Common genetic variants related to genomic integrity and risk of papillary thyroid cancer

Gila Neta¹, Alina V.Brenner, Eric M.Sturgis¹, Ruth M.Pfeiffer², Amy A.Hutchinson³, Briseis Aschebrook-Kilfoy⁴, Meredith Yeager³, Li Xu⁴, William Wheeler⁵, Michael Abend⁶, Elaine Ron¹, Margaret A.Tucker⁷, Stephen J.Chanock⁸ and Alice J.Sigurdson

Radiation Epidemiology Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Department of Health and Human Services, Rockville, MD 20852-7244, USA, ¹Department of Head and Neck Surgery, The University of Texas M. D. Anderson Cancer Center, Houston, TX, USA, ²Biostatistics Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Department of Health and Human Services, Rockville, MD, USA, ³Core Genotyping Facility, SAIC-Frederick Inc., NCI-Frederick, Frederick, MD, USA, ⁴Occupational and Environmental Epidemiology Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Department of Health and Human Services, Rockville, MD, USA, ⁵Information Management Services Inc., Silver Spring, MD, USA, ⁶Bundeswehr Institute of Radiobiology, Munich, Germany, ⁷Genetic Epidemiology Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Department of Health and Human Services, Rockville, MD, USA and ⁸Laboratory of Translational Genomics, Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Department of Health and Human Services, Gaithersburg, MD, USA

DNA damage is an important mechanism in carcinogenesis, so genes related to maintaining genomic integrity may influence papillary thyroid cancer (PTC) risk. Candidate gene studies targeting some of these genes have identified only a few polymorphisms associated with risk of PTC. Here, we expanded the scope of previous candidate studies by increasing the number and coverage of genes related to maintenance of genomic integrity. We evaluated 5077 tag single-nucleotide polymorphisms (SNPs) from 340 candidate gene regions hypothesized to be involved in DNA repair, epigenetics, tumor suppression, apoptosis, telomere function and cell cycle control and signaling pathways in a case–control study of 344 PTC cases and 452 matched controls. We estimated odds ratios for associations of single SNPs with PTC risk and combined P values for SNPs in the same gene region or pathway to obtain gene region-specific or pathway-specific P values using adaptive rank-truncated product methods. Nine SNPs had P values <0.0005, three of which were in HDAC4 and were inversely related to PTC risk. After multiple comparisons adjustment, no SNPs remained associated with PTC risk. Seven gene regions were associated with PTC risk at P < 0.01, including HUS1, ALKBH3, HDAC4, BAK1, FAA1, CHK2, DACT3 and FZD6. Our results suggest a possible role of genes involved in maintenance of genomic integrity in relation to risk of PTC.

Introduction

Thyroid cancer incidence has been steadily increasing worldwide and is now the fifth most common cancer diagnosed in women in the USA (1). Papillary thyroid cancer (PTC), the most common histological type, accounts for the majority of this increased incidence (2). Exposure to ionizing radiation at a young age is the strongest known risk factor for papillary thyroid cancer (3). In addition, family history studies have suggested that thyroid cancer has a greater familial risk than other common cancers with the reported relative risk estimates of 3- to 4-fold or higher (4–8). Because radiation is known to induce DNA damage, an important mechanism in carcinogenesis, heritable polymorphic variation in genes related to maintenance of genomic integrity has been hypothesized to affect risk of thyroid cancer. While two genome-wide association studies have identified only thyroid transcription factor gene polymorphisms both in persons exposed (9) and unexposed (10) to ionizing radiation, candidate gene studies have identified promising polymorphisms in other pathways including DNA repair and other genomic integrity-related pathways (11–20). For example, several studies reported associations with thyroid cancer risk for single-nucleotide polymorphism (SNPs) in the DNA repair genes XRCC1 (13–15,18) and XRCC3 (11,16,18) as well as other loci (RET, PTEN) involved in the mitogen-activated protein (MAP) kinase and AKT-signaling pathways (12,20). While providing initial clues to determinants of genetic susceptibility to PTC, these early studies were relatively small or had a limited selection of genes, suggesting that additional genetic determinants of PTC risk remain to be discovered.

Here, we expand the scope of previous candidate gene studies of thyroid cancer by increasing the number and coverage of genes belonging to several pathways related to maintenance of genomic integrity. Specifically, we examined risk of PTC in relation to 5077 common tag SNPs in a selected set of 340 genes involved in DNA repair, cell cycle control, apoptosis, tumor suppression, telomere function, epigenetics, Wnt/beta-catenin signaling, MAP kinase signaling and PI3K/AKT signaling using data from a common genotyping platform that was developed to evaluate genetic susceptibility to a variety of rare cancers (21,22).

Materials and methods

Study population

Our cases included thyroid cancers diagnosed within the US Radiologic Technologists (USRT) cohort as described previously (23) and cancers diagnosed and treated at the University of Texas M. D. Anderson Cancer Center (UTMDACC) (13). Controls for both types of cases were selected from the USRT study (24). Briefly, the USRT study was initiated in 1984 and included 146 022 radiologic technologists nationwide who were certified for at least 2 years by the American Registry of Radiologic Technologists between 1926 and 1982. Three questionnaires were administered (1984–1989, 1994–1998 and 2003–2005) to collect information on health outcomes (including self-reports of thyroid cancer), demographic characteristics, medical history, work practices and other environmental risk factors. Participant response to questionnaires has consistently been ~70%. Thyroid cancer cases reported on any of the three questionnaires were recruited for blood collection. Of 415 living thyroid cancer cases eligible for recruitment in 2005, 66 would not agree to release their medical records, and the pathology records of an additional 25 could not be obtained. In addition, there were 54 refusals and 19 non-responders to an invitation to provide blood samples and 19 who sent their samples after recruitment had closed, leaving 232 confirmed thyroid cancer cases who donated blood samples for the study (55.9% collected from all self-reported cases or treated at the University of Texas M. D. Anderson Cancer Center (UTMDACC) (13). Controls for both types of cases were selected from the USRT study (24). Briefly, the USRT study was initiated in 1984 and included 146 022 radiologic technologists nationwide who were certified for at least 2 years by the American Registry of Radiologic Technologists between 1926 and 1982. Three questionnaires were administered (1984–1989, 1994–1998 and 2003–2005) to collect information on health outcomes (including self-reports of thyroid cancer), demographic characteristics, medical history, work practices and other environmental risk factors. Participant response to questionnaires has consistently been ~70%. Thyroid cancer cases reported on any of the three questionnaires were recruited for blood collection. Of 415 living thyroid cancer cases eligible for recruitment in 2005, 66 would not agree to release their medical records, and the pathology records of an additional 25 could not be obtained. In addition, there were 54 refusals and 19 non-responders to an invitation to provide blood samples and 19 who sent their samples after recruitment had closed, leaving 232 confirmed thyroid cancer cases who donated blood samples for the study (55.9% collected from all self-reported cases or treated at the University of Texas M. D. Anderson Cancer Center (UTMDACC) (13). Each UTMDACC PTC case provided demographic and exposure information, including radiation...
of the genes obtained from various publically available databases, including NCBI Entrez Gene (www.ncbi.nlm.nih.gov/sites/entrez) and GeneCards (www.genecards.org). The DNA repair pathway was further divided into narrower subpathway categories including direct reversal of damage, base excision repair, homologous recombination, mismatch repair, non-homologous end joining, other conserved damage response genes, etc. When assigning gene regions with a broad range of known functions, we allowed for allocation to multiple pathways. A complete list of all candidate gene regions evaluated in current analyses and their pathway allocation is available in supplementary Table A, Carcinogenesis Online.

To minimize potential for population stratification and phenotypic heterogeneity of thyroid cancer cases, we also excluded in the analysis individuals with non-European ancestry (n = 97) and cases with follicular thyroid cancer (n = 17) leaving 344 papillary thyroid cancer cases (n = 202 USRT and n = 142 UTMDACC) and 452 controls of European ancestry with validated genotyping results. Allele frequencies for papillary thyroid cancer cases of European ancestry were largely similar between the USRT and the UTMDACC study sites and between males and females, so these groups were combined for genetic analyses.

Statistical analysis
We organized the analytic approach to proceed from examining the relationship of individual SNPs with PTC risk to examining relationships at the gene region level, pathway (e.g. apoptosis) and subpathway (e.g. direct reversal of DNA repair specific categories or subpathways) level and, finally, overall. We chose this approach to maximize our ability to detect small SNP effects that could only be appreciated in the aggregate. For examples, several SNPs in the same gene region might not individually contribute to PTC risk, but all of the SNPs with small effects taken together at the gene region level could be significant.

SNP-based associations. Logistic regression models were used to calculate odds ratios and 95% confidence intervals of the association of PTC risk with each SNP genotype, coded as 0, 1, 2, with 0 denoting the homozygous genotype with common allele as the referent category. We calculated the linear predictor \( P_{\text{linear}} \) for SNP genotype in crude models and models adjusted for sex, attained age in four categories (<35, 35–44, 45–54, 55+ years) and year of birth (<1940, 1940–1949, 1950+) as an ordinal variable. We also evaluated the effect of additional adjustment for cigarette smoking (ever/never), alcohol consumption (ever/never), and body mass index as a continuous variable. The results from models with additional adjustments were essentially similar and we chose to present the results from basic models adjusted for sex, attained age and year of birth throughout the manuscript. We adjusted for multiple comparisons using the false discovery rate control method (27).

Gene region- and pathway-based analyses. We combined SNP-specific P values of linear trend within the same gene region using an adaptive rank- truncated product method (28). This method accounts for the linkage disequilibrium (LD) structure within the gene region and allows for flexibility in assumptions about the number of SNPs to include in the \( P \) value calculation. For the subset of genes with the combined \( P \) value for gene region \( (P_{\text{region}}) < 0.01 \), we evaluated pairwise indices of LD (\( D^2 \) and \( r^2 \)) in controls using the Haplovoy package (29) and used a haplotype-sliding window approach (with windows composed of three SNPs) to evaluate potential loci in small genetic regions that may have been overlooked with a single locus analysis (30). Gene region level \( P \) values were combined into the \( P \) values associated with the nine pathways as well as the DNA repair specific categories or subpathways \( (P_{\text{pathway}}) \). Lastly, we evaluated the significance of the overall group of genomic integrity pathways combining all gene region level based \( P \) values \( (P_{\text{overall}}) \). Subpathway/pathway-based analyses were repeated including and excluding gene regions with multiple allocations. Statistical analyses were conducted in SAS version 9.1 (SAS Institute, Cary, NC) and in R, except where otherwise noted.

Results
The characteristics of 344 cases of PTC and 452 controls are summarized in Table I. There was a lower proportion of females among cases compared with controls (79.7 versus 93.6%). The distribution of referent age in cases and controls was comparable. On average, cases were less likely to smoke or drink alcohol or to have a family history of cancer but more likely to have a family history of thyroid cancer or a higher body mass index.

SNP-based associations
While the observed distribution of \( P \) values of linear trend for all 5077 SNPs was not statistically different from the expected uniform (null)
distribution, there was some suggestion of departure from the null in the area of lowest P values (Figure 1). Nine SNPs in six gene regions were associated with risk of PTC at \( P_{\text{trend}} < 0.0005 \) (Table II; a complete list of all SNP-based \( P_{\text{trend}} \) values is available in supplementary Table C, Carcinogenesis Online). Three of the nine SNPs (rs6749348, rs507159 and rs7584828) associated with reduced risk of PTC were in the \( \text{HDAC4} \) gene and did not appear to be in LD with one another (\( D^2 \) range 29–47, \( r^2 \) range 0.01–0.06) (Figure 2). The remaining six SNPs were in the genetic regions \( \text{HUS1} \) (rs2708906), \( \text{BAK1} \) (rs493871), \( \text{ALKBH3} \) (rs10838192), \( \text{DACT3} \) (rs314659), \( \text{MGMT} \) (rs4751109) and \( \text{FAF1_CDKN2C} \) (rs11587909). After multiple comparisons adjustment, none of the SNP-based \( P_{\text{trend}} \) values remained significant.

### Gene region-based associations

Consistent with individual SNP analyses, six gene regions (\( \text{HUS1}, \text{HDAC4}, \text{ALKBH3}, \text{BAK1}, \text{FAF1_CDKN2C}, \text{DACT3} \)) containing the most significant SNPs were associated with PTC risk at \( P_{\text{region}} < 0.01 \) (Table III). In addition, one gene (\( \text{FZD6} \)) containing several moderately significant SNPs was associated with PTC risk at \( P_{\text{region}} < 0.01 \). Two of the seven promising gene regions are involved in DNA repair (\( \text{HUS1}, \text{ALKBH3} \)), two in apoptosis (\( \text{BAK1}, \text{FAF1_CDKN2C} \)), two in Wnt/beta-catenin-signaling (\( \text{FZD6}, \text{DACT3} \)) and one in the epigenetic (\( \text{HDAC4} \)) pathway. After multiple comparisons adjustment, none of the gene regions were significant at \( P_{\text{region}} < 0.10 \). The results of all gene region-based analyses are available in supplementary Table D, Carcinogenesis Online. For a more detailed description of the top-ranked genes, see supplementary Table B, Carcinogenesis Online.

### Table I. Characteristics of the study population by case and control status, USRT Study and University of Texas M. D. Anderson Cancer Center Study

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Cases ((N = 344))</th>
<th>Controls ((N = 452))</th>
<th>(P)-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(n) (%)</td>
<td>(n) (%)</td>
<td></td>
</tr>
<tr>
<td>Radio logical Technologists</td>
<td>202 (58.7)</td>
<td>452 (100.0)</td>
<td></td>
</tr>
<tr>
<td>M. D. Anderson Cancer Center</td>
<td>142 (41.3)</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>274 (79.7)</td>
<td>423 (93.6)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Male</td>
<td>70 (20.3)</td>
<td>29 (6.4)</td>
<td></td>
</tr>
<tr>
<td>Attained/referent age, year</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19–25</td>
<td>26 (7.6)</td>
<td>30 (6.6)</td>
<td></td>
</tr>
<tr>
<td>26–35</td>
<td>73 (21.2)</td>
<td>103 (22.8)</td>
<td></td>
</tr>
<tr>
<td>36–45</td>
<td>113 (32.8)</td>
<td>159 (35.2)</td>
<td>0.360</td>
</tr>
<tr>
<td>46–55</td>
<td>77 (22.4)</td>
<td>105 (23.2)</td>
<td></td>
</tr>
<tr>
<td>56–65</td>
<td>43 (12.5)</td>
<td>49 (10.8)</td>
<td></td>
</tr>
<tr>
<td>66–79</td>
<td>12 (3.5)</td>
<td>6 (1.3)</td>
<td></td>
</tr>
<tr>
<td>Smoking status(^a)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never</td>
<td>208 (60.6)</td>
<td>234 (52.2)</td>
<td>0.001</td>
</tr>
<tr>
<td>Current</td>
<td>87 (25.4)</td>
<td>101 (22.5)</td>
<td></td>
</tr>
<tr>
<td>Number of relatives with cancer(^b)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>170 (49.4)</td>
<td>171 (37.8)</td>
<td></td>
</tr>
<tr>
<td>One</td>
<td>107 (31.1)</td>
<td>180 (39.8)</td>
<td></td>
</tr>
<tr>
<td>Two</td>
<td>39 (11.3)</td>
<td>64 (14.2)</td>
<td>0.002</td>
</tr>
<tr>
<td>Three or more</td>
<td>8 (2.5)</td>
<td>33 (7.3)</td>
<td></td>
</tr>
<tr>
<td>Unknown</td>
<td>10 (2.9)</td>
<td>4 (0.9)</td>
<td></td>
</tr>
<tr>
<td>Number of relatives with thyroid cancer(^b)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>319 (92.7)</td>
<td>442 (97.8)</td>
<td></td>
</tr>
<tr>
<td>At least one</td>
<td>15 (4.4)</td>
<td>6 (1.3)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Unknown</td>
<td>10 (2.9)</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Adopted</td>
<td>0</td>
<td>4 (0.9)</td>
<td></td>
</tr>
<tr>
<td>BMI ((\text{kg/m}^2))(^a)</td>
<td>26.1 (6.0)</td>
<td>24.4 (4.7)</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

\(^a\)BMI, body mass index; \(n\), not applicable.

\(^b\)As of referent age. See text for description of referent age assignment.

---

### Table II. Odds ratio and 95% confidence interval, adjusted for sex, attained age and year of birth, for the tag SNPs associated with papillary thyroid cancer risk at \( P_{\text{trend}} < 0.0005 \)

<table>
<thead>
<tr>
<th>Target gene</th>
<th>SNP</th>
<th>Genotype</th>
<th>Odds ratio (95% CI)</th>
<th>(P_{\text{trend}}/P_{\text{FDR}})</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\text{HDAC4})</td>
<td>rs6749348</td>
<td>GG</td>
<td>0.41 (0.26–0.65)</td>
<td>0.0004/0.2359</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AA</td>
<td>0.28 (0.03–2.46)</td>
<td></td>
</tr>
<tr>
<td>(\text{HDAC4})</td>
<td>rs507159</td>
<td>GC</td>
<td>0.68 (0.49–0.93)</td>
<td>0.0002/0.2359</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CC</td>
<td>0.39 (0.23–0.67)</td>
<td></td>
</tr>
<tr>
<td>(\text{HDAC4})</td>
<td>rs7584828</td>
<td>GA</td>
<td>0.55 (0.38–0.79)</td>
<td>0.0004/0.2775</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AA</td>
<td>0.33 (0.08–1.28)</td>
<td></td>
</tr>
<tr>
<td>(\text{HDAC4})</td>
<td>rs2708906</td>
<td>AG</td>
<td>1.55 (1.11–2.18)</td>
<td>0.0001/0.2359</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GG</td>
<td>2.40 (1.51–3.82)</td>
<td></td>
</tr>
<tr>
<td>(\text{BAK1})</td>
<td>rs493871</td>
<td>GA</td>
<td>1.41 (0.98–2.04)</td>
<td>0.0001/0.2359</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AA</td>
<td>2.29 (1.49–3.50)</td>
<td></td>
</tr>
<tr>
<td>(\text{ALKBH3})</td>
<td>rs10838192</td>
<td>TC</td>
<td>1.74 (1.25–2.41)</td>
<td>0.0003/0.2775</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CC</td>
<td>2.33 (1.05–5.17)</td>
<td></td>
</tr>
<tr>
<td>(\text{DACT3})</td>
<td>rs314659</td>
<td>GA</td>
<td>1.59 (1.16–2.18)</td>
<td>0.0004/0.2775</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AA</td>
<td>2.18 (1.26–3.77)</td>
<td></td>
</tr>
<tr>
<td>(\text{MGMT})</td>
<td>rs4751109</td>
<td>CA</td>
<td>1.66 (1.15–2.39)</td>
<td>0.0004/0.2775</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CC</td>
<td>2.50 (1.24–5.03)</td>
<td></td>
</tr>
<tr>
<td>(\text{FAF1_CDKN2C})</td>
<td>rs11587909</td>
<td>CT</td>
<td>1.36 (0.99–1.88)</td>
<td>0.0004/0.2775</td>
</tr>
</tbody>
</table>

CI, confidence interval.

\(^a\)SNP-based linear \( P_{\text{trend}} \) calculated based on the three-level genotype (0, 1 and 2) in logistic regression models adjusted for sex, attained age and year of birth.

\(^b\)False discovery rate (FDR) corrected linear \( P_{\text{trend}} \).
Pathway-based associations

The apoptosis pathway was the only pathway significantly associated with risk of PTC ($P_{\text{pathway}} = 0.039$) (Table IV). However, after excluding four gene regions (including two of the top gene region-based hits BAK1 and FAF1_CDKN2C) overlapping with other pathways, the apoptosis pathway-based $P$ value was no longer significant, suggesting that the excluded gene regions were primarily responsible for the association. While the DNA repair pathway as a whole was not statistically significantly associated with PTC risk, the DNA repair sub-pathways direct reversal of DNA damage and other conserved damage response genes, were associated with risk ($P_{\text{pathway}} = 0.002$ and $P_{\text{pathway}} = 0.017$, respectively). Finally, the genomic integrity pathway as a whole was of borderline statistical significance ($P_{\text{overall}} = 0.074$ and $P_{\text{overall}} = 0.061$ including and excluding overlap, respectively).

Discussion

We evaluated the associations of tag SNPs in candidate genes from several interrelated pathways involved in maintenance of genomic integrity, including DNA repair, epigenetic mechanisms, telomere function, apoptosis, cell cycle control, tumor suppression and the MAPK, PI3K/AKT and Wnt/beta-catenin cell-signaling pathways with risk of PTC. The group of genes as a whole was suggestively associated with risk of PTC ($P_{\text{overall}} = 0.074/P_{\text{overall}} = 0.061$). We
found nine tag SNPs in seven gene regions that were associated with PTC at $P_{\text{trend}} < 0.0005$. The strongest associations were seen for SNPs in the histone deacetylase 4 gene $\text{HDAC4}$, in the DNA repair checkpoint gene $\text{HUS1}$ and in the apoptosis gene $\text{BAK1} (P_{\text{trend}} < 0.0001)$. Gene region-based analyses showed that all three of these gene regions ($\text{HDAC4}$, $\text{HUS1}$ and $\text{BAK1}$) were significant at $P_{\text{region}} < 0.005$. While after formal correction for multiple comparisons neither individual SNP- nor gene region-based results remained statistically significant, several of our findings with suggestive $P$ values are of potential biologic or clinical interest.

Specifically, the observed thyroid cancer associations with multiple polymorphisms in $\text{HDAC4}$ were intriguing. $\text{HDAC4}$ is a histone deacetylation gene that has the capacity to alter chromosome structure and silence gene transcription by limiting access of transcription factors to DNA, particularly tumor suppressor genes, thereby deactivating tumor suppression activity (31,32). Histone deacetylase inhibitors are demonstrated anticancer agents whose main mechanism of action is the transcriptional reactivation of tumor suppressor genes that have been turned off through histone deacetylation (33). Several studies of thyroid tumor cells have demonstrated the ability of $\text{HDAC}$ inhibitors to facilitate radioactive iodine uptake (34–36) and suppress growth and proliferation (37–40). However, no studies have been published on associations between $\text{HDAC}$ polymorphisms and thyroid cancer risk.

$\text{HUS1}$ has been linked with other cancers, namely breast and ovarian (41,42), and is part of the Rad9-Rad1-Hus1 (911) cell cycle checkpoint point complex that plays a key role in all checkpoint responses to DNA damage (43). In vitro studies have shown that human cells exposed to ionizing and ultraviolet radiation have higher levels of the 911 protein complex compared with unexposed cells (44,45) and that this relationship is dose dependent. No studies previously have reported a relationship between $\text{HUS1}$ polymorphisms and thyroid cancer.

We also found variants in two genes postulated to play a role in apoptosis, $\text{BAK1}$ and $\text{FAF1}$, had suggestive $P$ values for an association with an increased risk of PTC. Previous research suggests that these genes may play a role in carcinogenesis of certain cancers, including testicular cancer and chronic lymphocytic leukemia (46,47), myeloma (48) and mantle cell lymphoma (49). Moreover, $\text{BAK1}$ expression appears to be upregulated in thyroid tumor cells (50–2).

Our strongest pathway-based findings for the DNA repair direct reversal of damage pathway (53) was driven by two of our top SNP-based findings—tag SNPs rs10838192 in $\text{ALKB3}$ and rs4751109 in $\text{MGMT}$. While little is known about $\text{ALKB3}$ in relation to cancer, several candidate gene studies have linked $\text{MGMT}$ polymorphisms to risk of head and neck cancer (54,55), glioma (56–58) and esophageal cancer (59,60). Moreover, two studies have found an association between $\text{MGMT}$ hypermethylation and PTC (61,62).
In interpreting the results of our study, several strengths and limitations need to be considered. Our study had high participation rates minimizing potential for selection bias. Because survival rates for PTC are exceptionally high, survival bias is unlikely. To minimize concerns about population stratification, all analyses were limited to individuals of European ancestry. Moreover, cases from the two studies were similar with regard to age at diagnosis, smoking status, tumor size and allele frequencies. Although radiation exposure is an established risk factor for thyroid cancer, our cases are unlikely to be radiation related because doses are low, most of the exposure occurred in adulthood, and no dose–response has been observed in the USRT population (63,64), and only five of the UTMDC cases were exposed to radiation from self-report of radiotherapy (12). Thus, the results of our study should be internally valid. However, extrapolation of our findings to the general population requires caution. Thyroid cancer incidence in the USRT cohort was higher than in the general population, with a standardized incidence rate ratio of 1.7. This difference is probably due to increased screening in the USRT cohort as the proportion of small thyroid tumors was higher among the cohort (30%) compared with the general population (15%) based on the SEER registries database (25) but comparable with that in the UTMDC cases.

Given that PTC cases were less likely to be cigarette smokers or alcohol drinkers and had higher average body mass index that controls, we explored whether these factors may have influenced the associations of interest. When added to the models, these variables did not meaningfully change the risk estimates, and therefore, are unlikely to confound our main findings. Other strengths of our study include thorough selection of genes related to a variety of genomic integrity pathways (53,65–67) and nearly complete representation of DNA repair genes. Relative to genome-wide association studies, the coverage of selected gene regions was higher, although we could have missed important associations with SNPs not included within the genotyping platform. While among the larger studies with respect to the number of thyroid cancer cases and controls, our study had limited power to detect weak associations, especially for less common genetic variants. Another limitation is the use of tag SNPs that are themselves unlikely to be the disease-related SNPs but are assumed to be in LD with the causal variant. To address these limitations, we excluded SNPs with minor allele frequency <10% and relied on robust gene/pathway adaptive rank-truncated product methods combining SNP-specific P values of trend to confirm associations with risk of PTC.

In summary, our results suggest that genetic alterations in the pathway involved in maintenance of genomic integrity may contribute to thyroid cancer susceptibility.

Supplementary material
Supplementary Tables A–D can be found at http://carcin.oxfordjournals.org/

Funding
This research was supported in part by the Intramural Research Program of the Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health and by a grant from the American Thyroid Association (Principal Investigator: E.M.S.). This project has been funded in whole or in part with federal funds from the National Cancer Institute, National Institutes of Health, under contract number HHSN261200800001E.

Acknowledgements
We are grateful to the radiologic technologists who participated in the USRT Study; Jerry Reid of the American Registry of Radiologic Technologists for continued support of this study; Diane Kampa and Allison Iwan of the University of Minnesota for data collection and study coordination; Liliana Mugartegui for patient recruitment, data collection and study coordination at the University of Texas M. D. Anderson Cancer Center; Laura Bowen of Information Management Systems for biomedical computing statistical support. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products or organizations imply endorsement by the US Government.

Conflict of Interest Statement: None declared.

References
Genomic integrity gene variants and thyroid cancer


Received February 4, 2011; revised May 10, 2011; accepted May 17, 2011