XPA, haplotypes, and risk of basal and squamous cell carcinoma

Katie L.Miller¹², Margaret R.Karagas⁵, Peter Kraft¹⁶, David J.Hunter¹⁶, Paul J.Catalano³⁷, Steven H.Byler⁴ and Heather H.Nelson²⁺

¹Department of Epidemiology, ²Department of Environmental Health, ³Department of Biostatistics and ⁴Department of Genetics and Complex Diseases, Harvard School of Public Health, Boston, MA, USA, ⁵Section of Biostatistics and Epidemiology, Department of Community and Family Medicine and Norris Cotton Cancer Center, Dartmouth Medical School, Lebanon, NH, USA, ⁶Channing Laboratory, Department of Medicine, Brigham and Women’s Hospital and Harvard Medical School, Boston, MA, USA and ⁷Department of Biostatistics and Computational Biology, Dana Farber Cancer Institute, Boston, MA, USA

To whom correspondence should be addressed. Tel: +1 617 432 0107; Fax: +1 617 432 0107; Email: hnelson@hsph.harvard.edu

Nucleotide excision repair (NER) is instrumental in removing DNA lesions caused by ultraviolet (UV) radiation, the dominant risk factor for keratinocyte carcinoma, including basal cell carcinoma (BCC) and squamous cell carcinoma (SCC). We evaluated whether BCC or SCC risk was influenced by the A23G single nucleotide polymorphism (SNP) in Xeroderma pigmentosum group A (XPA), which codes for an essential protein in NER. We also investigated whether haplotypes of XPA, determined by seven haplotype-tagging SNPs, better define susceptibility to keratinocyte carcinoma. Incident cases of BCC and SCC from New Hampshire were identified through dermatologists and pathology laboratories. Population-based controls were frequency-matched to cases by gender and age. Cases of BCC (886) and of SCC (682) were compared as the 796. Models controlled for age, gender, pigmentation factors and severe sunburns and were restricted to Caucasians. Using GG as the reference, the A allele was less frequent among cases of BCC (ORAG = 0.82, 95% CI (0.66, 1.01); ORAA = 0.74, 95% CI (0.53, 1.03); trend test P = 0.03) and SCC (ORAG = 0.85, 95% CI (0.67, 1.07)); ORAA = 0.74, 95% CI (0.52, 1.05); trend test P = 0.05) than controls. Risk from ≥3 severe sunburns was elevated for those with the GG genotype only, and this interaction was nearly significant for BCC (P = 0.07). XPA genotype also modified a relationship between SCC and the amount of pigmentation (P = 0.02). Using a haplotype analysis identifying seven common XPA haplotypes, indicated that the A23G polymorphism alone captured the differences in susceptibility to keratinocyte carcinoma. The common G allele of the A23G polymorphism was associated with an increased risk of BCC and SCC and this polymorphism appeared to be the determining polymorphism in XPA that alters cancer susceptibility.

Introduction

Ultraviolet (UV) radiation from sunlight is the dominant risk factor for cancers of keratinocytes (KCs), including basal cell carcinoma (BCC) and squamous cell carcinoma (SCC) of the skin. UV induces DNA lesions, such as pyrimidine dimers and 6,4-photoproducts, which may lead to cancer if not repaired. The nucleotide excision repair (NER) pathway is necessary to remove these DNA lesions. A rare autosomal recessive condition, Xeroderma pigmentosum (XP), demonstrates the importance of this pathway. XP occurs when a gene (e.g., predominantly from Xeroderma pigmentosum group A through G) involved in NER contains a mutation on both copies of the gene that, when translated, results in a protein that is not capable of repairing photolesions. This leads to extreme photosensitivity and an estimated 1000-fold increased risk of KC with a much earlier age at onset (1,2). Null mutations in NER genes that result in XP are rare; however, these same genes are known to be highly polymorphic (3–6). Little is known about how these more common polymorphisms affect the risk of KC on a population level.

Xeroderma pigmentosum group A (XPA) is a gene that is necessary for NER. Null mutations in this gene lead to the most severe form of XP (7). In NER, XPA has a central role in interacting with a number of proteins, including RPA, TFIIH, and the ERCC1-XPF protein complex (8,9). A common polymorphism in XPA has been reported by several groups (10–12). The A23G polymorphism, also referred to as the XPA (−4) G-to-A polymorphism, is located in the 5′-untranslated region (UTR) and is four nucleotides upstream of the start codon. Polymorphisms in this area proximal to the start codon, referred to as the Kozak sequence, could have implications for the binding of the 40S ribosomal subunit and as a result influence protein levels in the cell (13,14).

One or more copies of the G allele resulted in significantly higher DNA repair capacity as measured by the host cell reactivation assay (15). Also, a reduced repair phenotype has been found to increase susceptibility to KC as well as other cancers (15–17). Epidemiologic studies have observed an increased risk of lung cancer with the A allele (15,18–20); however, XPA polymorphisms have not been studied in relation to KC risk.

Given the essential role of XPA in repairing UV lesions, we examined whether the A23G polymorphism is related to risk of BCC and SCC. Also, we investigated gene-environment interaction between UV exposure and this common polymorphism. Further, we conducted a haplotype analysis in order to determine whether additional polymorphisms are needed to identify those who are susceptible to KC.

Materials and methods

Study population

Newly diagnosed cases of histologically confirmed BCC and SCC in New Hampshire were identified through the collaboration of dermatologists,
dermatopathologists, and pathology laboratories throughout the state and bordering regions from July 1, 1993 to June 30, 1995 (series 1) and July 1, 1997 to March 30, 2000 (series 2) (21). Eligible cases were between 25 and 74 years of age, had a listed telephone number, and spoke English. Only living cases were enrolled; mortality and migration were not followed in this study. The number of controls selected was equal to the number of eligible cases, and controls were age-, sex-, and anatomic-site-matched to the cases. All eligible SCC cases and a ratio of approximately two to one BCC cases in series 1 and one to one ratio in series 2 were selected to take part in the study. The BCC cases were randomly sampled in order to ensure representativeness of age, sex and anatomic site for all incident BCCs within New Hampshire. A complete description of ascertainment of BCC and SCC cases has been described previously (21).

Population lists of New Hampshire residents obtained from the New Hampshire State Department of Transportation files were used to identify potential controls’ ages 25–64 years. Enrollment lists from the Center for Medicaid and Medicare Services provided a source of controls’ ages 65–74 years. Controls were frequency-matched to the combined case groups on age, sex, and anatomic site.

A personal interview (usually conducted in the participant’s home) covered demographic factors, pigmentation characteristics, sun exposure and sensitivity and other factors (22). All study protocol and materials were approved by the Dartmouth College Committee for the Protection of Human Subjects and all participants provided informed consent.

SNP selection
Information on coding variation across the length of XPA came from the Environmental Genome Project of the National Institute of Environmental Health Sciences (NIEHS) (http://egp.gs.washington.edu). This comprises 90 individuals of various ethnic backgrounds, including samples from subjects of European, African, Mexican, Native American, and Asian descent, although ethnicity information was definitively blinded (23). Sequencing of the XPA gene for 90 subjects identified 140 single nucleotide polymorphisms (SNPs), 72 of which had a minor allele frequency ≥0.05, our cutoff for inclusion in the haplotype analysis. Haploview determined that the coding variation for XPA was in one haplotype block (24). We used the Tagsnp program to estimate common haplotypes for XPA and determined that nine haplotype-tagging SNPs (htSNPs) were needed to reach an R² of 0.80 (25). One of these htSNPs was the A23G polymorphism.

Genotyping
Genotyping for eight htSNPs (rs1800975, rs3176633, rs3176689, rs2805835, rs3176719, rs1962592, rs3176751, and rs3176690) was conducted with ABI PRISM technology (Applied Biosystems, Foster City, California). Taqman primers, probes and conditions are available upon request. It was necessary to use PCR-RFLP for one htSNP, an insertion/deletion (rs3176649) preceeded by a SNP (rs3176648). The PCR reaction used the following primers: 5′-AGCCGAGTACACCTGCTT 3′ and 5′-AGAGTTGATAGCCAACCTGGTG 3′. The PCR product was digested with HhaI, which cleaves the sequence 5′-GCTG 3′ and produces 303 bp fragments. The reaction was incubated at 95°C for 2 min, then 35 PCR cycles (94°C for 30 s, 59°C for 30 s and 72°C for 45 s) and was completed with a hold at 72°C for 7 min. The Cac8I endonuclease cut the wild-type, regardless of the coding of the preceding SNP (rs3176648). The PCR reaction used the following primers: 5′-AGCTGAGTACACCTGCTT 3′ and 5′-AGAGTGATAGCCAACCTGGTG 3′. The PCR product was digested with HhaI, which cleaves the sequence 5′-GCTG 3′ and produces 303 bp fragments. The reaction was incubated at 95°C for 2 min, then 35 PCR cycles (94°C for 30 s, 59°C for 30 s and 72°C for 45 s) and was completed with a hold at 72°C for 7 min. The Cac8I endonuclease cut the wild-type, regardless of the coding of the preceding SNP (rs3176648). One of these htSNPs was the A23G polymorphism.

Results
The majority of the cases were men (Table I). The mean age of subjects in the study was 61.2 (standard error (SE) = 10.6) for controls, 58.8 (SE = 11.2) for BCC and 64.0 (SE = 8.8) for SCC. After adjusting for the matching factors age and gender, the number of severe sunburns (≥3 compared with the referent 0–2 severe sunburns) experienced in a lifetime was significantly higher for BCC (P < 0.0001) and SCC (P < 0.0001).

Table I. Selected characteristics of basal cell carcinoma and squamous cell carcinoma cases and controls

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Controls N = 796</th>
<th>BCC N = 886</th>
<th>SCC N = 682</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>485 (60.9%)</td>
<td>497 (56.1%)</td>
<td>434 (63.6%)</td>
</tr>
<tr>
<td>Female</td>
<td>311 (39.1%)</td>
<td>389 (43.9%)</td>
<td>248 (36.4%)</td>
</tr>
<tr>
<td>Age (SE)</td>
<td>61.2 (10.6)</td>
<td>58.8 (11.2)</td>
<td>64.0 (8.8)</td>
</tr>
<tr>
<td>Severe sunburns*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0–2</td>
<td>482 (61.6%)</td>
<td>405 (46.2%)</td>
<td>316 (47.2%)</td>
</tr>
<tr>
<td>≥3</td>
<td>301 (38.4%)</td>
<td>471 (53.8%)</td>
<td>354 (52.8%)</td>
</tr>
<tr>
<td>Pigment score*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quartile 1–dark</td>
<td>197.5 (24.8%)</td>
<td>85 (9.6%)</td>
<td>65 (9.5%)</td>
</tr>
<tr>
<td>Quartile 2</td>
<td>198.5 (24.9%)</td>
<td>173 (19.5%)</td>
<td>107 (15.7%)</td>
</tr>
<tr>
<td>Quartile 3</td>
<td>200.5 (25.2%)</td>
<td>242 (27.3%)</td>
<td>196 (28.7%)</td>
</tr>
<tr>
<td>Quartile 4–fair</td>
<td>199.5 (25.1%)</td>
<td>386 (43.6%)</td>
<td>314 (46.0%)</td>
</tr>
</tbody>
</table>

*Missing data on severe sunburns for <2% of subjects (13 Controls, 10 BCC, 12 SCC).

Pigment score was generated using a multivariate confounder score and combined data on the following pigment factors: skin color to first hour of intense sunshine, skin reaction to repeated sun exposure, hair color, eye color, skin color, and number of moles on the back; higher pigment score represented lower pigmentation and melanin production. The distribution of pigment score in controls represents the average for the pigment score in controls as generated separately for BCC and SCC.
compared with controls, and pigment score was also a significant predictor of BCC ($P < 0.0001$) and SCC ($P < 0.0001$).

For the A23G polymorphism, the variant A allele was more frequent in controls (34%) than either BCC (31%) or SCC (31%) (Table II). The distribution of the XPA genotypes in controls was in Hardy–Weinberg equilibrium. Genotypes were genotyped, and overall genotyping for the seven htSNPs was 92% complete (rs1800975, rs3176633, rs3176649, rs3176689, rs2805835, rs3176719, and rs1962592; and within SNP, wild-type = 0 and variant = 1). Nine htSNPs were originally identified from the multi-ethnic NIEHS data, and two of these SNPs were not observed in the Caucasian population. The seven remaining htSNPs were genotyped, and overall genotyping for the seven htSNPs was 92% complete (rs1800975, rs3176633, rs3176649, rs3176689, rs2805835, rs3176719, and rs1962592). All seven htSNPs were in Hardy–Weinberg equilibrium. The results from the haplotype analysis are presented in Table III. The A23G polymorphism corresponds to the first SNP (A23G) in the haplotype. Seven common haplotypes were identified in this study, and the eighth group represents the pooling of ‘rare’ haplotypes (individual haplotypes with a frequency <0.05). The global test on the haplotypes was statistically significant for BCC ($P = 0.03$), but not for SCC ($P = 0.41$). Haplotypes containing the A23G polymorphism were associated with a lower risk of KC compared with the common, referent haplotype containing all wild-type polymorphisms, confirming results of the A23G polymorphism alone. We tested for a difference between haplotypes 2 and 4 and found there was no statistical difference in the haplotypes containing the A23G polymorphism. Therefore, for the interaction models, we only used the A23G polymorphism. Further, because the association with KC was approximately the same for the AA and AG genotypes, we collapsed on the A allele.

In the main effect model for severe sunburns ($\geq 3$ compared with $0–2$), there was a 50% significantly increased risk for BCC (OR, 1.50, 95% CI, 1.22–1.85) and a 60% significantly increased risk of SCC (OR, 1.59, 95% CI, 1.27–2.00) (Table IV). Within strata of severe sunburns, the association between XPA genotype and BCC and SCC was statistically significant for those who reported experiencing three or more severe sunburns (BCC: OR, 0.65, 95% CI, 0.48–0.88; SCC: OR, 0.69, 95% CI, 0.50–0.96) (Table IV). When we modeled the joint effects of XPA genotype and severe sunburns, the increased risk of BCC and SCC occurring among those with $\geq 3$ sunburns was stronger for the GG genotype than those with one or two copies of the A allele. A test for statistical interaction was not significant for SCC ($P = 0.17$) but approached significance for BCC ($P = 0.07$).

Because UV exposure at the target cell can vary by pigmentation, we next examined the influence of pigmentation on the risk of KC (Table V). There was an increasing risk of BCC with increasing pigment score (higher pigment score corresponds to lower pigmentation and greater UV exposure), which was as strong as a 6-fold increased risk for the fourth quartile compared with the first (Quartile 2: OR, 1.99, 95% CI, 1.43–2.77; Quartile 3: OR, 2.52, 95% CI, 1.82–3.47; Quartile 4: OR,
When we examined the joint effects of XPA and pigment score, there was an increasing risk of BCC with increasing pigment score for those with the GG genotype as well as those with the AG or AA genotypes (Table V). Within each pigment score quartile, the risk of BCC for those with genotypes containing the A allele was lower than those with the GG genotype. For SCC, the risk among those with the GG genotype increased with increasing pigment score. For genotypes containing an A allele, the second through fourth quartiles with the AG or AA genotypes had a pattern of increasing risk of SCC with increasing pigment score, and these ORs were lower than those of the homozygous wild-type for the same pigment score quartile. However, in the first pigment score quartile, the risk of BCC for those with genotypes containing an A allele was lower than those with the GG genotype. The test for interaction also had 3 degrees of freedom.

When we examined the joint effects of XPA and pigment score, there was an increasing risk of BCC with increasing pigment score for those with the GG genotype as well as those with the AG or AA genotypes (Table V). Within each pigment score quartile, the risk of BCC for those with genotypes containing the A allele was lower than those with the GG genotype. For SCC, the risk among those with the GG genotype increased with increasing pigment score. For genotypes containing an A allele, the second through fourth quartiles with the AG or AA genotypes had a pattern of increasing risk of SCC with increasing pigment score, and these ORs were lower than those of the homozygous wild-type for the same pigment score quartile. However, in the first pigment score quartile, the risk of SCC was elevated among those with genotypes containing the A allele and of borderline statistical significance (OR, 1.72, 95% CI, 0.97–3.06).
for interaction between pigmentation and XPA genotype was statistically significant for SCC only (test for interaction, 3 degrees of freedom: SCC $P = 0.02$; BCC $P = 0.62$).

**Discussion**

We found the common G allele of the A23G XPA polymorphism associated with an increased risk of BCC and SCC. Accounting for coding variation across the full length of XPA in the form of a haplotype analysis did not reveal any additional information about susceptibility to KC. When UV exposure was taken into account, the increased risk of the GG genotype appeared to be influenced by severe sunburns. When we used pigmentation factors as a measure of susceptibility to UV exposure at the target cell, we observed a statistically significant pigment-XPA interaction for SCC.

The distribution of alleles in controls in this study (A allele frequency: 34%) was similar to that of other Caucasian populations. This includes a population in Germany (33%) and a population in Kentucky (26%) (10,18,20). Compared with other ethnicities, the observed frequency was similar to a sample of African Americans (30%) and Mexican Americans (39%), although it differed from that found in a Korean population (48%), suggesting that allele frequency could vary by ethnicity (9,15,19).

Previous studies have found that the AA genotype leads to an increased risk of lung cancer (15,18–20). In contrast, in our study of skin cancer we observed that the AA genotype occurred less frequently among BCC and SCC cases compared with controls. This contrast is consistent with prior investigations of DNA repair polymorphisms where the variant alleles are risk factors for certain cancers but are associated with a relative risk below the null for KC studies. For example, this has been reported for XRCC1 and XPC (22,30–34). An explanation for this is based on the greater apoptotic response of keratinocytes to DNA damage (35). As described previously (22), keratinocytes that carry the allele for less efficient repair may have greater DNA damage, leading to apoptosis prior to cell division, and presenting in epidemiologic studies as relative risks below the null for KCs. Another possible explanation for the different findings from previous studies of the XPA polymorphism may relate to differences in exposure. Cellular response to damage from smoking compared with UV could elicit different responses from NER and other repair and damage response pathways.

Our results from the analysis of the joint effects of XPA and UV-related factors support this hypothesis. The degree of pigmentation determines the amount of UV reaching the nucleus of keratinocytes. In our analysis, a lower pigment score was synonymous with more pigment, therefore allowing less UV radiation to reach a cell’s DNA. An increased risk with the AG and AA genotypes compared with the GG genotype was confined to the lowest pigment quartile for SCC. These results suggest that for those with more pigment, limited UV radiation reaches the nucleus to trigger the cells with lower repair capacity (A allele) to undergo apoptosis, resulting in the increased risk of SCC. In contrast, as pigmentation decreased, i.e., in the higher quartiles of the score variable, we see that the GG genotype again conferred greater susceptibility to SCC and BCC. When we examined the combined effects of the number of severe sunburns and XPA genotype, we observed similar trends. Those with the GG genotype and a greater number of severe sunburns were at increased risk of disease. Based on the finding that the G allele resulted in better DNA repair capacity (15), it appears possible that keratinocytes with the GG genotype may have sufficient repair following UV exposure to allow for mitosis, whereas those with an A allele are more likely to undergo apoptosis.

A common criticism of epidemiologic studies of SNPs is that the results could be explained by the SNP being in linkage disequilibrium with other SNPs on the gene. Therefore, in this analysis we examined haplotypes of XPA to account for coding variation across the length of the gene. Our htSNP selection process focused on identifying common haplotypes that influence the risk of KC. However, the relationship between common variation across XPA and its influence on KC has not been studied previously. Therefore, we wanted to examine the impact of these common haplotypes on the risk of BCC and SCC. However, the analysis supported the findings from the A23G polymorphism and did not suggest that common haplotypes provided additional information about susceptibility to BCC or SCC. Thus, we have greater confidence in the findings related to this polymorphism, and feel that the association is unlikely to be explained by linkage disequilibrium within the gene. Of course, this analysis focused on common haplotypes, which selected htSNPs to explain 80% of the variation in haplotypes. Rare haplotypes may not be strongly correlated with the seven observed SNPs. These data cannot distinguish between an association due to the G allele at A23G and an association due to multiple rare haplotypes in strong LD with the G allele. However, we consider it unlikely that multiple rare disease mutations would arise and remain in strong association with the G allele if that allele played no causal role.

Genotyping for this gene was 92% complete for all seven htSNPs and over 97% for the A23G polymorphism. Missing genotype was not related to case status, therefore, it is unlikely that this unobserved information would have changed our results. Over 80% of subjects enrolled in the study provided a DNA sample. There were no differences in demographic, sunburn or pigment characteristics between those who provided a sample and those who did not. We do not think selection bias is a major concern for this study since the participation from incident cases of KC in New Hampshire were over 80%, and identification of controls through Medicare files and drivers’ license records are believed to cover over 90% of the New Hampshire population (36). In order for the non-participants to bias this analysis, they would have to differ from participants in terms of XPA genotype, which is unlikely.

Our study suggests that the common GG genotype of the XPA A23G polymorphism increases the risk of BCC and SCC. Further, we found evidence that XPA genotype may modify the association between KC and UV exposure. More research into the apoptotic threshold in keratinocytes is needed to understand the mechanism through which this association occurs.

**Acknowledgement**

Support: NIH R01CA082354, R01CA57494, and CA006515, and NIEHS T32 ES007155.

**Conflict of Interest Statement:** None declared.

**References**


Received December 19, 2005; revised February 8, 2006; accepted February 14, 2006