Apoptosis-related single nucleotide polymorphisms and the risk of non-small cell lung cancer in women

Anand Pathak^{1*}, Angela S. Wenzlaff², Paula L. Hyland³, Michele L. Cote², Greg R. Keele², Susan Land⁴, Matthew L. Boulton⁵ and Ann G. Schwartz²

*Correspondence: pathaka@mail.nih.gov

CrossMark

← Click for updates

¹Clinical Genetics Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, 9609 Medical Center Drive, Rockville, MD, USA.

²Population Studies and Disparities Research Program, Karmanos Cancer Institute, Wayne State University, Detroit, MI, USA.

³Genetic Epidemiology Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, 9609 Medical Center Drive, Rockville, MD, USA.

⁴Applied Genomics Technology Center and Department of Obstetrics & Gynecology, Wayne State University School of Medicine, CS Mott Center. Detroit MI, USA.

⁵Preventive Medicine Residency, Department of Epidemiology, University of Michigan School of Public Health, MI, USA.

Abstract

Background: Germline apoptosis-related single nucleotide polymorphisms (SNPs) have been shown to contribute to the risk of developing non-small cell lung cancer (NSCLC). However, very few studies have looked specifically at apoptosis-related SNPs in a racially-stratified analysis of white and African-American women.

Methods: We examined the risk of developing NSCLC associated with 98 germline SNPs in 32 apoptosis-related genes among women in a population-based case-control study from the Detroit metropolitan area. We examined 453 cases of NSCLC and 478 control subjects. We used an unconditional logistic regression with a dominant model, stratified by race, and adjusted for age, pack-years smoked, ever/never smoking status, family history of lung cancer, history of COPD, BMI and education.

Results: Our logistic regression identified 3 significant apoptosis-related SNPs in whites (*APAF-1*, rs1007573; *CD40* rs3765459, and *CD40* rs1535045), and 7 significant SNPs (*ATM*, rs1801516; *BAK1*, rs513349; *TNF*, rs1800629; *TP63*, rs6790167; *TP63*, rs7613791, *TP63*, rs35592567 and *TP63*, rs3856775) in African-Americans. In a downstream analysis, these SNPs were further prioritized utilizing the False Positive Report Percentage (FPRP) methodology and backwards elimination. In whites, *APAF-1* (rs1007573), *CD40* (rs3765459) and *CD40* (rs1535045) were all found to be significant by FPRP. In African-Americans, *TP63* SNPs rs6790167 and rs7613791 were found to have a significant FPRP. In parallel, a backward elimination procedure was used on the 3 significant SNPs in whites and 7 significant SNPs in African-Americans. This procedure identified *APAF-1* rs1007573 (OR=1.86, 95% CI:1.17-2.95) and *CD40* rs1535045 (OR=0.58, 95% CI:0.40-0.84) as significant independent predictors of risk among whites, and *ATM* rs1801516 (OR=24.15, 95% CI:3.50-166.55), *TNF* rs1800629 (OR=0.42, 95% CI:0.18-0.99) and *TP63* rs6790167 (OR: 2.85, 95% CI:1.33-6.09) as significant, independent predictors in African-Americans.

Conclusion: In whites, only SNPs *APAF-1* rs1007573 and *CD40* rs1535045 were significant by both FPRP and backwards elimination, while in African-Americans, only *TP63* rs6790167 was significant by both methodologies. Thus, we have identified three promising variants associated with increased risk of NSCLC that warrant additional investigation in future studies.

Keywords: Non-small cell lung cancer (NSCLC), single nucleotide polymorphisms, apoptosis, females, *APAF-1*, *CD40*, *ATM*, *TNF*, *TP63*

Introduction

Lung cancer is one of the leading causes of cancer in the United States with 159,480 deaths and 228,190 new cases estimated in 2013 [1]. While smoking is the principal risk factor for lung cancer, a genetic component is also known to contribute to risk. A better understanding of germline mutations that alter lung cancer risk may add to the future diagnosis and prevention of this disease. We hypothesize that attenuation of apoptosis-related pathways could contribute to uncontrolled cellular proliferation and the development of lung cancer. To date, a number of targeted studies have examined the relationship

between a limited number of apoptosis-related single nucleotide polymorphisms (SNPs) and lung cancer. Specifically, some focused studies have implicated SNPs in *CASP8* [2], *MDM2* [3], and *TGFB1* [4] as associated with altered risk of developing non-small cell lung cancer (NSCLC).

Genome wide association studies (GWAS) have permitted agnostic interrogation of the entire genome for the association between SNPs and NSCLC. A GWAS study performed on 2331 lung cancer cases and 3,077 controls in China revealed significant associations of SNPs in apoptosis-related genes (*TP63* and *TNFRSF19*) and NSCLC [5]. In a study of 1,952 European cases

and 1,438 European controls, a SNP in the apoptosis-related gene *BAT3* (BCL2-associated athanogene 6) was found to be significantly associated with lung cancer [6,7]. The remainder of the lung cancer GWAS studies did not implicate any other apoptosis-related gene, though nicotinic acid receptor genes [8], inflammatory genes such as *IL1RAP* and *CRP* [9] and a telomerase associated gene *TERT* [5,10,11] all were significantly associated with lung cancer. However, none of these GWAS studies included a focus on either African-Americans or women.

To evaluate the genetic contributions to lung cancer risk in African-American and white women, we examined 98 SNPs in 32 apoptosis-related genes, with the *a priori* hypothesis that variations in apoptosis-related genes may have a significant effect on the risk of NSCLC in women. The current state of the literature supports a prominent role for apoptosis in lung cancer and provides the rationale behind this apoptosis focused study.

Methods

Study participants

A population-based case-control study was conducted with the cases identified through the Metropolitan Detroit Cancer Surveillance System, part of the NCI's Surveillance, Epidemiology and End Results (SEER) program, as described previously [12]. Eligible cases were women between 18 to 74 years of age, diagnosed with histologically confirmed NSCLC between November 1, 2001 and October 31, 2005, and residing in Wayne, Oakland or Macomb counties. A total of 453 cases were used for this study. In total, 54.5% of eligible cases completed the in-person interview. Fifteen cases who reported race other than African-American or white were excluded from further analysis. Age (5-year intervals) and race-matched population-based controls living in the same geographic area were identified by random digit dialing. The majority (70%) of those who were eligible for screening agreed to participate in the same interview administered to the cases. Eleven controls reported a race other than African-American or white and were excluded from further analysis. Data from a total of 478 control subjects were utilized in the study.

Questionnaire data collection

Questionnaire data collection was carried out as described in a previous report from this study [12]. A detailed questionnaire covering personal health history, exposure history and family history was administered to consenting participants. Never-smokers were defined as those who had smoked <100 cigarettes in their lifetime. The diagnoses of COPD, chronic bronchitis and emphysema were combined to obtain a recoded composite COPD variable. Incident COPD diagnoses within the year prior to lung cancer diagnosis were not included. The COPD and BMI variables were by self-report. In addition, the education variable was categorized into less than 12 years, 12 to 16 years and greater than 16 years, to facilitate appropriate statistical analysis.

Blood collection and genotyping

After blood was drawn into EDTA containing Vacutainer Plus tubes, DNA was purified from whole blood using manufacturer's protocols for Qiagen AutoPure LS Genomic DNA Purification System (Gentra Systems). After purification, 250 ng of genomic DNA was submitted to the Applied Genomics Technology Core at Wayne State University. The Illumina Golden Gate assay Cancer SNP Panel (Catalog # GT-17-211) consisting of 1421 SNPs in 400 genes selected from the National Cancer Institute's Cancer Genome Anatomy SNP500 Cancer Database, was used. Data were analyzed using Illumina's BeadStudio software. Of the 1421 SNPs, 8 SNPs had more than 5% missing, thus 99.4% of the SNPs were called with greater than 95% of the samples genotyped.

Statistical analysis

Cases were first compared with controls by the Student's *t*-test stratified by race for age, pack-years smoked and body mass index (BMI). Differences in categorical variables (current smoker/never smoker, family history of lung cancer/no family history, COPD/no COPD, and education) were determined using a Chi-Square test. Hardy Weinberg Equilibrium (HWE) was tested in whites and African-Americans separately for 105 SNPs. Seven SNPs not in HWE were excluded from further analysis; a total of 98 SNPs were examined in this study. Next, we excluded any SNP that was not called in at least 90% of subjects. We then reran the clustering algorithm in BeadStudio before exporting the data. In addition, we calculated linkage disequilibrium R^2 values for all combinations of SNPs.

Controlling for age, smoking pack-years, ever/never smoking status, COPD history, family history of lung cancer, BMI and education, a dominant mode of transmission was modeled utilizing unconditional logistic regression in SAS version 9.1.3 (SAS Institute). The odds ratios obtained reflected the risk in the heterozygote and variant genotype group *versus* the homozygote "wild-type" genotype. Overfitting was assessed utilizing $1 - \hat{\gamma}$ [12,13]. The association between significant genes and ever/never smoking status was assessed by stratification by ever and never smoking status among whites only, as there were insufficient samples within African-Americans to support such stratification. Next, for the significant SNPs, we utilized backwards elimination, to discern which SNPs were significantly associated with NSCLC, while accounting for all significant SNPs ($p < 0.05$) among whites and African-Americans, respectively.

The False Positivity Report Probability (FPRP) was determined as described previously [14]. For this analysis, the *a priori* likelihood for the SNPs in *CD40* and *APAF-1* was set at 0.1 because of a lack of previous work suggesting association in SNPs in these genes and lung cancer. For *TP63*, there were several GWAS studies [5,10,15-17] that have shown strong evidence for association between *TP63* and lung cancer, so we judged it appropriate to set a strong *a priori* hypothesis of significance for this gene (0.25). A FPRP of <20% was

considered to be significant.

SNP function annotation

To explore whether any of the significant SNPs identified might have potential regulatory functions in lung and other tissues/cells, we used custom tracks on the UCSC Genome browser (<http://genome.ucsc.edu>) to screen NIH Roadmap and ENCODE data of the implicated SNP regions for evidence for regulatory relevance [18,19], such as overlapping with chromatin marks and interactions, CpG-site methylation, protein binding and transcription factor binding motifs. We also used the online tools HaploReg (<http://www.broadinstitute.org/mammals/haploreg/haploreg.php>) and RegulomeDB (<http://regulome.stanford.edu>) as a complementary analysis and to confirm the location of each SNP in relation to annotated protein-coding genes and/or non-coding RNA genes.

Informed consent

The study was approved by the Wayne State University institutional review board and each participant provided written informed consent.

Results

Participant characteristics

Key participant characteristics stratified by race are provided in **Table 1**. There were no significant differences in age between the cases and controls for both whites and African-Americans. Approximately 90% of the cases were current smokers among both whites and African-Americans; mean pack-years smoked was significantly greater in white and African-American cases relative to controls (46.17 vs. 12.49, $p < 0.001$ and 27.03 vs. 12.1, $p < 0.0001$, respectively). Cases were significantly more likely to report a family history of lung cancer relative to controls for both whites (24% versus 11.7%, $p < 0.0001$), and African-Americans (28.7% versus 10.8%, $p < 0.0015$). In addition, cases were more than twice as likely as controls to report a COPD diagnosis in whites (33.98% vs. 14.10%, $p < 0.0001$) and African-Americans (23.40% cases vs. 12.75% controls, $p < .05$). Controls had greater BMI than cases ($p < 0.0001$) among both whites and African-Americans. Finally, for the categorical education variable defined, white controls were significantly different from cases ($\chi^2 = 45.06$ $p < 0.0001$). However, this trend was not statistically significant among African-Americans ($\chi^2 = 5.0357$, $p = 0.08$).

Apoptosis-related SNPs and risk for NSCLC by race

We stratified our analysis by race since the allele frequency distribution for 74.3% of the SNPs we differed significantly between African-American and white controls, as measured by χ^2 analysis. **Table 2** summarizes significant SNPs for whites and African-Americans as determined by our stratified multivariate logistic regression model, after adjustment for the covariates of age, pack-years smoked, ever/never smoking status, family history of lung cancer, history of COPD, BMI

and education. **Supplement Table S1** list the odds ratios and p -values for all SNPs, among whites and African-Americans, respectively. The FPRP [14] was assessed for significant SNPs in whites and African-Americans. Among whites, SNPs in *CD40* (rs1535045 and rs3765459) and *APAF-1* (rs1007573) were significant with an *a priori* hypothesis of 0.1, having FPRP values $< 20\%$. Among African-Americans, two of the four *TP63* SNPs (rs6790167 and rs7613791), were found to have acceptable FPRP (less than 20%) for noteworthiness, given the strong *a priori* hypothesis (0.25) from the abundant literature implicating *TP63* in lung cancer. These *TP63* SNPs were ~12 kb apart at the 3' end of *TP63* gene, with an R^2 of 0.77 and a Lewontin's D' of 0.908. The *CD40* SNPs rs1535045 and rs3765459 were also in linkage disequilibrium ($R^2 = 0.87$); they were ~9.3 kb apart in an intronic region of *CD40*.

In parallel, we also conducted backwards elimination for all significant SNPs. In whites, *APAF-1* SNP rs1007573 and *CD40* SNP rs1535045 remained significant in the multivariable model. For African-Americans, the only SNPs that remained significant in the model after this selection procedure were *ATM* rs1801516, *TNF* rs1800629 and *TP63* rs6790167.

In summary, only two SNPs in whites were significant by both FPRP and backward elimination (*APAF-1* rs1007573 and *CD40* rs1535045), and only one SNP (*TP63* rs6790167) was significant by both methodologies in African-Americans. These three SNPs appear to be the most promising NSCLC risk-modifying apoptosis-related SNPs in this study.

Stratification by ever/never smoking status

To investigate whether there might be interaction between smoking status and the contribution of a SNP to the development of NSCLC, we built separate logistic regression models for never-smokers and ever-smokers among white subjects. Basic characteristics of this white population, stratified by smoking status, are listed in **Table 3**. An interaction was detected between smoking status and *APAF-1*. Among never-smokers, the association between carrying the heterozygote or variant homozygote genotype (rs1007573) and lung cancer was $OR = 0.72$ (95% CI: 0.28-1.88), while the OR for this SNP among smokers was 2.86 (95% CI: 1.58-5.16; $p = 0.01$), after adjustment. We found very similar associations for the *CD40* SNPs in both never-smokers and smokers; however, these findings were only significant for smokers. The OR was non-significant in *CD40* rs3765459 among never-smokers 0.62 (95% CI: 0.29-1.33) and significantly protective for smokers ($OR = 0.58$; 95% CI: 0.37-0.89). Analogous results were found for the *CD40* SNP rs1535045 ($OR = 0.66$; 95% CI: 0.31-1.40) for non-smokers and smokers ($OR = 0.52$; 95% CI: 0.33-0.80).

Functional annotation of SNPs

Based on ENCODE data, rs228652 (*TP63*), rs100753 (*APAF1*) and rs1535035 (*CD40*) overlapped with open chromatin in a number of ENCODE cell lines including A459 lung cancer cells (**Supplement Table S2** and **Figures 1** and **2**). Furthermore,

Table 1. Characteristics of Participants.

Variable	White			African Americans		
	Cases (n=359)	Controls (n=376)	P-value	Cases (n=95)	Controls (n=102)	P-value
Age (Mean, SD)	60.65 (9.4)	59.60 (9.2)	0.127	57.71 (9.3)	58.08 (9.4)	0.7856
Pack Years Smoked (Mean, SD)	46.17 (31.9)	12.49 (20.4)	<0.0001	27.03 (21.8)	12.1 (20.6)	<0.0001
Ever Smoker (n, %)	324 (90.3%)	185 (40.2%)	<0.0001	86 (91.5%)	54 (52.9%)	<0.0001
Family History of Lung Cancer (n, %)	86 (24.0%)	44 (11.7%)	<0.0001	27 (28.7%)	11 (10.8%)	0.0015
History of COPD* (n, %)	122 (34.0%)	53 (14.1%)	<0.0001	22 (23.4%)	13 (12.8%)	0.0516
BMI (Mean, SD)	26.24 (5.7)	28.71 (6.7)	<0.0001	26.73 (6.7)	32.31 (8.2)	<0.0001
Education (Mean, SD)	12.89 (2.3)	14.3 (2.7)	<0.0001	12.33 (2.7)	13.5 (2.6)	0.0303

*COPD: Chronic Obstructive Pulmonary Disease including diagnosis of emphysema and chronic bronchitis.

Table 2. Evaluation of SNPs significantly associated with non-small cell lung cancer with the False Discovery Reporting Percentage Method.

Race	Gene	Refseq	OR*** (95% CI)	0.1 A Priori FPRP (OR 2.0)	0.25 A Priori FPRP (OR 2.0)
Whites:	<i>APAF-1</i>	rs1007573	1.86 (1.17-2.95)	0.108*	0.0399**
	<i>CD40</i>	rs3765459	0.62 (0.42-0.90)	0.105*	0.038**
	<i>CD40</i>	rs1535045	0.58 (0.40-0.84)	0.045*	0.016**
African Americans:	<i>ATM</i>	rs1801516	24.15 (3.50-166.55)	0.659	0.392
	<i>BAK1</i>	rs513349	0.35 (0.12-0.98)	0.629	0.361
	<i>TNF</i>	rs1800629	0.42 (0.18-0.99)	0.548	0.288
	<i>P63</i>	rs6790167	2.85 (1.33-6.09)	0.254	0.102 **
	<i>P63</i>	rs7613791	2.70 (1.26-5.82)	0.31	0.13**
	<i>P63</i>	rs35592567	0.44 (0.20-0.95)	0.471	0.235
	<i>P63</i>	rs3856775	2.13 (1.03-4.43)	0.467	0.226

*Significant with an a priori hypothesis that the SNP may be associated with lung cancer.

**Significant with an a priori hypothesis that the SNP has a high likelihood of being associated with lung cancer.

*** Adjusted for age, pack years smoked, ever/never smoking status, family history of lung cancer, history of COPD, BMI and education.

Table 3. Characteristics of White Participants Stratified by Smoking Status.

Variable	White		
	Smokers (n=509)	Non-Smokers	P-value
Age (Mean, SD)	60.28 (9.3)	59.7 (9.2)	0.463
Family history of Lung Cancer (n, %)	95 (18.7%)	35 (15.5%)	0.298
History of COPD* (n,%)	152 (29.9%)	23 (10.9%)	<0.0001
BMI (Mean, SD)	26.89 (5.9)	28.89 (7.0)	<0.0001

*COPD: Chronic Obstructive Pulmonary Disease including diagnosis of emphysema and chronic bronchitis.

the DNA regions containing rs100753 (*APAF1*) and rs1535035 (*CD40*) overlapped with weak enhancer and/or promoter regions, respectively in normal fetal lung cells. Chromatin interaction paired-end tags (ChiA-PET) data can define two different DNA regions genomically far from each other or

on different chromosomes, but spatially close to each other in the nucleus due to interaction for regulatory functions. The DNA region containing rs100753 was enriched for CTCF binding in MCF-7 and K562 cells and also has the potential to form independent chromatin interactions/loops with the 3'UTR and the gene body of the downstream gene *ANKS1B*, albeit in K562 cells only. The chromatin interaction regions in K562 cells can be observed as open chromatin sites (DNaseI sites and FAIRE data) in A549 lung cancer cells (**Figure 1**). SNP rs1535035 (*CD40*) also mapped to a DNA region enriched for Pol2 enzyme in K562 and HCT-116 cells and characterized a strong chromatin interaction with the upstream promoter region of *NCOA5* (**Figure 2**). A summary of the analyses for each of the SNPs is presented in **Supplement Table 2**.

Discussion

The results of candidate gene SNP analysis [2-4] and large-scale GWAS studies [5-7,10] have both strongly pointed to apoptosis-related variations as key modifiers of lung cancer

Figure 1. UCSC Genome Browser track display of ENCODE and Roadmap data from specific cells and tissues.

UCSC Genome Browser image of 3' region of APAF-1 and gene body of ANKS1B on human assembly hg19 showing chr12q23.1 region, NHGRI SNPs, RefSeq Genes, CpG island annotation, Roadmap Chromatin State Segmentation by a Hidden Markov Model (HMM) for Roadmap fetal lung cells, ChIA-PET interactions and enrichment for CTCF in K562 and MCF-7 cells and open chromatin (peaks (Pk) and signals (DS and Sg) for DNaseI-seq and FAIRE data) in A549 lung cancer cells. The HMM utilizes computationally integrated ChIP-seq data for histone marks with associated segment color indicating a weak enhancer annotation for the overlapping SNP region in fetal lung cells/tissue. The overlapping SNP region is also marked by open chromatin in A549 cells and is enriched for CTCF binding and chromatin interactions in MCF-7 cells. Note all MCF-CTCF ChIA-PET interactions are not shown in full for this genomic region or the overlapping SNP region and the data segment has been adapted for the purpose of this illustration to show the relevant interactions only (bars connected by a line; intensity indicative of strength of interaction). The SNP rs1007573 is colored red.

risk. In this study, we assessed the significance of 98 SNPs in 32 apoptosis-related genes selected from the 400 genes on the Illumina GoldenGate Assay Cancer SNP Panel. Given that allele frequencies differed for 74% of the SNPs initially analyzed between African-Americans and whites, we stratified

analysis of our candidate SNPs by race.

Associations differed significantly by race. In total, only five SNPs in three genes reached a level of significance, as determined by the False Positivity Report Probability Method. One SNP in *APAF-1* and two SNPs in *CD40* were found to be

significant by FPRP among whites and two SNPs in *TP63* were found to be significant by FPRP among African-Americans. Fewer SNPs were significant by *both* FPRP and backwards elimination for lung cancer risk: two SNPs in whites (*APAF-1* rs1007573 and *CD40* rs1535045) and one SNP in African-Americans (*TP63* rs6790167).

We observed an association between SNPs in *TP63* (rs6790167: OR=2.85; 95% CI:1.33-6.09; rs7613791: OR=2.70; 95% CI:1.26-5.82) and NSCLC in African-Americans. These SNPs were located within introns at the 3' end of *TP63*. Previous GWAS studies have shown an association of *TP63* SNP rs4488809 in intron 1 and an increased risk of lung cancer [5,10]. *TP63* is a member of the p53 tumor suppressor gene family and is involved in response to cellular stress [10] and is known to be induced by DNA damage [20,21]. *TP63* activation may be an important part of the response mechanism to DNA damage caused by cigarette smoking.

The observed association between *APAF-1* SNP rs1007573 and NSCLC was significant by FPRP in white females. *APAF-1* plays a critical role in the apoptosome by activating cleavage of *CASP9* in the apoptosis pathway. *APAF-1* levels have been reported to be decreased [22,23] and increased [24] in NSCLC. Nuclear localization of *APAF-1* has been associated with improved 5-year survival and lower stage at diagnosis [25]. In addition, *APAF-1* has been implicated in activation of *CHK1* and cell cycle arrest under conditions of DNA damage [26], implicating a non-apoptosome related function for *APAF-1*. Taken together, *APAF-1* appears to be functionally important in the cellular response to stress; polymorphisms that decrease *APAF-1* levels and *APAF-1* nuclear localization may lead to more aggressive tumor formation. From this study, *APAF-1* SNP rs1007573, was only significant among smokers; thus, it is possible that the smoking associated mutagenic effects are further potentiated by aberrations in the *APAF-1* signaling axis. In K562 and MCF-7 cells, rs1007573 was shown to overlap with a region enriched for CTCF binding, which can act as a transcriptional repressor or insulator. The SNP overlapping region also has the potential to form a weak CTCF-mediated chromatin interaction with the proximal 3'UTR of ankyrin repeat and sterile alpha motif domain containing 1B (*ANKS1B*), as well as two stronger interactions with the gene body of *ANKS1B* or upstream of four of its variant mRNA transcripts. A recent paper has demonstrated a protective effect between SNPs in *ANKS1B* and lung cancer risk in Caucasians [27].

The protective effect we observed of *CD40* SNPs rs3765459 and rs1535045 against NSCLC has not been previously reported. In this study we observed that rs1535045 overlaps with an open chromatin site in A459 lung cancer cells. In other ENCODE cell lines this SNP region also overlapped with a repressive H3K27me3 mark and a H3K4me3 mark suggesting a cell-type-specific bivalent chromatin domain that is silenced, but poised for activation. Notably, the rs1535045 containing region is enriched for POL2 binding in both MCF-7 and HCT-116 cells, and in MCF-7 cells has the potential to form a POL2-mediated

chromatin loop with the promoter region of the nearby nuclear receptor coactivator 5 (*NCOA5*) gene. *CD40* is a member of the tumor necrosis factor (*TNF*) receptor superfamily (*TNFRSF*). Studies have shown that *CD40* is expressed in NSCLC and in lung cancer cell lines [28]. In a study of 129 surgical biopsy samples, increased *CD40* expression was associated with a poorer prognosis [28]. Other studies, however, have shown that gene transfer of *CD40* in NSCLC produces a direct inhibition of growth [29] and suppression of tumor proliferation [30]. This discrepancy regarding the function of *CD40* in NSCLC should be investigated in future studies. It is important to note that the protective effect of the *CD40* SNPs were only significant among smokers. The direction and magnitude of the effect was similar in never-smokers but numbers in this group were too small to make a definitive statement.

Strength and limitations

Our case-control study has several key strengths. The population-based in-person interview design, plus the collection of detailed data on exposure history, comorbidities and family history, permitted us to account for possible confounders that could obscure the relationship between candidate SNPs and NSCLC risk. In addition, we had a sizeable number of African-Americans in this study, permitting us to stratify our analysis by race. Furthermore, this study is unique in that it focuses primarily at women, and identifies germline NSCLC risk modifiers in this select population.

Our study has a number of limitations. First of all, in comparison to the GWAS studies, our sample size was relatively small, reducing our power to detect significant SNPs of small effect or low frequency alleles. Secondly, our interview response rates among the cases was relatively modest at 54.5% only. Third, we were only able to carry out stratified analysis by ever/never smoking status in whites because of limited sample size in African-Americans. Future large-scale studies should address the possible interaction between smoking and the presence of the *TP63* SNPs. The majority of our cases were lung adenocarcinomas and the results of our study may not be generalizable to other histologic subtypes of NSCLC [31]. It is important to note that none of the *p*-values adjusted for multiple comparisons in our study were significant. However, as noted by Wacholder et al., [14], the *a priori* hypothesis of a SNP's importance needs to be factored in when determining significance in a study with multiple comparisons. Thus, we utilized the Wacholder method, using existing evidence in the literature to find SNPs significant by FPRP.

Also, while we observe open chromatin domains in the UCSC tracks for A459 lung cancer cells, which coincided with the CHIA-PET protein binding and interactions data described for *APAF-1* rs1007573 and *CD40* rs1535045 from other cancer cell types; the potential effect of the variant G allele of rs1007573 or the T allele of rs1535045 on CTCF or POL2 binding, respectively and/or chromatin interactions across the specific genomic region in lung tissue is not known.

Testing these hypotheses and the tissue-specific nature of a potential regulatory function for these SNPs remain interesting biological questions for the future. Also, detailed mechanistic work examining the effect of these SNPs under conditions of cellular stress is needed.

Conclusions

Our investigation of apoptosis-related SNPs in this case-control study of African-American and white women from the Detroit Metropolitan area has identified three promising variants associated with increased risk of NSCLC (whites: *APAF-1* rs1007573 and *CD40* rs1535045; African-Americans: *TP63* rs6790167). These variants are significant by the FPRP methodology and also remained in the model as independent risk factor upon backwards elimination. Replication of our findings in larger, additional well-powered studies will be required to validate that these variants truly modify lung cancer risk. Like this study, future association studies should also account for important covariates that influence lung cancer risk. In summary, we have carried out a study that has rigorously assessed the lung cancer risk-modifying effects of apoptosis-related SNPs, while accounting for critical covariates, identifying 3 promising SNPs with potential regulatory functions that warrant further investigation in larger and more representative populations.

List of Abbreviations

SNPs: Single nucleotide polymorphisms
NSCLC: Non-small cell lung cancer
FPRP: False Positive Report Percentage
GWAS: Genome Wide Association Study

Additional files

[Supplement Table S1](#)
[Supplement Table S2](#)

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

Authors' contributions	AP	ASW	PLH	MLC	GRK	SL	MLB	AGS
Research concept and design	✓	✓	--	✓	--	✓	--	✓
Collection and/or assembly of data	--	✓	--	--	--	--	--	--
Data analysis and interpretation	✓	--	--	✓	--	--	--	✓
Writing the article	✓	--	✓	--	--	--	✓	✓
Critical revision of the article	✓	--	✓	--	--	--	✓	✓
Final approval of article	✓	--	--	--	--	--	✓	✓
Statistical analysis	✓	--	--	--	✓	--	--	--

Acknowledgement and funding

This work was supported by the National Institutes of Health [R01-CA87895 to A.G.S. and contracts N01-PC35145 to A.G.S. and P30CA22453]. This research was also supported in part by the

Intramural Research Program of the Division of Cancer Epidemiology and Genetics of the NIH and NCI.

Publication history

EIC: G.J. Peters, VU University Medical Center, Netherlands.

Received: 31-Aug-2013 Revised: 06-Jan-2014

Accepted: 17-Jan-2014 Published: 31-Jan-2014

References

1. Siegel R, Naishadham D and Jemal A. **Cancer statistics, 2013.** *CA Cancer J Clin.* 2013; **63**:11-30. | [Article](#) | [PubMed](#)
2. Hart K, Landvik NE, Lind H, Skaug V, Haugen A and Zienoldiny S. **A combination of functional polymorphisms in the *CASP8*, *MMP1*, *IL10* and *SEPS1* genes affects risk of non-small cell lung cancer.** *Lung Cancer.* 2011; **71**:123-9. | [Article](#) | [PubMed](#)
3. Liu G, Wheatley-Price P, Zhou W, Park S, Heist RS, Asomaning K, Wain JC, Lynch TJ, Su L and Christiani DC. **Genetic polymorphisms of *MDM2*, cumulative cigarette smoking and nonsmall cell lung cancer risk.** *Int J Cancer.* 2008; **122**:915-8. | [Article](#) | [PubMed](#)
4. Teixeira AL, Araujo A, Coelho A, Ribeiro R, Gomes M, Pereira C and Medeiros R. **Influence of *TGFβ1*+869T>C functional polymorphism in non-small cell lung cancer (NSCLC) risk.** *J Cancer Res Clin Oncol.* 2011; **137**:435-9. | [Article](#) | [PubMed](#)
5. Hu Z, Wu C, Shi Y, Guo H, Zhao X, Yin Z, Yang L, Dai J, Hu L and Tan W et al. **A genome-wide association study identifies two new lung cancer susceptibility loci at 13q12.12 and 22q12.2 in Han Chinese.** *Nat Genet.* 2011; **43**:792-6. | [Article](#) | [PubMed](#)
6. Wang Y, Broderick P, Webb E, Wu X, Vijayakrishnan J, Matakidou A, Qureshi M, Dong Q, Gu X, Chen WV, Spitz MR, Eisen T, Amos CI and Houlston RS. **Common 5p15.33 and 6p21.33 variants influence lung cancer risk.** *Nat Genet.* 2008; **40**:1407-9. | [Article](#) | [PubMed Abstract](#) | [PubMed Full Text](#)
7. Broderick P, Wang Y, Vijayakrishnan J, Matakidou A, Spitz MR, Eisen T, Amos CI and Houlston RS. **Deciphering the impact of common genetic variation on lung cancer risk: a genome-wide association study.** *Cancer Res.* 2009; **69**:6633-41. | [Article](#) | [PubMed Abstract](#) | [PubMed Full Text](#)
8. Liu P, Vikis HG, Wang D, Lu Y, Wang Y, Schwartz AG, Pinney SM, Yang P, de Andrade M and Petersen GM et al. **Familial aggregation of common sequence variants on 15q24-25.1 in lung cancer.** *J Natl Cancer Inst.* 2008; **100**:1326-30. | [Article](#) | [PubMed Abstract](#) | [PubMed Full Text](#)
9. Amos CI, Wu X, Broderick P, Gorlov IP, Gu J, Eisen T, Dong Q, Zhang Q, Gu X and Vijayakrishnan J et al. **Genome-wide association scan of tag SNPs identifies a susceptibility locus for lung cancer at 15q25.1.** *Nat Genet.* 2008; **40**:616-22. | [Article](#) | [PubMed Abstract](#) | [PubMed Full Text](#)
10. Miki D, Kubo M, Takahashi A, Yoon KA, Kim J, Lee GK, Zo JI, Lee JS, Hosono N, Morizono T, Tsunoda T, Kamatani N, Chayama K, Takahashi T, Inazawa J, Nakamura Y and Daigo Y. **Variation in *TP63* is associated with lung adenocarcinoma susceptibility in Japanese and Korean populations.** *Nat Genet.* 2010; **42**:893-6. | [Article](#) | [PubMed](#)
11. McKay JD, Hung RJ, Gaborieau V, Boffetta P, Chabrier A, Byrnes G, Zaridze D, Mukeria A, Szeszenia-Dabrowska N and Lissowska J et al. **Lung cancer susceptibility locus at 5p15.33.** *Nat Genet.* 2008; **40**:1404-6. | [Article](#) | [PubMed Abstract](#) | [PubMed Full Text](#)
12. Van Dyke AL, Cote ML, Wenzlaff AS, Chen W, Abrams J, Land S, Giroux CN and Schwartz AG. **Cytokine and cytokine receptor single-nucleotide polymorphisms predict risk for non-small cell lung cancer among women.** *Cancer Epidemiol Biomarkers Prev.* 2009; **18**:1829-40. | [Article](#) | [PubMed Abstract](#) | [PubMed Full Text](#)
13. Harrell FE, Jr., Lee KL and Mark DB. **Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors.** *Stat Med.* 1996; **15**:361-87. | [Pdf](#) | [PubMed](#)
14. Wacholder S, Chanock S, Garcia-Closas M, El Ghormli L and Rothman N. **Assessing the probability that a positive report is false: an approach for molecular epidemiology studies.** *J Natl Cancer Inst.* 2004; **96**:434-42. |

[Article](#) | [PubMed](#)

15. Lan Q, Hsiung CA, Matsuo K, Hong YC, Seow A, Wang Z, Hosgood HD, 3rd, Chen K, Wang JC and Chatterjee N et al. **Genome-wide association analysis identifies new lung cancer susceptibility loci in never-smoking women in Asia.** *Nat Genet.* 2012; **44**:1330-5. | [Article](#) | [PubMed](#)
16. Shiraishi K, Kunitoh H, Daigo Y, Takahashi A, Goto K, Sakamoto H, Ohnami S, Shimada Y, Ashikawa K, Saito A, Watanabe S, Tsuta K, Kamatani N, Yoshida T, Nakamura Y, Yokota J, Kubo M and Kohno T. **A genome-wide association study identifies two new susceptibility loci for lung adenocarcinoma in the Japanese population.** *Nat Genet.* 2012; **44**:900-3. | [Article](#) | [PubMed](#)
17. Wang Y, Broderick P, Matakidou A, Vijaykrishnan J, Eisen T and Houlston RS. **Variation in TP63 is associated with lung adenocarcinoma in the UK population.** *Cancer Epidemiol Biomarkers Prev.* 2011; **20**:1453-62. | [Article](#) | [PubMed](#)
18. Rosenbloom KR, Sloan CA, Malladi VS, Dreszer TR, Learned K, Kirkup VM, Wong MC, Maddren M, Fang R, Heitner SG, Lee BT, Barber GP, Harte RA, Diekhans M, Long JC, Wilder SP, Zweig AS, Karolchik D, Kuhn RM, Haussler D and Kent WJ. **ENCODE data in the UCSC Genome Browser: year 5 update.** *Nucleic Acids Res.* 2013; **41**:D56-63. | [Article](#) | [PubMed](#) | [Abstract](#) | [PubMed Full Text](#)
19. Meyer LR, Zweig AS, Hinrichs AS, Karolchik D, Kuhn RM, Wong M, Sloan CA, Rosenbloom KR, Roe G and Rhead B et al. **The UCSC Genome Browser database: extensions and updates 2013.** *Nucleic Acids Res.* 2013; **41**:D64-9. | [Article](#) | [PubMed Abstract](#) | [PubMed Full Text](#)
20. Katoh I, Aisaki KI, Kurata SI, Ikawa S and Ikawa Y. **p51A (Tap63gamma), a p53 homolog, accumulates in response to DNA damage for cell regulation.** *Oncogene.* 2000; **19**:3126-30. | [Article](#) | [PubMed](#)
21. Petitjean A, Ruptier C, Tribollet V, Hautefeuille A, Chardon F, Cavard C, Puisieux A, Hainaut P and Caron de Fromental C. **Properties of the six isoforms of p63: p53-like regulation in response to genotoxic stress and cross talk with DeltaNp73.** *Carcinogenesis.* 2008; **29**:273-81. | [Article](#) | [PubMed](#)
22. Krepela E, Prochazka J, Fiala P, Zatloukal P and Selinger P. **Expression of apoptosome pathway-related transcripts in non-small cell lung cancer.** *J Cancer Res Clin Oncol.* 2006; **132**:57-68. | [Article](#) | [PubMed](#)
23. Yang L, Mashima T, Sato S, Mochizuki M, Sakamoto H, Yamori T, Oh-Hara T and Tsuruo T. **Predominant suppression of apoptosome by inhibitor of apoptosis protein in non-small cell lung cancer H460 cells: therapeutic effect of a novel polyarginine-conjugated Smac peptide.** *Cancer Res.* 2003; **63**:831-7. | [Article](#) | [PubMed](#)
24. Krepela E, Prochazka J, Liul X, Fiala P and Kinkor Z. **Increased expression of Apaf-1 and procaspase-3 and the functionality of intrinsic apoptosis apparatus in non-small cell lung carcinoma.** *Biol Chem.* 2004; **385**:153-68. | [Article](#) | [PubMed](#)
25. Besse B, Cande C, Spano JP, Martin A, Khayat D, Le Chevalier T, Tursz T, Sabatier L, Soria JC and Kroemer G. **Nuclear localization of apoptosis protease activating factor-1 predicts survival after tumor resection in early-stage non-small cell lung cancer.** *Clin Cancer Res.* 2004; **10**:5665-9. | [Article](#) | [PubMed](#)
26. Zermati Y, Mouhamad S, Stergiou L, Besse B, Galluzzi L, Boehrer S, Pauleau AL, Rosselli F, D'Amelio M, Amendola R, Castedo M, Hengartner M, Soria JC, Cecconi F and Kroemer G. **Nonapoptotic role for Apaf-1 in the DNA damage checkpoint.** *Mol Cell.* 2007; **28**:624-37. | [Article](#) | [PubMed](#)
27. Lin J, Lu C, Stewart DJ, Gu J, Huang M, Chang DW, Lippman SM and Wu X. **Systematic evaluation of apoptotic pathway gene polymorphisms and lung cancer risk.** *Carcinogenesis.* 2012; **33**:1699-706. | [Article](#) | [PubMed](#) | [Abstract](#) | [PubMed Full Text](#)
28. Ishikawa K, Miyamoto M, Yoshioka T, Kato T, Kaji M, Ohbuchi T, Hirano S, Itoh T, Dosaka-Akita H and Kondo S. **Up-regulation of CD40 with juxtacrine activity in human nonsmall lung cancer cells correlates with poor prognosis.** *Cancer.* 2008; **113**:530-41. | [Article](#) | [PubMed](#)
29. Tong AW and Stone MJ. **Prospects for CD40-directed experimental therapy of human cancer.** *Cancer Gene Ther.* 2003; **10**:1-13. | [Article](#) | [PubMed](#)
30. Yamada M, Shiroko T, Kawaguchi Y, Sugiyama Y, Egilmez NK, Chen FA and

Bankert RB. **CD40-CD40 ligand (CD154) engagement is required but not sufficient for modulating MHC class I, ICAM-1 and Fas expression and proliferation of human non-small cell lung tumors.** *Int J Cancer.* 2001; **92**:589-99. | [Article](#) | [PubMed](#)

31. Shields PG. **Molecular epidemiology of smoking and lung cancer.** *Oncogene.* 2002; **21**:6870-6. | [Article](#) | [PubMed](#)

Citation:

Pathak A, Wenzlaff AS, Hyland PL, Cote ML, Keele GR, Land S, Boulton ML and Schwartz AG. **Apoptosis-related single nucleotide polymorphisms and the risk of non-small cell lung cancer in women.** *J Cancer Ther Res.* 2014; **3**:1.
<http://dx.doi.org/10.7243/2049-7962-3-1>