A population-based study of DNA repair gene variants in relation to non-melanoma skin cancer as a marker of a cancer-prone phenotype

Ingo Ruczinski 1,4, Timothy J. Jorgensen 1,4, Yin Yao Shugart 1, Yvette Berthier Schaad 2,5, Bailey Kessing 5, Judith Hoffman-Bolton 6,4, Kathy J. Helzlsouer 7, W.H. Linda Kao 8, Lee Whelless 9, Lesley Francis 10, Rhoda M. Alan 10, Paul T. Strickland 10, Michael W. Smith 11 and Anthony J. Alberg 12

1Department of Biostatistics, The Johns Hopkins University Bloomberg School of Public Health, Baltimore, MD, USA
2Department of Radiation Medicine, Georgetown University School of Medicine, Washington, DC, USA
3Division of Intramural Research Program, National Institute of Mental Health, Bethesda, MD, USA
4Department of Epidemiology, The Johns Hopkins University Bloomberg School of Public Health, Baltimore, MD, USA
5Laboratory of Genomic Diversity, SAIC-Frederick, NCI-Frederick, Frederick, MD
6George W. Comstock Center for Public Health Research and Prevention, Washington County, MD, USA
7Mercy Medical Center, Baltimore, MD, USA
8Hollings Cancer Center and Division of Epidemiology and Biostatistics, Department of Medicine, Medical University of South Carolina, Charleston, SC, USA
9Department of Dermatology, Boston University School of Medicine, Boston, MA, USA
10Department of Environmental Health Sciences, The Johns Hopkins University Bloomberg School of Public Health, Baltimore, MD, USA
11Genetics and Genomics Group, Advanced Technology Program, SAIC-Frederick, Inc., NCI-Frederick, Frederick, MD, USA
4To whom correspondence should be addressed. Tel: +843 876 2439; Fax: +843 876 2344; Email: alberg@musc.edu

For unknown reasons, non-melanoma skin cancer (NMSC) is associated with increased risk of other malignancies. Focusing solely on DNA repair or DNA repair-related genes, this study tested the hypothesis that DNA repair gene variants contribute to the increased cancer risk associated with a personal history of NMSC. From the parent CLUE II cohort study, established in 1989 in Washington County, MD, the study consisted of a cancer-free control group (n = 2296) compared with three mutually exclusive groups of cancer cases ascertained through 2007: (i) Other (non-NMSC) cancer only (n = 2349); (ii) NMSC only (n = 694) and (iii) NMSC plus other cancer (n = 577). The frequency of minor alleles in 759 DNA repair gene single nucleotide polymorphisms (SNPs) was compared in these four groups. Comparing those with both NMSC and other cancer versus those with no cancer, 10 SNPs had allelic trend P-values < 0.01. The two top-ranked SNPs were both within the thymine DNA glycosylase gene (TDG). One was a non-synonymous coding SNP (rs2888808) [per allele odds ratio (OR) 1.40, 95% confidence interval (CI) 1.16–1.70; P-value = 0.0006] and the other was an intronic SNP in high linkage disequilibrium with rs2888808 (rs4135150). None of the associations had a P-value < 6.6 × 10−5, the threshold for statistical significance after correcting for multiple comparisons. The results pinpoint DNA repair genes most likely to contribute to the NMSC cancer-prone phenotype. A promising lead is genetic variants in TDG, thymine DNA glycosylase gene; XP, xeroderma pigmentosum.

Introduction

Non-melanoma skin cancer (NMSC) is by far the most common human cancer (1). NMSC is usually treated by local excision and is rarely fatal, but a personal history of NMSC is associated with increased risk of other malignancies (2). In a meta-analysis of prospective studies with individual-level data, a prior NMSC diagnosis was associated with a 50% greater risk of developing another type of cancer (2). NMSC was associated with a broad spectrum of malignancies (2), suggesting that NMSC may be a marker of a cancer-prone phenotype. NMSC cases comprise two major histologic types, basal cell carcinoma (BCC) and squamous cell carcinoma (SCC), and the overall increased risk of other malignancies associated with NMSC has been consistently observed for both BCC and SCC. However, it is not known why a personal history of NMSC is associated with increased risk of non-cutaneous malignancies, and there has been no previous hypothesis-driven research on this topic. The fact that a personal history of NMSC affects overall cancer risk across a broad spectrum of malignancies potentially makes NMSC a cross-cutting marker of overall cancer risk that could potentially provide valuable clues to human carcinogenesis.

Major DNA repair defects cause cancer-prone phenotypes. For example, mutated DNA repair genes cause cancer-prone syndromes that involve multiple cancers, such as xeroderma pigmentosum (XP) (3,4) and Bloom’s syndrome (5,6). XP is an autosomal recessive condition caused by rare, high-penetrant mutations in nucleotide excision repair (NER) genes. XP directly demonstrates that defective DNA repair can be a potent cause of genetic predisposition to both BCC and other malignancies. The DNA repair deficiency in XP is characterized by extreme sensitivity to ultraviolet radiation, causing >2000-fold increased skin cancer rates before age 20 years (4). XP patients also have a marked excess of internal malignancies (4). The fact that rare mutations in DNA repair genes cause major cancer susceptibility in XP patients suggests that even common, low-penetrant DNA repair gene variants may jointly affect susceptibility to NMSC and other malignancies in the general population (7,8). However, except for rare, autosomal recessive cancer syndromes caused by inherited mutations in DNA repair genes, the contribution of common allelic variants of DNA repair genes to cancer risk in the general population remains incompletely characterized (e.g. ref. 7).

DNA repair is accomplished through multiple interacting biochemical pathways entailing overlapping functions and cross-talk; collectively, these comprise >140 genes (Table I, column 1). All of these pathways have been implicated in carcinogenesis. The NER pathway is emphasized for NMSC because NER removes cyclobutane pyrimidine dimers and 6-4 photoproducts, carcinogenic DNA lesions unique to ultraviolet radiation exposure. Ultraviolet radiation-induced DNA lesions are not restricted to pyrimidine dimers but also entail oxidative lesions (8,9), demonstrating that even protection against skin cancer involves multiple DNA repair pathways. Thus, in the present study, all known human DNA repair genes were considered candidates for association with the NMSC-associated cancer-prone phenotype.

Consistent with other cohort studies, we previously reported that a prior NMSC diagnosis was associated with significantly elevated risk of other malignancies (10). Building on this finding, the present study focused solely on DNA repair and DNA repair-related genes and employed the same cohort to test the novel hypothesis that germ-line polymorphisms in DNA repair genes contribute to the cancer-prone phenotype characterized by the diagnosis of both NMSC and another type of cancer in the same individual.

Abbreviations: BCC, basal cell carcinoma; LD, linkage disequilibrium; MAF, minor allele frequency; NER, nucleotide excision repair; NMSC, non-melanoma skin cancer; SNP, single nucleotide polymorphism; SCC, squamous cell carcinoma; TDG, thymine DNA glycosylase gene; XP, xeroderma pigmentosum.

1These authors contributed equally to this work
Materials and Methods

This prospective, population-based study was embedded within the CLUE II (named after the campaign slogan of ‘Give us a clue to cancer and heart disease’) cohort. CLUE II is a community-based cohort study established in Washington County, MD in 1989 (11). In this study, the independent variables were single nucleotide polymorphisms (SNPs) in DNA repair genes. The dependent variable was the occurrence of cancer classified according to the joint outcome of NMSC and other cancer, with a cancer-free group to compare with those with cancer other than NMSC only, NMSC only, and both NMSC plus other cancer. The study of all non-NMSC cancers combined is consistent with the prior observation that the excess risk of other malignancies observed in those with a personal history of NMSC is not confined to a few specific malignancies but rather seems to apply to an overall heightened cancer risk.

Study population

The CLUE II cohort was established when baseline data were collected in May through November 1989 from volunteers who were mostly residents of Washington County, MD. The campaign was designed to collect blood samples from as many adult residents as possible in the Washington County and surrounding (30 mile radius) area. Brief medical histories and blood pressures were taken, and 20 ml of blood was drawn into heparinized vacutainers. Specimens were refrigerated at once and processed within 24 h. Buffy coats were placed in storage at −70 °C. During the baseline data collection, participants completed a questionnaire that included information on age, race, sex, cigarette smoking, height, weight and years of schooling. This study was approved by the Institutional Review Boards of the Johns Hopkins University and Medical University of South Carolina.

From the entire cohort of 30726 participants, 6589 were selected for inclusion in the present study. This included all those with a confirmed cancer diagnosis as of 30 September 2007 (plus 96 cases added 1 April 2008) and a cancer-free comparison group that was a 10% age-stratified random sample of adult CLUE II participants, plus 250 controls added for a lung cancer substudy. Once selected, NMSC or other cancer diagnoses that occurred by 31 December 2010 contributed to the categorization of study endpoints.

Of the 6589 subjects selected for inclusion, genotyping could not be attempted for 312, and 263 with ≥5% genotyping failures were excluded. Among the remaining 6014 subjects, genetic heterogeneity was assessed using principal components. Comparison with the three HapMap populations identified 98 (1.6%) subjects of non-European ancestry, who were excluded due to concerns about population stratification and differential risks of NMSC. To account for any residual ancestral differences, the first three principal components were adjusted for in all analyses. The final study population of 5916 was classified into four categories: (i) cancer-free group (n = 2296) for comparison to those with a pathologically confirmed diagnosis of any cancer other than NMSC (‘other cancer only’, n = 2349); (ii) NMSC with no other cancer diagnosis (‘NMSC only’, n = 694) and (iv) NMSC plus another type of cancer (‘NMSC plus other cancer’, n = 577).

Some genes may be associated with more than one pathway; if so, the gene was arbitrarily assigned to the most recognized pathway.
typing platforms for those who overlapped with the present study population (per SNP number of participants with duplicate genotype results ranged from 3943 to 4757); the overall concordance was 98%.

Statistical analyses

For each SNP, an exact test was used to test whether genotypic frequencies in the cancer-free controls departed from HWE (15). Logistic regression was used to estimate odds ratios, confidence intervals and P-values for the associations between each SNP and cancer risk. Odds ratios were estimated using the cancer-free group as the referent category. The hypothesis tested was that minor alleles in DNA repair genes would be most prevalent in the NMSC plus other cancer group and least prevalent in the cancer-free comparison group. Given this hypothesis, the SNP screening strategy focused solely on the comparison of the NMSC plus other cancer group to the cancer-free comparison group. The additive genetic model (SNPs coded as having 0, 1 or 2 copies of the minor allele) was used to screen the associations with P-values < 0.01 in the NMSC plus other cancer category. For the SNPs screened with the additive genetic model P-value < 0.01 for the NMSC plus other cancer group, the associations between these SNPs and the risk of other cancer only and NMSC only were also evaluated. The associations between these SNPs and the NMSC cancer-prone phenotype were then further characterized in greater detail by assessing the dominant and recessive phenotypic models of inheritance.

Analyses were carried out in the statistical environment R (http://cran.r-project.org/). All statistical tests were two-sided. Assuming a P-value of 0.05 represents a statistically significant association for a single comparison, a Bonferroni multiple comparison correction for the 759 DNA repair gene SNPs was used to adjust for the number of comparisons. Given the number of SNPs tested, a P-value of 0.05 was corrected to 0.00006 (6.6 × 10−5). The top ranked SNP (rs2888805) was non-synonymous in the coding region of the thymine DNA glycosylase gene (TDG) (per minor allele odds ratio for the ‘both’ category (ORBoth) 1.40, 95% confidence interval (CI) 1.16–1.70). As expected, this SNP was also associated with the cancer-prone phenotype; each of these SNP pairs comprised one non-synonymous coding SNP and one intronic SNP that were in high LD and had identical associations with the NMSC cancer-prone phenotype. The results for all 759 SNPs are summarized in Supplementary Table I, available at Carcinogenesis Online.

A measure of the internal consistency of this finding was that the second ranked SNP was an intronic TDG SNP (rs4135150) in high LD (r² = 0.99) with rs2888805 and with virtually identical results, the prevalence of ever-smokers was highest (56%) in the other cancer group, and current smoking was less prevalent in study groups that included NMSC. Other than NMSC, the most common malignancies were breast, prostate, lung and colorectal cancer.

Table III lists results for the 10 SNPs with additive model P-values < 0.01 for the cancer-prone phenotype (‘both NMSC plus other cancer’) versus the ‘no cancer’ comparison. The smallest P-value was 6 × 10−4, greater than the Bonferroni threshold of 6.6 × 10−5; thus, none of the observed associations were statistically significant after correction for multiple comparisons. The 10 SNPs represented the following 6 pathways and 8 genes: 2 nucleotide excision repair (ERCC8, ERCC3), 2 homologous recombination repair (PALB2, DMC1) and 1 each from base excision repair (TDG), direct reversal repair (MGM1), DNA damage signal transduction (CHEK2) and mismatch repair (MSH6). Both TDG and PALB2 had two SNPs associated with the cancer-prone phenotype; each of these SNP pairs comprised one non-synonymous coding SNP and one intronic SNP that were in high LD and had identical associations with the NMSC cancer-prone phenotype.

Baseline descriptive characteristics of the four study groups are summarized in Table II. The cancer-free group was significantly younger (43 years) than the three groups that included cancer, whose average ages ranged from 57 to 62 years. The NMSC plus other cancer group was 56% male, notably higher than the other study groups.

Results

Baseline descriptive characteristics of the four study groups are summarized in Table II. The cancer-free group was significantly younger (43 years) than the three groups that included cancer, whose average ages ranged from 57 to 62 years. The NMSC plus other cancer group was 56% male, notably higher than the other study groups. Age group 45–54 y had the highest prevalence of ever-smokers (56%) in the other cancer only group, and current smoking was less prevalent in study groups that included NMSC. Other than NMSC, the most common malignancies were breast, prostate, lung and colorectal cancer.

Table III lists results for the 10 SNPs with additive model P-values < 0.01 for the cancer-prone phenotype (‘both NMSC plus other cancer’) versus the ‘no cancer’ comparison. The smallest P-value was 6 × 10−4, greater than the Bonferroni threshold of 6.6 × 10−5; thus, none of the observed associations were statistically significant after correction for multiple comparisons. The 10 SNPs represented the following 6 pathways and 8 genes: 2 nucleotide excision repair (ERCC8, ERCC3), 2 homologous recombination repair (PALB2, DMC1) and 1 each from base excision repair (TDG), direct reversal repair (MGM1), DNA damage signal transduction (CHEK2) and mismatch repair (MSH6). Both TDG and PALB2 had two SNPs associated with the cancer-prone phenotype; each of these SNP pairs comprised one non-synonymous coding SNP and one intronic SNP that were in high LD and had identical associations with the NMSC cancer-prone phenotype. The results for all 759 SNPs are summarized in Supplementary Table I, available at Carcinogenesis Online.

A measure of the internal consistency of this finding was that the second ranked SNP was an intronic TDG SNP (rs4135150) in high LD (r² = 0.99) with rs2888805 and with virtually identical results, the prevalence of ever-smokers was highest (56%) in the other cancer only group, and current smoking was less prevalent in study groups that included NMSC. Other than NMSC, the most common malignancies were breast, prostate, lung and colorectal cancer.

Table III lists results for the 10 SNPs with additive model P-values < 0.01 for the cancer-prone phenotype (‘both NMSC plus other cancer’) versus the ‘no cancer’ comparison. The smallest P-value was 6 × 10−4, greater than the Bonferroni threshold of 6.6 × 10−5; thus, none of the observed associations were statistically significant after correction for multiple comparisons. The 10 SNPs represented the following 6 pathways and 8 genes: 2 nucleotide excision repair (ERCC8, ERCC3), 2 homologous recombination repair (PALB2, DMC1) and 1 each from base excision repair (TDG), direct reversal repair (MGM1), DNA damage signal transduction (CHEK2) and mismatch repair (MSH6). Both TDG and PALB2 had two SNPs associated with the cancer-prone phenotype; each of these SNP pairs comprised one non-synonymous coding SNP and one intronic SNP that were in high LD and had identical associations with the NMSC cancer-prone phenotype. The results for all 759 SNPs are summarized in Supplementary Table I, available at Carcinogenesis Online.

A measure of the internal consistency of this finding was that the second ranked SNP was an intronic TDG SNP (rs4135150) in high LD (r² = 0.99) with rs2888805 and with virtually identical results, the prevalence of ever-smokers was highest (56%) in the other cancer only group, and current smoking was less prevalent in study groups that included NMSC. Other than NMSC, the most common malignancies were breast, prostate, lung and colorectal cancer.
indicating that both SNPs were equally informative about the strength of the allelic association because they are physically linked to each other. A total of six SNPs were successfully genotyped that tagged regions of TDG, rs2888805 and rs4150454 plus four more SNPs, rs4964435, rs6539116, rs4135113 and rs2583274. None of these other four were in high LD with one another (all r² < 0.10), and none were associated with the NMSC cancer-prone phenotype except rs4964435, which was nominally significantly inversely associated with cancer risk (per minor allele OR both 0.81; 95% CI 0.67–0.98, P-value 0.029).

This pattern of associations suggests that the TDG haplotype blocks tagged by these other four SNPs (rs4964435, rs6539116, rs4135113, rs2583274) are not likely to contribute to the functional changes in TDG that may be affecting the risk of NMSC plus another cancer. Rather, to the extent the observed associations with rs2888805 and rs4150454 are genuine, the results point toward a true functional variant residing within the haplotype block tagged by these two SNPs but unlikely to reside in the other haplotype blocks.

As with TDG, PALB2 also had two SNPs in high LD (r² = 0.99) and consequently nearly identical ORs. The two PALB2 SNPs were rs152451 (ORboth 1.35; 95% CI 1.10–1.66, non-synonymous coding, rank no. 4) and rs447529 (ORboth 1.35; 95% CI 1.10–1.65, intronic, rank no. 5). The remaining associated SNPs comprised one SNP each from ERCC8 (ORboth 1.35, rank no. 3), ERCC3 (ORboth 0.82, rank no. 7), DMC1 (ORboth 0.84, rank no. 8), CHEK2 (ORboth 1.25, rank no. 9) and MSH6 (ORboth 1.21, rank no. 10). To account for the remote possibility that these associations between genetic markers and cancer could potentially be confounded by demographic or lifestyle characteristics, additional adjustments for age, gender, education, body mass index, and cigarette smoking status at baseline were tested. The adjustments made only negligible differences, and these slight differences tended to be in the direction of slightly strengthening the observed associations for ORboth (Supplementary Table 2, available at Carcinogenesis Online). Due to the younger age distribution of the cancer-free comparison group, the associations were also assessed with age restricted to those ≥35 years at baseline, which excluded 799 (35%) from the cancer-free group but only 8 (1%) from the NMSC plus other cancer group. The associations observed in this age-restricted analysis were very similar to the overall results (Supplementary Table 3, available at Carcinogenesis Online).

Table IV summarizes associations for the same 10 SNPs according to genotype (two degrees of freedom) and dominant and recessive modes of inheritance. Nine SNPs (all except rs797690) were more consistent with the dominant than recessive model based on the deviances in the respective logistic regression models. For the top-ranked SNPs in TDG (rs2888805: OR both 1.53; rs4150454: OR both 1.54), ERCC8 (rs1038144: OR both 1.44) and PALB2 (rs152451: OR both 1.40; rs447529: OR both 1.39), the ORs for the dominant model ranged from 1.4 to 1.5 with P-values ranging from 0.0001 to 0.0036. For the TDG SNPs, in the model with two degrees of freedom, the significance of the parameter for the one versus zero variant allele comparison was near (rs2888805: 8.2 × 10⁻³) or below (rs4150454: 6.1 × 10⁻³) the Bonferroni significance threshold.

**Discussion**

The present study was carried out to test the hypothesis that germ-line variants in DNA repair genes were associated with the NMSC cancer-prone phenotype. SNPs were genotyped in known DNA repair genes in a population-based study to determine if germ-line polymorphisms contribute to joint susceptibility to both NMSC and other cancers. The results revealed several variant alleles that may contribute to a cancer-prone phenotype characterized by susceptibility to both NMSC and other malignancies.

An interesting observation was that carrying a non-synonymous coding SNP (rs2888805) in the base excision repair gene TDG was shown to be associated with a 50% increased risk of developing NMSC plus another cancer. This polymorphic SNP leads to a methionine residue at position 367, or rarely leucine, rather than the more common valine (16). This residue does not fall within the known active sites for the protein or targets of acetylation, and its relevance to overall protein function is currently unknown. In the absence of direct experimental data, bioinformatics databases were queried to assess for the putative functionality and, consistent with the information above, TDG SNP rs2888805 was predicted to be benign in the PolyPhen-2, Condel and SIFT bioinformatics databases. For the SNP in high LD with rs2888805, rs4150454, the most current alignment has this SNP residing in GLTD2, not TDG. rs4150454 is in intron 8 (of 8) in GLTD2, but was selected as a tagging SNP for the 3′ UTR of TDG. rs2888805 is in exon 10 (of 10) in TDG. The TDG and GLTD2 genes run in opposite directions, so the two SNPs are only approximately 3000 bp apart. The most current alignment also has three other TDG tagging SNPs (rs4964435, rs6539116, rs2583274) physically residing in GLTD2.

TDG was the first known mismatch-specific DNA repair gene; TDG repairs thymine mispaired with guanine, a mispairing that is both highly mutagenic and common due to the propensity of 5-methylcytosine to

---

**Table III. DNA repair SNPs (n = 10) with P-value < 0.01 for the allelic trend test comparing the ‘both NMSC plus other cancer’ to the ‘no cancer’ category, ordered by P-value**

<table>
<thead>
<tr>
<th>SNP</th>
<th>Gene</th>
<th>Pathway</th>
<th>Type</th>
<th>MAF</th>
<th>P(HWE)</th>
<th>OR 95% CI</th>
<th>OR 95% CI</th>
<th>OR 95% CI</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs2888805</td>
<td>TDG</td>
<td>BER</td>
<td>Nonsyn</td>
<td>0.10</td>
<td>0.25</td>
<td>0.98 (0.85–1.12)</td>
<td>1.07 (0.88–1.30)</td>
<td>1.40 (1.16–1.70)</td>
<td>0.0006</td>
</tr>
<tr>
<td>rs4150454</td>
<td>TDG</td>
<td>BER</td>
<td>Intron</td>
<td>0.10</td>
<td>0.38</td>
<td>0.98 (0.86–1.13)</td>
<td>1.08 (0.89–1.31)</td>
<td>1.40 (1.15–1.70)</td>
<td>0.0007</td>
</tr>
<tr>
<td>rs1038144</td>
<td>ERCC8</td>
<td>NER</td>
<td>Intron</td>
<td>0.09</td>
<td>0.23</td>
<td>1.07 (0.93–1.24)</td>
<td>0.96 (0.77–1.19)</td>
<td>1.38 (1.12–1.71)</td>
<td>0.0030</td>
</tr>
<tr>
<td>rs152451</td>
<td>PALB2</td>
<td>HR</td>
<td>Nonsyn</td>
<td>0.09</td>
<td>0.64</td>
<td>0.95 (0.83–1.10)</td>
<td>1.01 (0.82–1.25)</td>
<td>1.35 (1.10–1.66)</td>
<td>0.0043</td>
</tr>
<tr>
<td>rs447529</td>
<td>PALB2</td>
<td>HR</td>
<td>Intron</td>
<td>0.09</td>
<td>0.64</td>
<td>0.95 (0.82–1.10)</td>
<td>1.01 (0.82–1.24)</td>
<td>1.35 (1.10–1.65)</td>
<td>0.0046</td>
</tr>
<tr>
<td>rs11016857</td>
<td>MGMT</td>
<td>DR</td>
<td>Intron</td>
<td>0.43</td>
<td>0.43</td>
<td>0.96 (0.88–1.04)</td>
<td>1.04 (0.92–1.18)</td>
<td>0.83 (0.73–0.95)</td>
<td>0.0061</td>
</tr>
<tr>
<td>rs4150454</td>
<td>ERCC3</td>
<td>NER</td>
<td>Intron</td>
<td>0.39</td>
<td>0.44</td>
<td>1.01 (0.92–1.11)</td>
<td>0.89 (0.77–1.03)</td>
<td>0.82 (0.70–0.95)</td>
<td>0.0088</td>
</tr>
<tr>
<td>rs16500024</td>
<td>DMC1</td>
<td>HR</td>
<td>Intron</td>
<td>0.46</td>
<td>0.50</td>
<td>1.00 (0.92–1.09)</td>
<td>0.99 (0.88–1.12)</td>
<td>0.84 (0.73–0.96)</td>
<td>0.0088</td>
</tr>
<tr>
<td>rs2347443</td>
<td>CHEK2</td>
<td>Transduc</td>
<td>Intron</td>
<td>0.16</td>
<td>0.92</td>
<td>1.09 (0.98–1.22)</td>
<td>1.01 (0.85–1.19)</td>
<td>1.25 (1.06–1.48)</td>
<td>0.0090</td>
</tr>
<tr>
<td>rs797690</td>
<td>MSH6</td>
<td>MMR</td>
<td>5′ UTR</td>
<td>0.26</td>
<td>0.43</td>
<td>1.10 (1.00–1.21)</td>
<td>1.13 (0.98–1.30)</td>
<td>1.21 (1.05–1.40)</td>
<td>0.0096</td>
</tr>
</tbody>
</table>

Allelic trend test odds ratios (ORs) and 95% CI are summarized for each SNP for comparisons of the following groups with the comparison group with no cancer: (i) other cancer only, (ii) NMSC only and (iii) both NMSC plus another cancer; all results adjusted for residual confounding by ancestry by including first three principal components.

1BER, base excision repair; NHEJ, non-homologous end-joining; NER, nucleotide excision repair; DR, direct reversal; HR, homologous recombination; Transduc, DNA damage signal transduction; MMR, mismatch repair.

2Nonsyn, non-synonymous coding; 5′ UTR, 5′ untranslated region.

3MAF, minor allele frequency.

4P(HWE), P-value for Hardy–Weinberg equilibrium.

---

NA repair gene variants and NMSC cancer-prone phenotype
Table IV. Odds ratios (OR) and 95% Confidence Intervals (CI) for the association between SNPs with additive model $P$-value $< 0.01$ and risk of both NMSC plus other cancers compared with comparison group with no cancer, according to genotype and one or more versus zero minor alleles (dominant model) and two versus less than two minor alleles (recessive model)

<table>
<thead>
<tr>
<th>SNP</th>
<th>Gene</th>
<th>No. of Minor alleles (no. of both/no cancer)</th>
<th>Genotype OR (95% CI)</th>
<th>Dominant model2,3 OR (95% CI) $P$-value</th>
<th>Recessive Model2,4 OR (95% CI) $P$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs2888805</td>
<td>TDG</td>
<td>0 (432/1920)</td>
<td>1.0 (referent)</td>
<td>1.0 (referent)</td>
<td>1.0 (referent)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 (136/379)</td>
<td>1.57 (1.26–1.96)</td>
<td>1.53 (1.23–1.90) 0.0001</td>
<td>0.98 (0.45–2.14) 0.96</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2 (832)</td>
<td>1.07 (0.49–2.35)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs4135150</td>
<td>TDG</td>
<td>0 (433/1886)</td>
<td>1.0 (referent)</td>
<td>1.0 (referent)</td>
<td>1.0 (referent)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 (137/411)</td>
<td>1.58 (1.27–1.98)</td>
<td>1.54 (1.24–1.91) 0.0001</td>
<td>0.88 (0.39–2.01) 0.76</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2 (7/31)</td>
<td>0.97 (0.42–2.21)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs1038144</td>
<td>ERCC8</td>
<td>0 (452/1924)</td>
<td>1.0 (referent)</td>
<td>1.0 (referent)</td>
<td>1.0 (referent)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 (121/1354)</td>
<td>1.45 (1.15–1.83)</td>
<td>1.44 (1.15–1.81) 0.0017</td>
<td>1.03 (0.34–3.11) 0.96</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2 (4/15)</td>
<td>1.10 (0.36–3.33)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs152451</td>
<td>PALB2</td>
<td>0 (444/1893)</td>
<td>1.0 (referent)</td>
<td>1.0 (referent)</td>
<td>1.0 (referent)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 (126/384)</td>
<td>1.40 (1.12–1.75)</td>
<td>1.40 (1.12–1.74) 0.0033</td>
<td>1.25 (0.50–3.13) 0.64</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2 (6/19)</td>
<td>1.33 (0.53–3.35)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs447529</td>
<td>PALB2</td>
<td>0 (445/1892)</td>
<td>1.0 (referent)</td>
<td>1.0 (referent)</td>
<td>1.0 (referent)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 (126/384)</td>
<td>1.39 (1.11–1.75)</td>
<td>1.39 (1.11–1.74) 0.0036</td>
<td>1.24 (0.49–3.13) 0.64</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2 (6/19)</td>
<td>1.33 (0.53–3.34)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs11016857</td>
<td>MGMT</td>
<td>0 (216/714)</td>
<td>1.0 (referent)</td>
<td>1.0 (referent)</td>
<td>1.0 (referent)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 (268/1159)</td>
<td>0.76 (0.62–0.93)</td>
<td>0.75 (0.62–0.91) 0.0034</td>
<td>0.85 (0.66–1.08) 0.18</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2 (92/531)</td>
<td>0.72 (0.55–0.95)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs1450454</td>
<td>ERCC3</td>
<td>0 (229/631)</td>
<td>1.0 (referent)</td>
<td>1.0 (referent)</td>
<td>1.0 (referent)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 (276/1126)</td>
<td>0.82 (0.67–0.98)</td>
<td>0.80 (0.65–0.98) 0.0278</td>
<td>0.77 (0.58–1.02) 0.07</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2 (72/338)</td>
<td>0.67 (0.49–0.92)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs11650024</td>
<td>DMC1</td>
<td>0 (204/660)</td>
<td>1.0 (referent)</td>
<td>1.0 (referent)</td>
<td>1.0 (referent)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 (277/184)</td>
<td>0.77 (0.63–0.95)</td>
<td>0.76 (0.62–0.92) 0.0054</td>
<td>0.84 (0.66–1.06) 0.15</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2 (102/471)</td>
<td>0.72 (0.55–0.94)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs2347443</td>
<td>CHEK2</td>
<td>0 (379/1646)</td>
<td>1.0 (referent)</td>
<td>1.0 (referent)</td>
<td>1.0 (referent)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 (179/588)</td>
<td>1.33 (1.09–1.62)</td>
<td>1.32 (1.09–1.60) 0.0051</td>
<td>1.16 (0.68–1.98) 0.58</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2 (18/62)</td>
<td>1.26 (0.74–2.16)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs797690</td>
<td>MSH6</td>
<td>0 (305/1312)</td>
<td>1.0 (referent)</td>
<td>1.0 (referent)</td>
<td>1.0 (referent)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 (222/855)</td>
<td>1.11 (0.92–1.35)</td>
<td>1.19 (0.99–1.43) 0.0654</td>
<td>1.61 (1.14–2.26) 0.0063</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2 (50/129)</td>
<td>1.68 (1.18–2.38)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1Numbers do not sum to 5916 due to genotyping failures.
2Adjusted for residual confounding by ancestry by including first three principal components.
3One or two minor alleles versus zero minor alleles.
4Two minor alleles versus zero or one minor allele.
The other of rs152451 heterozygotes ranged from 9% to 22%, consistently of familial or bilateral breast cancer documented that the prevalence bioinformatics databases indicated it was likely to be benign. Studies the PALB2 protein, but queries of the PolyPhen-2, Condel and SWIFT repair genes may contribute to this general cancer-prone why a personal history of NMSC is associated with enhanced suscep-

cancer group may reflect a greater likelihood of genetically driven, environmental exposures likely play a more prominent role in markers associated with a predisposition to cancer. In contrast, the confinement of the risk associations to the NMSC plus other cancer group could be expected to predominate. This is analogous to other situations for which cancers are also more strongly genetically driven than the average-risk population, such as phenotypic markers defined by early age of cancer onset or a positive family history of cancer. Studying the relationship between NMSC and risk of other malignancies thus provides a fertile research model for identifying markers of potential susceptibility to human cancer.

The present study had multiple strengths, including addressing a novel question embedded within a well-established community-based cohort study. No previous study has tested a hypothesis to understand why a personal history of NMSC is associated with enhanced susceptibility to other cancers. The hypothesis that germ-line polymorphisms in DNA repair genes may contribute to this general cancer-prone phenotype was rigorously tested via thorough coverage of allelic variants in DNA repair pathway genes. Using the co-occurrence of NMSC and another cancer as a sentinel for a cancer-prone phenotype is novel and may enhance detection of common allelic variants that might otherwise escape notice as markers of cancer susceptibility. This is particularly true when one considers that in contrast to the complications usually encountered when studying multiple primary cancers, NMSC has the advantages that it is rarely fatal and therefore less susceptible to survival biases and, because most tumors are locally excised, its treatment is not associated with risk of other cancers.

Nevertheless, caution is warranted in drawing inferences. Even though a cohort of ~30,000 followed-up for almost two decades enables the conclusion of a total of 3,620 cancer cases, the primary study inferences centered on the both NMSC plus another cancer category, which was limited to 577 cancer cases. A resulting weakness was limited statistical power, a likely reason that no associations survived conservative significance thresholds imposed by corrections for multiple comparison testing. The sample size constraints of the both NMSC plus other cancer category highlight challenges of this line of inquiry and emphasize the need for replication of our findings. These same sample size constraints precluded meaningful subgroup analyses, such as separate analyses for BCC and SCC, but the totality of the current evidence suggests that the NMSC cancer-prone phenotype is consistently observed for both histologic types.

A limitation of the genetic association study approach is that we have identified statistical associations, but confirmation of the specific SNPs that might be involved would require additional experiments not included in the present study. In this discovery phase research, the results identified the genetic regions of DNA repair genes where variants are most likely to be associated with the risk of NMSC plus another cancer. If the observed associations are genuine, the association of these SNPs with the NMSC cancer-prone phenotype allows us to infer that allelic variations in TSG and the other genes identified may affect cancer development but it does not allow us infer causality for the SNPs themselves, because they are simply markers for genes regions that contain many other sequence variations. Thus, a limitation of this study is that further research will be needed to identify and characterize the specific functional variants within these gene regions and the protein change in function that might modify cellular DNA repair capacity. In the absence of direct experimental data, bioinformatics databases were queried to assess for the putative functionality of the two non-synonymous SNPs. For both TSG SNP rs2888805 and PALB2 SNP rs152451, the scores for the PolyPhen-2, Condel and SIFT bioinformatics databases were consistent in indicating that these SNPs were expected to be benign.

In conclusion, the findings have generated new leads suggesting that common variants in DNA repair genes may affect the increased overall cancer risk experienced by individuals with a personal history of NMSC. The strongest association was for a non-synonymous coding SNP in TSG, suggesting that base excision repair, possibly via regulation of the epigenome, may be important in the NMSC cancer-prone phenotype.

Supplementary material
Supplementary Tables 1–3 can be found at http://carcin.oxfordjournals. org/

Funding
National Cancer Institute (NCI) (R01 CA105069, HHSN26120080001E); Intramural Research Program of the NCI, Center for Cancer Research.

Acknowledgements
Cancer incidence data were provided by the Maryland Cancer Registry, Center for Cancer Surveillance and Control, Department of Health and Mental Hygiene, which is funded by the State of Maryland, the Maryland Cigarette Restraint
Fund and the National Program of Cancer Registries of the CDC to make the cancer registry data available. This publication does not necessarily reflect the views or policies of the NCI, NIMH, NIH, US DHHS, the US government or the Maryland Cancer Registry, nor does mention of trade names, commercial products or organizations imply endorsement by the US government.

Conflict of Interest Statement: None declared.

References


Received December 30, 2011; revised May 2, 2012; accepted May 8, 2012