

A Large-scale genetic association study of esophageal adenocarcinoma risk

Chen-yu Liu¹, Michael C.Wu², Feng Chen¹,
Monica Ter-Minassian¹, Kofi Asomaning¹, Rihong Zhai¹,
Zhaoxi Wang¹, Li Su¹, Rebecca S.Heist^{1,3},
Matthew H.Kulke⁴, Xihong Lin², Geoffrey Liu^{1,5,6} and
David C.Christiani^{1,3,*}

¹Environmental and Occupational Medicine and Epidemiology Program, Department of Environmental Health, Harvard School of Public Health, Boston, MA 02115, USA, ²Department of Biostatistics, Harvard School of Public Health, Boston, MA 02115, USA, ³Department of Medicine, Massachusetts General Hospital, Harvard Medical School, Boston, MA02114, USA, ⁴Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, MA 02115, USA, ⁵Division of Medical Oncology and Haematology, Department of Medicine, Princess Margaret Hospital, Toronto, Ontario, Canada M5G 2M9 and ⁶Division of Applied Molecular Oncology, Department of Medical Biophysics, Ontario Cancer Institute, University of Toronto, Toronto, Ontario, Canada M5G 2M9

*To whom correspondence should be addressed. Environmental and Occupational Medicine and Epidemiology Program, Department of Environmental Health, Harvard School of Public Health, 665 Huntington Avenue, Building I Room 1401, Boston, MA 02115, USA.

Tel: +1 617 432 3323;

Fax: +1 617 432 3441;

Email: dchris@hsph.harvard.edu

The incidence of esophageal adenocarcinoma (EA) has been increasing rapidly, particularly among white males, over the past few decades in the USA. However, the etiology of EA and the striking male predominance is not fully explained by known risk factors. To identify susceptible genes for EA risk, we conducted a pathway-based candidate gene association study on 335 Caucasian EA cases and 319 Caucasian controls. A total of 1330 single-nucleotide polymorphisms (SNPs) selected from 354 genes were analyzed using an Illumina GoldenGate assay. The genotyped common SNPs include missense and exonic SNPs, SNPs within untranslated regions and 2 kb 5' of the gene, and tagSNPs for genes with little functional information available. Logistic regression adjusted for potential confounders was used to assess the genetic effect of each SNP on EA risk. We also tested gene-gender interactions using the likelihood ratio tests. We found that the genetic variants in the apoptosis pathway were significantly associated with EA risk after correcting for multiple comparisons. SNPs of rs3127075 in *Caspase-7 (CASP7)* and rs4661636 in *Caspase-9 (CASP9)* genes that play a critical role in apoptosis were found to be associated with an increased risk of EA. A protective effect of SNP rs572483 in the *progesterone receptor (PGR)* gene was observed among women carrying the variant *G* allele [adjusted odds ratio (OR) = 0.19; 95% confidence interval (CI) = 0.08–0.46] but was not observed among men (adjusted OR = 1.38; 95% CI = 0.95–2.00). In conclusion, this study suggests that the genetic variants of *CASP7* and *CASP9* in the apoptosis pathway may be important predictive markers for EA susceptibility and that *PGR* in the sex hormone signaling pathway may be associated with the gender differences in EA risk.

Introduction

The incidence of esophageal adenocarcinoma (EA) is increasing more rapidly than any other cancer in the USA (1,2) and is typically detected at an advanced stage with a 5 year mortality of >80% (3,4). EA is found seven times more often in men than in women in almost all

Abbreviations: CI, confidence interval; EA, esophageal adenocarcinoma; GERD, gastroesophageal reflux disease; *PGR*, *progesterone receptor*; OR, odds ratio; SNP, single-nucleotide polymorphism.

developed countries. In addition, it is very unusual for women before menopause to develop EA, suggesting the possibility that female sex hormones may protect against the development of EA (5–7). However, the etiology of EA and the striking male predominance is not fully understood.

Both genetic and environmental factors are suspected contributors to cancer development. Barrett's esophagus is a syndrome that develops among a subgroup of patients with chronic gastroesophageal reflux disease (GERD) (8). Although Barrett's esophagus is the only known premalignancy of EA, only a small fraction (0.5–1% annually) of Barrett's esophagus patients subsequently develop adenocarcinoma (9). The other potential risk factors include smoking, alcohol and obesity (10–13). However, only a fraction of individuals with these risk factors develop EA. For instance, the progression from GERD to Barrett's and Barrett's to carcinoma occurs among very few patients (14).

Systematic review of published common genetic variants and EA risk have been reported (15). While these studies have provided some promising results, a large-scale candidate gene analysis has not been performed to evaluate genetic susceptibility to EA risk. Therefore, we conducted a pathway-based candidate gene association study on 335 Caucasian EA cases and 319 Caucasian controls. A total of 1330 single-nucleotide polymorphisms (SNPs) in 354 genes were successfully genotyped using an Illumina GoldenGate assay.

Materials and methods

Study population and DNA sample collection

This study was approved by the Human Subjects Committees of Massachusetts General Hospital, Dana-Farber Cancer Institute and the Harvard School of Public Health (Boston, MA). Cases and controls were over the age of 18 years. Written informed consent was obtained from all subjects prior to study participation. Patients with incident, histologically confirmed EA were recruited at Massachusetts General Hospital between 1999 and 2005 and at Dana-Farber Cancer Institute between 2004 and 2005; patients with secondary or recurrent tumors were excluded. There was no restriction on tumor stage. Both Massachusetts General Hospital and Dana-Farber Cancer Institute are networked hospitals with similar practice patterns. Controls were selected from healthy friends and non-blood-related family members (usually spouses) of a multi-cancer susceptibility study conducted between 1999 and 2003 at Massachusetts General Hospital. All controls never had any diagnosis of cancer (16). Details of the control population have been described previously (16,17).

An in-person interview was conducted by a trained interviewer immediately after enrollment. A modified questionnaire (18) was used to obtain information on subjects' socio-demographic characteristics and a detailed smoking and alcohol exposure assessment.

Blood samples were collected from all participants at the time of recruitment. DNA was extracted from peripheral blood samples using the Puregene DNA Isolation Kit (Gentra Systems/QIAGEN, Valencia, CA).

SNP selection and genotyping

The selected candidate genes were in the pathways of carcinogens/procarcinogens metabolism, DNA repair, cell cycle control, apoptosis, inflammation, cell growth, cell signaling, angiogenesis, metastasis, sex hormone signaling, immunity, molecular transport, DNA methylation and telomere maintenance. These genes were selected based on published evidence of their relationship to carcinogenesis (19–30).

The genotyped common SNPs include missense and exonic SNPs, SNPs within untranslated regions and 2 kb 5' of the gene and tagSNPs for genes with little functional information available. The common non-synonymous SNPs were selected using SNP500Cancer Project (<http://snp500cancer.nci.nih.gov>) and International HapMap Project (<http://hapmap.org>). Potential functional non-synonymous SNPs from cancer-associated genes were selected from the PICS (Predicted Impact of Coding SNPs) database (http://www.icr.ac.uk/cancgen/molgen/MolPopGen_PICS_database.htm) (31) and FASTSNP (<http://fastsnp.ibms.sinica.edu.tw>) (32). SNPs on the Illumina Cancer Panel were selected with priority (http://www.illumina.com/products/cancer_snp_panel.ilmn). TagSNPs were selected using the r^2 -based Tagger program (33) with pairwise $r^2 \geq 0.80$ and minor allele frequency $\geq 5\%$ in the HAPMAP Caucasian population (CEU). SNPs probably to be problematic for genotyping

were eliminated based on the Illumina genotyping performance score for BeadArray assays.

Genotyping was performed on the Illumina GoldenGate assay at the Broad Institute, Cambridge, MA, by laboratory personnel without the knowledge of case-control status. Duplicate samples ($n = 48$) were randomly selected for quality control. The concordance rate of the replicate samples for all assays was >99%.

A total of 1536 SNPs were genotyped. After exclusion of 162 SNPs from unstable assays, 24 SNPs with low minor allele frequency (<0.01) and 20 SNPs with substantial deviation from Hardy–Weinberg equilibrium ($P < 0.0001$) and 1330 SNPs in 354 genes were left for analysis (Supplementary Table S1 is available at *Carcinogenesis* Online).

Statistical analysis

Among the study subjects genotyped, 31 (8%) of the EA cases and 21 (6%) of the controls were removed from the data analysis since they had >10% missing genotype information. We further restricted the analysis to Caucasians (>90% of the study subjects) and the study subjects with complete information on age, gender and smoking status (97% cases and 93% controls).

Genotype frequencies among controls were tested for departure from Hardy–Weinberg equilibrium by the Pearson χ^2 test (degree of freedom = 1). The selected principal characteristics were evaluated using χ^2 , Fisher exact and t -tests as appropriate.

Several potential confounders were adjusted for in the multiple regression models: age, gender, smoking status, body mass index at 18 years old, alcohol intake and prior history of GERD. As no confounding was seen (changes in 'estimated effects' <0.2) by these variables alone or together on the associations of interest, the results presented are only adjusted for age and gender.

To control the overall type I error rate while adjusting for correlation between SNPs, empirical P -values based on global random permutation tests were computed. Specifically, we jointly permuted the case-control status and the demographic covariates of each subject and recorded the minimal P -value for each permuted dataset. The distribution of the minimal P -values obtained from 10 000 permuted datasets was used to derive the empirical significance of the observed test statistic ($P_{\text{permutation}}$). The adjusted global-wide P -values were determined as $P_{\text{adjusted}} = P(P_{\text{observed}} \leq P_{\text{permutation}})$. All reported P -values are based on two-sided tests. An adjusted P -value <0.05 was considered to be statistically significance. Data were analyzed using the SAS® software version 9.1 (SAS Institute, Inc., Cary, NC) and R version 2.1.0 (The R Project for Statistical Computing, Vienna, Austria; <http://www.r-project.org/>).

Pathway-based analysis was performed by the logistic kernel machine regression test for genetic data (Michael, C., Wu, P.K., Michael, P. Epstein, Deanne, M. Taylor, Stephen, J. Chanock, David, J. Hunter and Xihong Lin, in preparation). Specifically, for each pathway, we analyzed all the SNPs within the pathway simultaneously and tested for their joint effect on the outcome, while adjusting for the potential confounders (age and gender). We applied the logistic kernel machine regression test of Liu *et al.* (34) using the identical-by-state kernel function (35) to measure genetic pathway similarity between subjects. This approach tests whether pairwise genetic similarity in the pathway (defined as the proportion of alleles shared identical-by-state across the pathway) is correlated with phenotypic similarity while adjusting for confounding covariates. Subjects with missing genotype values, for the SNPs in the pathway, were initially omitted from the pathway analysis. Pathway analyses in which the missing values were first imputed to the most common genotype (for each SNP) generated results that were concordant with the complete-case-only pathway analyses. The Bonferroni correction (36) was used to control for multiple comparisons.

Individual SNP analysis was performed among the 1330 SNPs by using multiple logistic regressions with additive genetic effect models. Dominant genetic effect models were used to test for interactions. Likelihood ratio tests were used to test the significance of the interaction terms.

Results

Demographics

A total of 335 cases and 319 controls were included in the final analyses. The distributions of selected characteristics among study subjects are summarized in Table I. Cases were older than controls and had a greater proportion of males. All cases and controls were Caucasians. Body mass index at 18 years was higher in cases than in controls. Not surprisingly, more smokers, alcohol drinkers and higher GERD prevalence were seen in the cases, as these are known risk factors for the development of EA.

Pathway and SNP effects

A total of 1330 SNPs in 354 genes were successfully genotyped. Under the pathway-based analysis, we categorized the 354 genes into

Table I. Distribution of study subjects by demographic and risk factors

Variable ^a	EA		
	Case ($n = 335$)	Control ($n = 319$)	P -value
Age (years) (range)	64 (31–91)	57 (30–83)	<0.0001
Sex			<0.0001
Male (%)	295 (88.1)	181 (56.7)	
Female (%)	40 (11.9)	138 (43.3)	
Smoking			<0.0001
Never (%)	66 (19.7)	112 (35.1)	
Ever (%)	269 (80.3)	207 (64.9)	
Body mass index at age 18			0.0003
<25 (%)	231 (69.0)	261 (82.3)	
25–29 (%)	84 (25.0)	47 (14.8)	
≥30 (%)	20 (6.0)	9 (2.9)	
Alcohol intake			0.35
Never (%)	26 (11.6)	41 (18.7)	
Ever (%)	199 (88.4)	178 (81.3)	
Prior history of GERD			<0.0001
No (%)	143 (49.8)	45 (65.2)	
Yes (%)	144 (50.2)	24 (34.8)	

^aNumber and percent of group (%). Percents are rounded.

Table II. Results of the pathway analyses

Pathway	Number of genes	Number of SNPs	Nominal P -value ^a	Bonferroni P -value
Carcinogens /procarcinogens metabolism	71	201	0.0119	0.1663
DNA repair	51	133	0.2218	1.0000
Cell cycle control	48	166	0.0417	0.5832
Apoptosis	41	234	0.0004	0.0062
Inflammation	37	224	0.0725	1.0000
Cell growth	30	63	0.6744	1.0000
Cell signaling	28	83	0.5267	1.0000
Angiogenesis	15	145	0.2296	1.0000
Metastasis	15	18	0.4344	1.0000
Sex hormone signaling	11	65	0.0583	0.8160
Immunity	10	20	0.8972	1.0000
Molecular transport	8	16	0.2100	1.0000
DNA methylation	4	14	0.2667	1.0000
Telomere maintenance	1	5	0.0174	0.2432

^aPathway-based analysis assessed by the logistic kernel machine regression adjusted for age and sex. Significant pathway effect after correcting for multiple comparisons is indicated in bold.

14 pathways (Table II): carcinogens/procarcinogens metabolism, DNA repair, cell cycle control, apoptosis, inflammation, cell growth, cell signaling, angiogenesis, metastasis, sex hormone signaling, immunity, molecular transport, DNA methylation and telomere maintenance. As expected, several genes overlap between pathways. Only the apoptosis pathway remained significant after correcting for multiple comparisons (nominal $P = 0.0004$, Bonferroni $P = 0.006$).

TagSNPs of rs3127075 in *CASP7* (nominal $P = 0.0001$, permutation $P < 0.05$) and rs4661636 in *CASP9* (nominal $P = 0.0004$, permutation $P = 0.2$) are the top SNPs in the apoptosis pathway associated with EA risk. The minor allele frequencies of *CASP7* rs3127075 and *CASP9* rs4661636 among the control population were 0.15 and 0.33, respectively. An increased risk was observed among subjects with polymorphisms rs3127075 in *CASP7* [odds ratio (OR) = 1.93; 95% confidence interval (CI) = 1.40–2.67] and rs4661636 in *CASP9* (OR = 1.60; 95% CI = 1.24–2.05) after adjustment for age and gender. No confounding effects of smoking status, body mass index at 18 years old, alcohol intake and prior history of GERD were observed by adjusting for these variables alone or together on the associations of interest.

We tested for possible joint effect between the polymorphisms rs3127075 in *CASP7* and rs4661636 in *CASP9* as caspase-7 is one of the downstream effectors activated by caspase-9 in the apoptosis pathway. Compared with the subjects with homozygous wild type of both *CASP7* rs3127075 and *CASP9* rs4661636, the subjects who carry the variant alleles of both *CASP7* rs3127075 and *CASP9* rs4661636 have 3.52 times higher risk of having EA (Table III).

Interaction effects

We conducted further analyses to evaluate the association of these SNPs with smoking and gender. There was no significant gene-smoking interaction effect after adjusting for multiple comparisons. Significant gene-gender interaction was observed among the polymorphism rs572483 in *PGR* (Table IV; nominal $P < 0.0001$, permutation $P < 0.05$). A protective effect was observed among women carrying the variant *G* allele (adjusted OR = 0.19; 95% CI = 0.08–0.46) but was not observed among men (adjusted OR = 1.38; 95% CI = 0.95–2.00).

Discussion

To our knowledge this is the first large-scale systematic pathway-based study to investigate the role of common genetic variation in susceptibility to EA. In this case-control study, we analyzed 14 pathways that are potentially important in carcinogenic processes and found that the genetic variants in the apoptosis pathway to be significantly associated with EA risk after correction for multiple comparisons. Polymorphisms of rs3127075 in *CASP7* (nominal $P = 0.0001$, permutation $P < 0.05$) and rs4661636 in *CASP9* (nominal $P = 0.0004$, permutation $P = 0.2$) are the top SNPs in the apoptosis pathway associated with EA risk.

Apoptosis, a genetically controlled process of programmed cell death, provides a protective mechanism by removing DNA-damaged cells that could either interfere with normal function or lead to neoplastic proliferation (37,38). Resistance to apoptosis is an important indicator related to esophageal carcinogenesis and chemotherapy resistance (38–42). Apoptosis activation occurs through the caspase

family in both mitochondrial and death receptor pathways (43). *CASP7* knockout mice were reported to be resistant to endotoxin-induced apoptosis (44). An inactivation mutation of the caspase genes, *CASP7*, has been found in esophageal cancer cells (42). It is plausible that SNPs in multiple genes in this pathway could affect cancer risk, as caspase-8, -9 and -10 initiate a cascade of caspase activation by cleaving downstream effectors such as caspase-7, which consequently causes cell death (45,46). Polymorphisms in *CASP7* and *CASP9* genes have found to be associated with several diseases including cancers (47–50) but have not been investigated in the EA risk. Polymorphisms of rs3127075 in *CASP7* and rs4661636 in *CASP9* are the top SNPs in the apoptosis pathway were associated with EA risk in our study. Our tagSNP *CASP7* rs12416109 is in complete linkage disequilibrium with *CASP7* rs2227309 ($r^2 = 1$; $D' = 1$), which has been found to be associated with different *CASP7* expression level and rheumatoid arthritis by Garcia-Lozano *et al.* (51), although the association of *CASP7* rs2227309 with rheumatoid arthritis was not replicated in another study (52). The tagSNP *CASP7* rs12416109 in our study did not show significant association with EA risk (data not shown). The polymorphism rs4661636 in *CASP9* was found to be associated with non-Hodgkin's lymphoma in a pooled analysis of three population-based case-control studies (53). In a Korean population, genetic polymorphisms in *CASP7* and *CASP9* were also found to be associated with survival in early-stage non-small-cell lung cancer and with worse outcome with the two genotypes combined (54). Different genetic polymorphisms in *CASP7* and *CASP9* have been found to be associated with these different diseases, possibly because the reported SNPs are in linkage disequilibrium with a nearby functional SNP. The observed associations in our study provide additional molecular epidemiologic evidence supporting the proposed role of apoptosis genes in EA risk.

Another interesting finding of the study is the gene-gender interaction effect of the *PGR* gene. A protective effect was observed among women carrying the variant *G* allele (adjusted OR = 0.19; 95% CI = 0.08–0.46) but was not observed among men (adjusted OR = 1.38; 95% CI = 0.95–2.00). Since the high male to female rate ratios are not probably to be explained by the different lifestyle risk factors among males and females (55), the different levels of hormones between genders may provide some further insights into the EA mechanisms. These results should be treated with caution due to the limited female cases in the current study. Further studies are needed to investigate the role of rs572483 in *PGR* gene in relation to the EA susceptibility.

The controls were selected to represent the general middle-aged adult distribution of Massachusetts population (56). While the case and control groups differed in several risk factors (Table I), we did not find the observed genetic associations confounded by these risk factors. No confounding bias was detected using multiple logistic regression models adjusting for these variables alone or in combination on the associations of interest.

It is important that the observed associations be confirmed in a larger, independent study. Although functionality is not known for all of the genotyped SNPs on our platform, our results are biologically plausible given the connections between the variants in apoptosis

Table III. The frequencies and adjusted ORs of genotypes of *CASP7* rs3127075 and *CASP9* rs4661636 to the susceptibility of EA

Polymorphism	Case (%)	Control (%)	Adjusted OR (95% CI) ^a
<i>CASP7</i> rs3127075			
<i>GG</i>	67 (20.0)	99 (31.0)	Ref
<i>GC or CC</i>	51 (15.2)	41 (12.9)	1.98 (1.13–3.47)
<i>GG</i>	122 (26.4)	133 (41.7)	1.51 (0.98–2.32)
<i>GC or CC</i>	95 (28.4)	46 (14.4)	3.52 (2.12–5.87)*
<i>CASP9</i> rs4661636			
<i>CC</i>			
<i>CT or TT</i>			

^aORs assessed by logistic regression adjusted for age and gender.
* $P < 0.0001$.

Table IV. The frequencies and adjusted odds ratio of *PGR* rs572483 polymorphism stratified by gender

Polymorphism	Women			Men		
	Case (%)	Control (%)	Adjusted OR (95%CI) ^a	Case (%)	Control (%)	Adjusted OR (95% CI) ^a
<i>PGR</i> rs572483 ^b						
<i>AA</i>	30 (75.0)	55 (39.9)	Ref	130 (44.7)	95 (52.5)	Ref
<i>AG or GG</i>	10 (25.0)	83 (60.1)	0.19 (0.08–0.46)*	161 (55.3)	86 (47.5)	1.38 (0.95–2.00)

^aORs assessed by logistic regression adjusted for age.

^bSignificant interaction effect (nominal $P < 0.0001$, permutation $P < 0.05$) according to the likelihood ratio tests.

* $P < 0.0005$.

genes and EA risk. However, associations with any specific SNP should be interpreted with caution until functionality is identified and these results are replicated.

Supplementary material

Supplementary Table S1 can be found at <http://carcin.oxfordjournals.org/>

Funding

National Institutes of Health (R01CA109193, R03CA110822, R01CA074386, ES00002); Canadian Institutes of Health Research operating grant; Kevin Jackson Memorial Fund; Alan B. Brown Chair in Molecular Genomics; National Cancer Institute of Canada Dorothy Lamont Award; Posluns Family Fund Foundation; Flight Attendant Medical Research Institute Award No. 062459. The Broad Institute Center for Genotyping and Analysis is supported by grant from the National Center for Research Resources (U54 RR020278).

Acknowledgements

The authors thank the physicians, staff and patients from the Massachusetts General Hospital Thoracic Oncology Center and the Dana-Farber Cancer Institute, Gastrointestinal Oncology Center for their cooperation and participation and Dr Daniel Mirel of the Broad Institute. We thank Andrea Shafer and Salvatore Mucci for technical assistance.

Conflict of Interest Statement: None declared.

References

- Farrow,D.C. *et al.* (1996) Determinants of survival following the diagnosis of esophageal adenocarcinoma (United States). *Cancer Causes Control*, **7**, 322–327.
- Pohl,H. *et al.* (2005) The role of overdiagnosis and reclassification in the marked increase of esophageal adenocarcinoma incidence. *J. Natl Cancer Inst.*, **97**, 142–146.
- Brown,L.M. *et al.* (2002) Epidemiologic trends in esophageal and gastric cancer in the United States. *Surg. Oncol. Clin. N. Am.*, **11**, 235–256.
- Holmes,R.S. *et al.* (2007) Epidemiology and pathogenesis of esophageal cancer. *Semin. Radiat. Oncol.*, **17**, 2–9.
- Devesa,S.S. *et al.* (1998) Changing patterns in the incidence of esophageal and gastric carcinoma in the United States. *Cancer*, **83**, 2049–2053.
- Bollschweiler,E. *et al.* (2001) Demographic variations in the rising incidence of esophageal adenocarcinoma in white males. *Cancer*, **92**, 549–555.
- Vizcaino,A.P. *et al.* (2002) Time trends incidence of both major histologic types of esophageal carcinomas in selected countries, 1973–1995. *Int. J. Cancer*, **99**, 860–868.
- Wang,K.K. *et al.* (2008) Updated guidelines 2008 for the diagnosis, surveillance and therapy of Barrett's esophagus. *Am. J. Gastroenterol.*, **103**, 788–797.
- Maley,C.C. *et al.* (2005) Natural selection in neoplastic progression of Barrett's esophagus. *Semin. Cancer Biol.*, **15**, 474–483.
- Cameron,A.J. *et al.* (1995) Adenocarcinoma of the esophagogastric junction and Barrett's esophagus. *Gastroenterology*, **109**, 1541–1546.
- Drewitz,D.J. *et al.* (1997) The incidence of adenocarcinoma in Barrett's esophagus: a prospective study of 170 patients followed 4.8 years. *Am. J. Gastroenterol.*, **92**, 212–215.
- Katzka,D.A. *et al.* (2000) Gastroesophageal reflux disease and Barrett's esophagus. *Med. Clin. North Am.*, **84**, 1137–1161.
- Sharma,P. *et al.* (1997) Dysplasia in short-segment Barrett's esophagus: a prospective 3-year follow-up. *Am. J. Gastroenterol.*, **92**, 2012–2016.
- Edelstein,Z.R. *et al.* (2007) Central adiposity and risk of Barrett's esophagus. *Gastroenterology*, **133**, 403–411.
- Cheung,W.Y. *et al.* (2009) Genetic variations in esophageal cancer risk and prognosis. *Gastroenterol. Clin. North Am.*, **38**, 75–91, viii.
- Miller,D.P. *et al.* (2002) Combinations of the variant genotypes of GSTP1, GSTM1, and p53 are associated with an increased lung cancer risk. *Cancer Res.*, **62**, 2819–2823.
- Miller,D.P. *et al.* (2006) An association between glutathione S-transferase P1 gene polymorphism and younger age at onset of lung carcinoma. *Cancer*, **107**, 1570–1577.
- Ferris,B.G. (1978) Epidemiology Standardization Project (American Thoracic Society). *Am. Rev. Respir. Dis.*, **118**, 1–120.
- Environmental Genome Project GeneSNPs*. [database on the Internet]. Utah Genome Depot at the University of Utah, Salt Lake City, UT. <http://www.genome.utah.edu/genesnps/> (16 June 2006, date last accessed).
- Kanehisa,M. *et al.* (2000) KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.*, **28**, 27–30. <http://www.genome.jp/kegg/pathway.html>.
- Biocarta* [database on the Internet]. Biocarta, San Diego, CA. <http://www.biocarta.com/genes/index.asp>.
- Maglott,D. *et al.* (2007) Entrez Gene: gene-centered information at NCBI. *Nucleic Acids Res.*, **35**, D26–D31.
- Rebhan,M. *et al.* (1997) GeneCards: integrating information about genes, proteins and diseases. *Trends Genet.*, **13**, 163. <http://www.genecards.org/index.shtml>.
- The Gene Ontology Consortium (2000). Gene ontology: tool for the unification of biology. *Nat Genet.*, **25**, 25–29.
- Gene Ontology Consortium. (2006) The Gene Ontology (GO) project in 2006. *Nucleic Acids Res.*, **34**(Database issue), D322–D326.
- Becker,K.G. *et al.* (2004) The genetic association database. *Nat. Genet.*, **36**, 431–432.
- Booker,S.M. (2001) Environmental genome project: a positive sequence of events. *Environ. Health Perspect.*, **109**, A22–A23.
- Kanehisa,M. *et al.* (2006) From genomics to chemical genomics: new developments in KEGG. *Nucleic Acids Res.*, **34**(Database issue), D354–D357.
- Rebhan,M. *et al.* (1998) GeneCards: a novel functional genomics compendium with automated data mining and query reformulation support. *Bioinformatics*, **14**, 656–664.
- Safran,M. *et al.* (2002) GeneCards 2002: towards a complete, object-oriented, human gene compendium. *Bioinformatics*, **18**, 1542–1543.
- Rudd,M.F. *et al.* (2005) The predicted impact of coding single nucleotide polymorphisms database. *Cancer Epidemiol. Biomarkers Prev.*, **14**(11 Pt 1), 2598–2604.
- Yuan,H.Y. *et al.* (2006) FASTSNP: an always up-to-date and extendable service for SNP function analysis and prioritization. *Nucleic Acids Res.*, **34**(Web Server issue), W635–W641.
- Ke,X. *et al.* (2003) Efficient selective screening of haplotype tag SNPs. *Bioinformatics*, **19**, 287–288.
- Liu,D. *et al.* (2008) Estimation and testing for the effect of a genetic pathway on a disease outcome using logistic kernel machine regression via logistic mixed models. *BMC Bioinformatics*, **9**, 292.
- Kwee,L.C. *et al.* (2008) A powerful and flexible multilocus association test for quantitative traits. *Am. J. Hum. Genet.*, **82**, 386–397.
- Adams,W.T. *et al.* (1987) Statistical test for the comparison of samples from mutational spectra. *J. Mol. Biol.*, **194**, 391–396.
- Raff,M. (1998) Cell suicide for beginners. *Nature*, **396**, 119–122.
- Thompson,C.B. (1995) Apoptosis in the pathogenesis and treatment of disease. *Science*, **267**, 1456–1462.
- Brabender,J. *et al.* (2009) Death-associated protein kinase (DAPK) promoter methylation and response to neoadjuvant radiochemotherapy in esophageal cancer. *Ann. Surg. Oncol.*, **16**, 1378–1383.
- Papineni,S. *et al.* (2009) Tolfenamic acid inhibits esophageal cancer through repression of specificity proteins and c-Met. *Carcinogenesis*, **30**, 1193–1201.
- Syrgos,K.N. *et al.* (2008) Targeted therapy for oesophageal cancer: an overview. *Cancer Metastasis Rev.*, **27**, 273–288.
- Soung,Y.H. *et al.* (2003) Inactivating mutations of CASPASE-7 gene in human cancers. *Oncogene*, **22**, 8048–8052.
- Boatright,K.M. *et al.* (2003) Caspase activation. *Biochem. Soc. Symp.*, **70**, 233–242.
- Lamkanfi,M. *et al.* (2009) Caspase-7 deficiency protects from endotoxin-induced lymphocyte apoptosis and improves survival. *Blood*, **113**, 2742–2745.
- Budihardjo,I. *et al.* (1999) Biochemical pathways of caspase activation during apoptosis. *Annu. Rev. Cell Dev. Biol.*, **15**, 269–290.
- Nagata,S. (1999) Biddable death. *Nat. Cell Biol.*, **1**, E143–E145.
- Gangwar,R. *et al.* (2009) Caspase 9 and caspase 8 gene polymorphisms and susceptibility to bladder cancer in north Indian population. *Ann. Surg. Oncol.*, **16**, 2028–34.

48. Lee, W.K. *et al.* (2009) Polymorphisms in the Caspase7 gene and the risk of lung cancer. *Lung Cancer*, **65**, 19–24.
49. Park, J.Y. *et al.* (2006) Caspase 9 promoter polymorphisms and risk of primary lung cancer. *Hum Mol Genet*, **15**, 1963–1971.
50. Xu, H.L. *et al.* (2009) Polymorphisms and haplotypes in the caspase-3, caspase-7, and caspase-8 genes and risk for endometrial cancer: a population-based, case-control study in a Chinese population. *Cancer Epidemiol. Biomarkers Prev.*, **18**, 2114–2122.
51. Garcia-Lozano, J.R. *et al.* (2007) Caspase 7 influences susceptibility to rheumatoid arthritis. *Rheumatology (Oxford)*, **46**, 1243–1247.
52. Teixeira, V.H. *et al.* (2008) Genetic and expression analysis of CASP7 gene in a European Caucasian population with rheumatoid arthritis. *J. Rheumatol.*, **35**, 1912–1918.
53. Lan, Q. *et al.* (2009) Genetic variation in caspase genes and risk of non-Hodgkin lymphoma: a pooled analysis of 3 population-based case-control studies. *Blood*, **114**, 264–267.
54. Yoo, S.S. *et al.* (2009) Polymorphisms in the CASPASE genes and survival in patients with early-stage non-small-cell lung cancer. *J. Clin. Oncol.*, **27**, 5823–5829.
55. Gammon, M.D. *et al.* (1997) Tobacco, alcohol, and socioeconomic status and adenocarcinomas of the esophagus and gastric cardia. *J. Natl Cancer Inst.*, **89**, 1277–1284.
56. Biener, L. *et al.* (2000) *Massachusetts Adult Tobacco Survey: Technical Report & Tables*. Center for Survey Research. University of Massachusetts, Boston, MA 2002.

Received January 20, 2010; revised April 2, 2010; accepted April 29, 2010