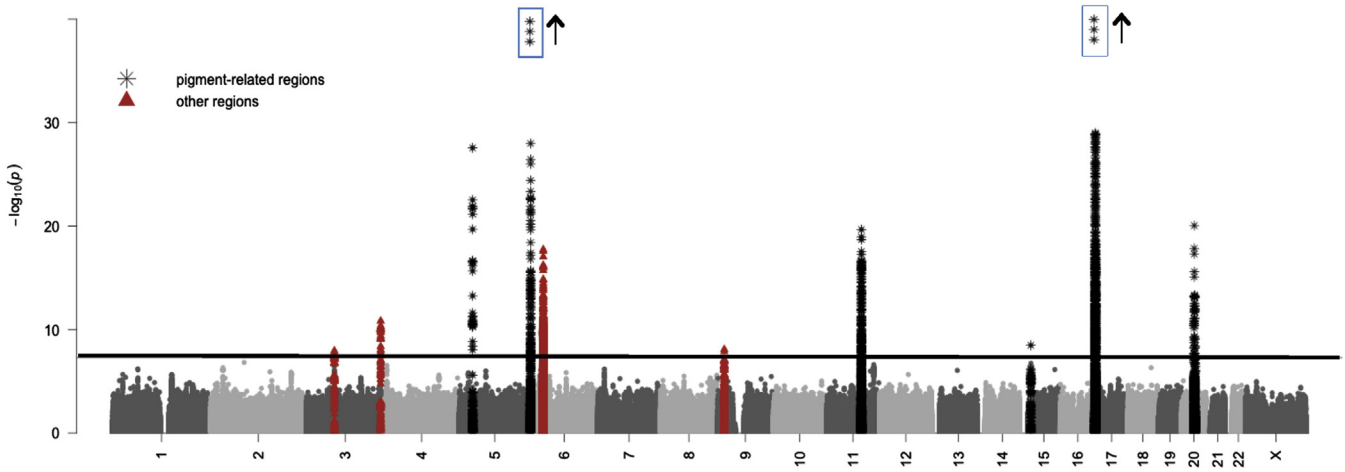

Figure 1. Manhattan plot showing $-\log_{10}$ *P*-values of squared Cochran-Armitage trend statistics. The horizontal line represents the threshold *P*-value of 5×10^{-8}. Markers within 50 kb of SNPs with *P*-values $< 5 \times 10^{-8}$ are indicated with black asterisks for those in pigment-related regions and in red triangles for those in other regions. The *y*-axis is truncated at $P = 10^{-30}$, although three SNPs at the 6p25 locus have *P*-values between 10^{-30} and 10^{-97}, and 72 SNPs at the 16q24 locus have *P*-values between 10^{-30} and 10^{-44}. SNPs, single-nucleotide polymorphisms.

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*, has been associated with SCC, BCC, and CMM (Barrett et al., 2011; Duffy et al., 2010a; 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 (Nan et al., 2011), 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[1]

						Initial screen			Replication phase			Combined		
Locus	SNP[2]	Gene	Minor allele	MAF[3]	Info[4]	OR[5]	CI[5]	*P*-value[6]	OR	CI	*P*-value	OR	CI	*P*-value
A. Pigment-related loci														
Chr 5p13	rs16891982	SLC45A2	C	0.45	Typed	0.53	0.47–0.60	1.64×10^{-24}	0.48	0.34–0.68	2.88×10^{-5}	0.52	0.47–0.59	2.77×10^{-28}
Chr 6p25	rs12203592	IRF4	T	0.17	1.00	1.54	1.48–1.61	2.45×10^{-83}	1.71	1.49–1.95	5.95×10^{-15}	1.56	1.49–1.62	8.29×10^{-97}
Chr 11q14	rs1126809	TYR	A	0.28	1.00	1.19	1.16–1.25	1.20×10^{-20}	1.08	0.96–1.22	2.09×10^{-1}	1.19	1.15–1.24	2.18×10^{-20}
Chr 15q13	rs12916300	HERC2	C	0.26	1.00	0.89	0.85–0.93	1.84×10^{-7}	0.82	0.72–0.93	2.40×10^{-3}	0.88	0.85–0.92	3.30×10^{-9}
Chr 16q24	rs4268748	DEF8	C	0.26	0.85	1.34	1.28–1.40	3.24×10^{-41}	1.28	1.13–1.45	7.85×10^{-5}	1.33	1.28–1.39	1.75×10^{-44}
Chr 20q11	rs6059655	RALY	A	0.08	0.99	1.30	1.22–1.38	5.18×10^{-17}	1.49	1.25–1.78	1.14×10^{-5}	1.32	1.24–1.39	9.03×10^{-21}
B. Other loci														
Chr 3p13	rs62246017	FOXP1	A	0.33	0.94	1.12	1.08–1.16	1.70×10^{-8}	1.07	0.95–1.2	0.26×10^{-1}	1.11	1.07–1.16	1.16×10^{-8}
Chr 3q28	rs6791479	TPRG1/ TP63[7]	T	0.43	1.00	1.12	1.08–1.16	2.57×10^{-9}	1.21	1.09–1.35	5.03×10^{-4}	1.13	1.09–1.16	1.47×10^{-11}
Chr 6p21	rs4455710	HLA-DQA1	T	0.38	1.00	1.17	1.12–1.21	1.80×10^{-16}	1.18	1.06–1.32	2.28×10^{-3}	1.17	1.13–1.21	1.86×10^{-18}
Chr 9p22	rs74664507	BNC2/ CNTLN[7]	T	0.44	0.95	0.90	0.87–0.93	3.64×10^{-8}	0.91	0.81–1.01	8.69×10^{-2}	0.90	0.87–0.93	8.24×10^{-9}

Abbreviations: SCC, squamous cell carcinoma; SNP, single-nucleotide polymorphisms.
[1]Regions containing SNPs with combined *P*-values $< 5 \times 10^{-8}$.
[2]SNP with strongest combined *P*-value in the region.
[3]MAF = minor allele frequency among control subjects.
[4]For imputed data, the info is the IMPUTE-2 information measure for imputation accuracy (Marchini and Howie, 2010).
[5]OR = odds ratio per minor allele, CI = 95% confidence interval.
[6]Based on the squared Cochran-Armitage trend test.
[7]Two flanking genes of an intergenic SNP.