Distribution Agreement

In presenting this thesis as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis. I retain all ownership rights to the copyright of the thesis. I also retain the right to use in future works (such as articles or books) all or part of this thesis.

Signature:

_____ 4/13/2015

Biwen Tao Date

Effects of Genetic Susceptibility Variants on the Risk of Pancreatic Cancer and the Two-way Interactions between Single Nucleotide Polymorphisms (SNPs)

By

Biwen Tao

MPH

Department of Epidemiology

Timothy L. Lash, DSc, MPH

Committee Chair

Effects of Genetic Susceptibility Variants on the Risk of Pancreatic Cancer and the Two-way Interactions between Single Nucleotide Polymorphisms (SNPs)

By

Biwen Tao

M.S.

University of Missouri

2013

Thesis Committee Chair: Timothy L. Lash, DSc, MPH

An abstract of

A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Public Health in
Department of Epidemiology
2015

Abstract

Effects of Genetic Susceptibility Variants on the Risk of Pancreatic Cancer and the Two-way Interactions between Single Nucleotide Polymorphisms (SNPs)

By Biwen Tao

Pancreatic cancer is one of the most lethal cancers, with an estimated 5-year relative survival rate of 6% in the United States, the lowest among all cancer sites. Although the etiology of pancreatic cancer is still not clear, an increasing number of studies suggest that genetic mutations are associated with pancreatic carcinogenesis. In contrast to mutations with significant functional consequences on the protein, SNPs have been considered as functionally insignificant. Even though the individual effect of a SNP is relatively limited, the additive effect of the combinations of functionally related SNPs may synergistically contribute to risk of pancreatic cancer. Previous genome-wide association studies (GWAS) identified a pancreatic cancer susceptibility locus on chromosome 13q22.1, which was considered to be specific for pancreatic cancer. In order to evaluate the risk of pancreatic cancer contributed by individual SNPs, this study investigated 39 SNPs from 14 important genes and 1 highly associated non-genic locus, based on established evidence from published studies and their availability in the PanScan dataset. To better understand the role of SNPs on non-genic region of chromosome 13q22.1 that is specifically associated with pancreatic cancer, this study also assessed the interaction between these two SNPs with other selected susceptibility variants. After adjusting for age and sex, significant associations are observed between 13 SNPs and the risk of pancreatic cancer. A positive additive interaction between rs9564966 on 13q22.1 and rs3790844 on NR5A2 is identified on the risk of pancreatic cancer, with relative excess risk due to interaction (RERI) of 1.56. Further investigation of this interaction is required to evaluate the probability of false positive.

Effects of Genetic Susceptibility Variants on the Risk of Pancreatic Cancer and the Two-way Interactions between Single Nucleotide Polymorphisms (SNPs)

By

Biwen Tao

M.S.

University of Missouri

2013

Thesis Committee Chair: Timothy L. Lash, DSc, MPH

A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Public Health
in Department of Epidemiology
2015

Table of Contents

Introduction

Background

Pancreatic cancer is one of the most lethal cancers, with an estimated 5-year relative survival rate of 6% in the United States, the lowest among all cancer sites (Siegel, Ma, Zou, & Jemal, 2014). Due to the lack of efficient early detection of pancreatic cancer, most patients are diagnosed at late stage when it is merely feasible for surgical resection without curative effect. With the poor prognosis of pancreatic cancer, more than 80% of patients die within a year of diagnosis, and 98% die within 5 years (Wang, Schrag, Brooks, & Dominici, 2014). Elevated body-mass index, current or recent cigarette smoking, heavy alcohol consumption, a diagnosis of diabetes mellitus and family history of pancreatic cancer are considered to be associated with increased pancreatic cancer risk (Arslan et al., 2010; Everhart & Wright, 1995; Hassan et al., 2007; Lynch et al., 2009; Michaud et al., 2010).

Although the etiology of pancreatic cancer is still not clear, an increasing number of studies suggest that genetic mutations are associated with pancreatic carcinogenesis (Lin et al., 2011). A variety of environmental risk factors can trigger metabolic activation pathways to form DNA adducts and induce mutations in essential carcinogen-metabolizing genes (Bardeesy & DePinho, 2002). These genetic mutations can further cause unbounded cell growth and tumor development. Besides sporadic (non-familial) pancreatic cancer cases caused by genetic mutation, inherited genetic factors play an important role in the clustering of familial pancreatic cancers, which accounts for 5-10% of pancreatic cancer patients (Hruban, Canto, Goggins, Schulick, & Klein, 2010). In other

words, pancreatic cancer, like all cancers, is essentially a genetic disease caused by both inherited and acquired genetic mutation.

In addition to these rare mutations, aggregated studies have shown that commonly occurring (>1%) single nucleotide polymorphisms (SNPs) also contribute to the incremented risk of pancreatic cancer. Recent advances in genome-wide association studies (GWAS) improved the identification of relatively common variants that may be associated with the increasing risk of disease. The first GWAS for pancreatic cancer (PanScan I and II studies) examined 550,000 variants in 3,851 cases and 3,934 unaffected controls drawn from 12 prospective cohort studies and 8 case studies (Petersen et al., 2010). This study identified a pancreatic cancer susceptibility locus on chromosome 13q22.1, which was considered to be specific for pancreatic cancer. However, there was no investigation conducted to evaluate the interaction between this pancreatic cancer susceptibility locus and variants of other pancreatic cancer associated SNPs on the risk of pancreatic cancer.

In contrast to mutations with significant functional consequences on the protein, SNPs have been considered as functionally insignificant. Even though the individual effect of a SNP is relatively limited, the additive effect of the combinations of functionally related SNPs may synergistically contribute to risk of pancreatic cancer. The gene-gene interaction or epistasis has been identified in breast cancer, prostate cancer and colon cancer (Beuten et al., 2009; Goodman et al., 2006; Onay et al., 2006).

In order to evaluate the risk of pancreatic cancer contributed by individual SNPs, I have studied 39 SNPs from 14 important genes and 1 highly associated non-genic locus, based on established evidence from published studies and their availability in the PanScan

dataset. To better understand the role of SNPs on non-genic region of chromosome 13q22.1 that is specifically associated with pancreatic cancer, I assessed the interaction between these two SNPs with other selected susceptibility variants. Selected SNPs are involved in carcinogen metabolism (NAT2, GSTP1, CYP1B1), DNA repair (XRCC1, XRCC2, XRCC3, OGG1), susceptibility variants based on previous studies of PanScan (ABO, NR5A2, CLPTM1, 1q32.1), and susceptibility variants found in other studies (BRCA2, FOXQ1, BICD1, DPP6).

Several epidemiological studies have examined the association between variants of carcinogen metabolism genes and risk of pancreatic cancer, yielding conflicting results (Bartsch et al., 1998; Duell et al., 2002; Jang, Cotterchio, Borgida, Gallinger, & Cleary, 2012; Jiao et al., 2007; Li et al., 2006; Liu et al., 2000; Vrana, Novotny, Holcatova, Hlavata, & Soucek, 2010; Yamada et al., 2014). It is difficult to clarify the association between variants of carcinogen metabolism genes and the risk of pancreatic cancer from these inconsistent results investigated under limited sample sizes.

Detailed descriptions of selected SNPs, relevant genes and characteristics are listed in Table 1.

Table 1. Characteristics of selected single-nucleotide polymorphisms (SNP)

SNP	Gene/ Location	Characteristics
rs9543325 rs9564966	**13q22.1**	A genome-wide association study identified slightly higher association with pancreatic cancer in this non-genic region
rs3790844 rs10919791 rs3790843 rs12029406 rs4465241	**NR5A2/1q32.1**	Nuclear receptor of intracellular transcription factors, liver receptor homolog 1 (LRH-1)
rs401681	**CLPTM1/TERT/5p15.33**	Cleft lip and palate associated transmembrane protein 1

Table 1. Characteristics of selected single-nucleotide polymorphisms (SNP)

SNP	Gene/ Location	Characteristics
rs505922	**ABO**	Genome-wide association study identifies variants in the ABO locus associated with susceptibility to pancreatic cancer
rs144848 rs4987117	**BRCA2**	Associated with an increased risk of breast, ovarian, prostate, and pancreatic cancer, especially for individuals with family history
rs9502893	**FOXQ1**	Member of the Fox family of transcription factor genes, which is under-expressed in pancreatic cancer
rs708224	**BICD1**	Protein bicaudal D homolog 1, involved in COPI-independent membrane transport from the Golgi apparatus to the endoplasmic reticulum
rs6464375	**DPP6**	Dipeptidyl aminopeptidase-like protein 6, a single-pass type II membrane protein that is a member of the S9B family in clan SC of the serine proteases
rs1041983 rs1208 rs1799929 rs1799930 rs1799931 rs1801280	**NAT2**	Functions to both activate and deactivate arylamine and hydrazine drugs and carcinogens Polymorphisms in this gene are responsible for the N-acetylation polymorphism in which human populations segregate into rapid, intermediate, and slow acetylator phenotypes
rs1138272 rs1695 rs6591256	**GSTP1**	Play an important role in detoxification by catalyzing the conjugation of many hydrophobic and electrophilic compounds with reduced glutathione
rs1800440 Rs1056836 rs1056837 rs28936701	**CYP1B1**	Cytochrome P450 proteins are monooxygenases which catalyze many reactions involved in drug metabolism and synthesis of cholesterol, steroids and other lipids.
rs254870 rs2548724 rs2548754 rs254877 rs3213255	**XRCC1**	X-ray repair cross-complementing protein 1, involved in DNA repair where it complexes with DNA ligase III
rs2040639 rs3218408 rs3218536	**XRCC2**	DNA repair protein, homologous recombination to maintain chromosome stability and repair DNA damage
rs12432907 rs861539	**XRCC3**	DNA repair protein, homologous recombination to maintain chromosome stability and repair DNA damage
rs10521330 rs125701	**OGG1**	DNA glycosylase enzyme, involved in base excision repair

Specific Aims

1) Evaluate the association between genetic polymorphisms of candidate susceptibility variants and the risk of pancreatic cancer.

2) Assess the interaction between identified pancreatic cancer unique susceptibility locus 13q22.1 and other candidate susceptibility variants on the risk of pancreatic cancer.

Methods

Study Population

This study is conducted by using the Pancreatic Cancer Cohort Consortium (PanScan) consent group 2 from the database of Genotypes and Phenotypes (dbGaP). Datasets have been obtained from the dbGaP created by National Center for Biotechnology Information (NCBI). The data of this distribution set is collected from 2443 participants of case-control study design, including 1190 cases and 1253 controls. Cases are defined as those individuals having primary adenocarcinoma of the exocrine pancreas, excluding those with non-exocrine pancreatic tumors.

Each included study has obtained informed consent from participants and approval from its Institutional Review Board (IRB) for each individual study. IRB certification permitting data sharing was also obtained according to NIH Policy for Sharing of Data Obtained in NIH-Supported or –Conducted Genome-Wide Association Studies (GWAS).

Data Collection

There are 2502 DNA samples included in the consent group 2, extracted from blood products. These are collected for genotyping with quality control measures at the Core Genotyping Facility of the National Cancer Institute. After excluding samples with duplicates or that failed to meet a sample completion rate cutoff, the final participant count available for the association analysis is 1190 cases and 1253 controls. After quality control, 551,766 SNPs are available for the association analysis.

Definition of Variables

Demographics: age (10-year categories), sex

Single-nucleotide polymorphisms (SNPs): detailed descriptions of selected SNPs are listed in Table 1.

Statistical Analysis

Multivariate unconditional logistic regression analysis was conducted to obtain age and sex adjusted odds ratio (OR) estimates with 95% confidence intervals (CIs) for the association between selected SNPs and the risk of pancreatic cancer. A P-value <0.05 is considered as statistically significant. The chi-squared test is applied to evaluate Hardy-Weinberg equilibrium (HWE) in genotype frequencies. The effect modification between pancreatic cancer susceptibility locus 13q22.1 and variants of each selected SNP is evaluated by conducting stratified analyses and using the likelihood ratio statistic comparing models with and without the interaction term at 5% significant level. I have also examined the relative excess risk due to interaction (RERI). Statistical analyses is conducted using SAS version 9.4 (SAS Institute, Cary, NC).

Ethics Approval

Ethics approval for this project was obtained from the institutional review board (IRB) of Emory University, Atlanta, Georgia, the United States.

Results

Characteristics of Study Population

There were 1190 cases and 1253 controls included in the analysis (Table 2). The proportions of males and females are similar between cases and controls, with 46.6% males in cases and 47.9% males in cases. The distribution of age groups is different between case and control groups according to the chi-square test (p=0.0216). Case group has less people younger than fifty years old (8.7%) compared to control group (10.9%). The age groups of 51-60 (25.5%) and ≥81 (7.9%) are a larger proportion in case group than in control group (51-60: 22.6%; ≥81: 5.7%).

Table 2. Characteristics of pancreatic cancer cases and controls

Variables	Overall (N=2443)		Cases (N=1190)		Controls (N=1253)		Chi-square
	n	%	n	%	n	%	p-value
Age							0.0216
≤50	241	9.9%	104	8.7%	137	10.9%	
51-60	586	24.0%	303	25.5%	283	22.6%	
61-70	776	31.8%	380	31.9%	396	31.6%	
71-80	674	27.6%	309	26.0%	365	29.1%	
≥81	166	6.8%	94	7.9%	72	5.7%	
Sex							0.5374
Female	1155	47.3%	555	46.6%	600	47.9%	
Male	1288	52.7%	635	53.4%	653	52.1%	

Association of Individual SNPs and Risk of Pancreatic Cancer

Unconditional logistic regression analysis was performed for all selected 39 SNPs with adjustment for age and sex. Table 3 shows the adjusted OR estimates and 95% CIs for all selected SNPs under co-dominant models. The homozygous for major allele is used as reference group. The SNPs with significant associations at 5% significance level are in

bold. Both of the selected SNPs (rs9543325 and rs9564966) on chromosome 13q22.1 show significant associations with increased risk of pancreatic cancer (rs9543325, CC, OR=1.74 (1.38, 2.19); rs9564966, AA, OR=1.64 (1.29, 2.09)). Four SNPs (rs3790844, rs10919791, rs3790843 and rs12029406) of NR5A2 show significant associations with decreased risk of pancreatic cancer (rs3790844, TC, OR=0.74 (0.63, 0.88), CC OR=0.56 (0.40, 0.79); rs10919791, AG, OR=0.72 (0.61, 0.86), AA, OR=0.54 (0.38, 0.77); rs3790843, AG, OR=0.74 (0.62, 0.88), AA, OR=0.65 (0.49, 0.87); rs12029406, TC, OR=0.75 (0.63, 0.89), TT, OR=0.65 (0.52, 0.83)). The selected SNP (rs505922) of the ABO gene shows significant association with increased pancreatic cancer risk (TC, OR=1.26 (1.06, 1.50), CC, OR=1.35 (1.04, 1.75)). The SNP (rs401681) on CLPTM1 gene shows significant correlation with increased risk of pancreatic cancer (TC, OR=1.27 (1.05, 1.53), TT, OR=1.66 (1.32, 2.09)). These results are consisted with previous GWAS studies including the current study population as a part of analysis.

Four of the selected SNPs (rs1041983, rs1799929, rs1799930 and rs1801280) on NAT2 gene showed significant associations with decreased risk of pancreatic cancer (rs1041983, TC, OR=0.83 (0.70, 0.98); rs1799929, TC, OR=0.80 (0.67, 0.96); rs1799930, AG, OR=0.84 (0.71, 0.99); rs1801280, TC, OR=0.79 (0.65, 0.94)). The SNP rs125701 on OGG1 shows significant correlation with increased risk of pancreatic cancer (AA, OR=2.25 (1.29, 3.92)).

Although BRCA2 is recognized as a high penetrance gene for pancreatic cancer, there is no significant association detected between selected SNPs (rs144848 and rs4987117) and risk of among current study population. Other SNPs on the genes (FOXQ1, BICD1, DPP6, GSTP1, CYP1B1, XRCC1, XRCC2, and XRCC3) with published evidence of

susceptibility variants show no significant association with pancreatic cancer risk in the study population.

Table 3. Adjusted odds ratios of individual SNPs and the risk of pancreatic cancer

Genetic polymorphism	Genotype	Cases (N=1190)	Controls (N=1253)	Adjusted OR*	95% CI	
SNP		**%**	**%**			
13q22.1						
rs9543325	TT	10.81	19.66	1.00		
	TC	22.15	23.71	1.17	0.98	1.40
	CC	15.72	7.94	**1.74**	1.38	2.19
rs9564966	GG	18.63	22.03	1.00		
	AG	21.33	22.81	1.11	0.94	1.32
	AA	8.76	6.43	**1.64**	1.29	2.09
NR5A2/1q32.1						
rs3790844	TT	31.23	28.65	1.00		
	TC	15.10	18.79	**0.74**	0.63	0.88
	CC	2.37	3.85	**0.56**	0.40	0.79
rs10919791	GG	32.58	29.79	1.00		
	AG	13.87	17.77	**0.72**	0.61	0.86
	AA	2.26	3.73	**0.54**	0.38	0.77
rs3790843	GG	27.32	24.62	1.00		
	AG	17.45	21.34	**0.74**	0.62	0.88
	AA	3.89	5.37	**0.65**	0.49	0.87
rs12029406	CC	20.14	17.31	1.00		
	TC	21.04	24.15	**0.75**	0.63	0.89
	TT	7.53	9.82	**0.65**	0.52	0.83
rs4465241	CC	33.01	36.69	1.00		
	TC	14.00	13.43	1.16	0.97	1.38
	TT	1.68	1.19	1.59	0.98	2.59
ABO						
rs505922	TT	18.71	22.68	1.00		
	TC	23.58	22.80	**1.26**	1.06	1.50
	CC	6.43	5.81	**1.35**	1.04	1.75
BRCA2						
rs144848	TT	26.20	26.73	1.00		
	TG	19.07	20.38	0.94	0.80	1.11
	GG	3.44	4.18	0.84	0.62	1.15
rs4987117	CC	46.23	48.44	1.00		
	TC	2.42	2.70	0.95	0.66	1.36
	TT	0.08	0.12	0.70	0.12	4.20
CLPTM1/TER1/5p15.33						
rs401681	CC	12.24	16.21	1.00		
	TC	24.60	25.58	**1.27**	1.05	1.53
	TT	11.87	9.50	**1.66**	1.32	2.09
FOXQ1						
rs9502893	TT	13.88	15.76	1.00		
	TC	25.26	24.27	1.18	0.98	1.42
	CC	9.58	11.26	0.97	0.77	1.21
BICD1						

Genetic polymorphism	Genotype	Cases (N=1190)	Controls (N=1253)	Adjusted OR*	95% CI	
SNP		**%**	**%**			
rs708224	GG	15.88	17.07	1.00		
	AG	24.44	25.46	1.05	0.88	1.25
	AA	8.39	8.76	1.04	0.82	1.32
DPP6						
rs6464375	CC	44.13	45.19	1.00		
	TC	4.42	5.85	0.77	0.59	1.01
	TT	0.16	0.25	0.68	0.19	2.44
NAT2						
rs1041983	CC	22.91	22.75	1.00		
	TC	19.59	23.52	**0.83**	0.70	0.98
	TT	6.23	5.00	1.21	0.93	1.59
rs1208	AA	17.11	16.78	1.00		
	AG	21.98	25.42	0.86	0.72	1.03
	GG	9.62	9.09	1.06	0.84	1.33
rs1799929	CC	18.02	16.83	1.00		
	TC	21.46	25.35	**0.80**	0.67	0.96
	TT	9.25	9.09	0.96	0.77	1.21
rs1799930	GG	25.59	25.51	1.00		
	AG	18.14	21.70	**0.84**	0.71	0.99
	AA	4.95	4.10	1.19	0.89	1.58
rs1799931	GG	45.15	48.26	1.00		
	AG	3.48	2.87	1.29	0.93	1.79
	AA	0.08	0.16	0.46	0.08	2.54
rs1801280	TT	17.10	15.74	1.00		
	TC	21.66	25.69	**0.79**	0.65	0.94
	CC	9.99	9.82	0.95	0.76	1.19
GSTP1						
rs1138272	CC	41.18	43.92	1.00		
	TC	7.25	7.20	1.08	0.87	1.36
	TT	0.29	0.16	1.80	0.53	6.19
rs1695	AA	22.27	22.51	1.00		
	AG	21.37	22.84	0.95	0.80	1.12
	GG	5.08	5.94	0.86	0.66	1.12
rs6591256	AA	19.20	18.65	1.00		
	AG	22.68	23.75	0.93	0.78	1.11
	GG	7.01	8.71	0.78	0.61	1.00
CYP1B1						
rs1800440	AA	32.27	35.05	1.00		
	AG	14.99	14.37	1.13	0.95	1.35
	GG	1.43	1.88	0.82	0.52	1.29
rs1056836	CC	16.10	17.70	1.00		
	GC	23.23	23.76	1.08	0.90	1.29
	GG	9.38	9.83	1.05	0.84	1.32
rs1056837	CC	16.05	17.73	1.00		

Genetic polymorphism	Genotype	Cases (N=1190)	Controls (N=1253)	Adjusted OR*	95% CI	
SNP		**%**	**%**			
	TC	23.30	23.87	1.08	0.90	1.29
	TT	9.38	9.66	1.07	0.85	1.35
XRCC1						
rs254870	GG	18.30	18.79	1.00		
	AG	22.51	24.72	0.92	0.78	1.10
	AA	7.90	7.78	1.03	0.81	1.31
rs2548724	GG	30.37	29.72	1.00		
	AG	15.92	18.46	0.84	0.71	1.00
	AA	2.42	3.11	0.78	0.54	1.11
rs2548754	CC	15.53	17.37	1.00		
	TC	24.54	24.33	1.12	0.94	1.34
	TT	8.60	9.63	1.00	0.79	1.27
rs254877	TT	31.70	33.33	1.00		
	TC	15.19	16.34	0.97	0.82	1.15
	CC	1.80	1.64	1.14	0.74	1.78
rs3213255	TT	18.09	18.22	1.00		
	TC	22.68	24.56	0.92	0.77	1.09
	CC	7.94	8.51	0.94	0.74	1.19
XRCC2						
rs2040639	GG	15.44	16.30	1.00		
	AG	23.27	25.11	0.98	0.82	1.18
	AA	10.04	9.83	1.08	0.86	1.36
rs3218408	TT	29.96	31.56	1.00		
	TG	16.70	17.11	1.02	0.86	1.21
	GG	2.05	2.62	0.81	0.55	1.19
rs3218536	GG	42.33	43.68	1.00		
	AG	6.02	7.12	0.87	0.69	1.11
	AA	0.37	0.49	0.75	0.31	1.80
XRCC3						
rs12432907	CC	29.20	32.06	1.00		
	TC	17.20	16.99	1.11	0.94	1.31
	TT	2.29	2.25	1.13	0.77	1.66
rs861539	CC	20.92	21.04	1.00		
	TC	21.08	23.17	0.91	0.76	1.08
	TT	6.71	7.08	0.94	0.74	1.20
OGG1						
rs10521330	AA	34.34	35.65	1.00		
	AG	12.85	13.92	0.96	0.80	1.15
	GG	1.51	1.72	0.93	0.59	1.47
rs125701	GG	35.46	37.31	1.00		
	AG	11.63	13.19	0.93	0.77	1.12
	AA	1.64	0.78	**2.25**	1.29	3.92

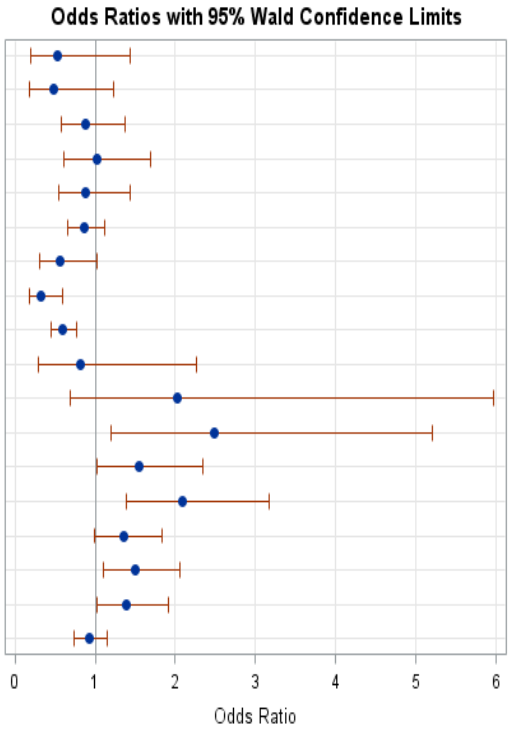
**Adjusted for age and sex*

Two-way Interaction between SNPs

The two-way interactions between two SNPs on 13q22.1 non-genic region and the rest of the SNPs was tested using multivariate logistic models. Assuming a multiplicative interaction effect on the logit scale, statistically significant interactions are selected using stepwise selection procedures at 5% significance level with likelihood ratio statistic. The only significant interaction term is observed between rs9564966 on 13q22.1 and rs3790844 on NR5A2 (p=0.0422).

Figure 1. Multiplicative interaction between rs9564966 on 13q22.1 and rs3790844 on NR5A2 on the risk of pancreatic cancer, OR of each stratum

Stratum	OR	95% CI	
rs3790844 CC vs TC at rs9564966=AA	0.54	0.20	1.43
rs3790844 CC vs TT at rs9564966=AA	0.48	0.18	1.24
rs3790844 TC vs TT at rs9564966=AA	0.89	0.57	1.38
rs3790844 CC vs TC at rs9564966=AG	1.02	0.62	1.70
rs3790844 CC vs TT at rs9564966=AG	0.88	0.54	1.44
rs3790844 TC vs TT at rs9564966=AG	0.86	0.67	1.12
rs3790844 CC vs TC at rs9564966=GG	0.56	0.30	1.03
rs3790844 CC vs TT at rs9564966=GG	0.33	0.18	0.60
rs3790844 TC vs TT at rs9564966=GG	0.59	0.45	0.78
rs9564966 AA vs AG at rs3790844=CC	0.82	0.29	2.27
rs9564966 AA vs GG at rs3790844=CC	2.03	0.69	5.97
rs9564966 AG vs GG at rs3790844=CC	**2.49**	**1.19**	**5.20**
rs9564966 AA vs AG at rs3790844=TC	1.55	1.03	2.34
rs9564966 AA vs GG at rs3790844=TC	2.10	1.38	3.17
rs9564966 AG vs GG at rs3790844=TC	1.35	1.00	1.83
rs9564966 AA vs AG at rs3790844=TT	1.51	1.11	2.05
rs9564966 AA vs GG at rs3790844=TT	1.40	1.02	1.91
rs9564966 AG vs GG at rs3790844=TT	**0.93**	**0.74**	**1.16**



To further investigate the statistically significant interaction term, the OR of each stratum is shown in Figure 1. There is no overlap of ORs' confidence intervals when rs9564966 AG vs GG stratified by rs3790844 of TT (OR=0.93 (0.74, 1.16)) and CC (OR=2.49 (1.19, 5.20)). The ratio of the ORs is significantly different from 1 (2.49/0.93=2.68), so there is modification on multiplicative odds ratio scale. Although the genotype AG of

rs9564966 doesn't have significant association with risk of pancreatic cancer, the presence of genotype CC of rs3790844 introduces a significant association between genotype AG of rs9564966 and increased risk of pancreatic cancer (OR=2.49 (1.19, 5.20)).

Relative excess risk due to interaction (RERI) is calculated based on the equation: $RERI=OR_{11}-OR_{10}-OR_{01}+1$ to test additive interaction. RERI=2.49-0.93-1.00+1= 1.56>0, so there is positive additive interaction between rs9564966 on 13q22.1 and rs3790844 on NR5A2 on the risk of pancreatic cancer.

Discussion

This study shows increased risk of pancreatic cancer associated with variants on 13q22.1, ABO, CLPTM1, and decreased risk associated with variants on NR5A2. These results are consistent with previous GWAS studies, which used the current study population as a part of analysis. This study also shows variants on carcinogen metabolism gene NAT2 gene are associated with decreased risk of pancreatic cancer, and increased risk with the variants on DNA repair OGG1 gene.

Previous studies indicate that BRCA2 is a high penetrance gene for elevated risk of pancreatic cancer, but this study failed to observe the association between selected SNPs (rs144848 and rs4987117) and pancreatic cancer. The rare occurrence of BRCA2 variants and its strong association to family history may make it difficult to detect under the frame of case-control study design. Previous reports on BRCA2 prevalence are mainly significant for patients with family history of pancreatic cancer.

The significant association between variants on DNA repair related gene OGG1 and incremented risk of pancreatic cancer suggests that such variant may affect the susceptibility of pancreatic cancer due to impaired pathway to remove oxidative DNA lesions (David, O'Shea, & Kundu, 2007).

Modest decrease in susceptibility to pancreatic cancer was observed among variants on carcinogen-metabolizing gene NAT2. No other association identified between variants on carcinogen-metabolizing genes and risk of pancreatic cancer. However, this may not be sufficient to rule out the association between variants on carcinogen-metabolizing genes and susceptibility of pancreatic cancer. Previous studies shows that the variants of

carcinogen-metabolizing genes may interact with environmental factors, including smoking and history of diabetes, to increase the risk of pancreatic cancer. Future studies should focus on its gene-environment interactions on the association with pancreatic cancer.

Although the statistical analysis shows significant association between identified SNPs and risk of pancreatic cancer, these variants may not necessarily be the causal factors for the pancreatic cancer. Further functional and fine-mapping studies are required to identify the causal variants and their role in the etiologic pathways.

The interaction between rs9564966 on 13q22.1 and rs3790844 on NR5A2 is statistically significant along with both additive and multiplicative effects. This interaction between these two SNPs is intricate. The variants on 13q22.1 tend to increase the risk of pancreatic while the variants on NR5A2 are likely to have a protective effect. The presence of minor allele of rs3790844 on NR5A2 makes heterozygous genotype of rs9564966 on 13q22.1 significantly associated with increased risk of pancreatic cancer. The individual variants of these two locus have opposite association with the risk of pancreatic cancer, but the combination of these two variants significantly increases the pancreatic cancer risk. Since the variants on 13q22.1 is nongenic region, it is difficult to evaluate potential pathway behind this statistically significant SNP-SNP interaction. Further investigation of this interaction is required to evaluate the probability of false positive.

References

Arslan, A. A., Helzlsouer, K. J., Kooperberg, C., Shu, X. O., Steplowski, E., Bueno-de-Mesquita, H. B., . . . Patel, A. V. (2010). Anthropometric measures, body mass index, and pancreatic cancer: a pooled analysis from the Pancreatic Cancer Cohort Consortium (PanScan). *Arch Intern Med, 170*(9), 791-802. doi: 10.1001/archinternmed.2010.63

Bardeesy, N., & DePinho, R. A. (2002). Pancreatic cancer biology and genetics. *Nat Rev Cancer, 2*(12), 897-909. doi: 10.1038/nrc949

Bartsch, H., Malaveille, C., Lowenfels, A. B., Maisonneuve, P., Hautefeuille, A., & Boyle, P. (1998). Genetic polymorphism of N-acetyltransferases, glutathione S-transferase M1 and NAD(P)H:quinone oxidoreductase in relation to malignant and benign pancreatic disease risk. The International Pancreatic Disease Study Group. *Eur J Cancer Prev, 7*(3), 215-223.

Beuten, J., Gelfond, J. A., Franke, J. L., Weldon, K. S., Crandall, A. C., Johnson-Pais, T. L., . . . Leach, R. J. (2009). Single and multigenic analysis of the association between variants in 12 steroid hormone metabolism genes and risk of prostate cancer. *Cancer Epidemiol Biomarkers Prev, 18*(6), 1869-1880. doi: 10.1158/1055-9965.epi-09-0076

David, S. S., O'Shea, V. L., & Kundu, S. (2007). Base-excision repair of oxidative DNA damage. *Nature, 447*(7147), 941-950. doi: 10.1038/nature05978

Duell, E. J., Holly, E. A., Bracci, P. M., Liu, M., Wiencke, J. K., & Kelsey, K. T. (2002). A population-based, case-control study of polymorphisms in carcinogen-

metabolizing genes, smoking, and pancreatic adenocarcinoma risk. *J Natl Cancer Inst, 94*(4), 297-306.

Everhart, J., & Wright, D. (1995). Diabetes mellitus as a risk factor for pancreatic cancer. A meta-analysis. *JAMA, 273*(20), 1605-1609.

Goodman, J. E., Mechanic, L. E., Luke, B. T., Ambs, S., Chanock, S., & Harris, C. C. (2006). Exploring SNP-SNP interactions and colon cancer risk using polymorphism interaction analysis. *Int J Cancer, 118*(7), 1790-1797. doi: 10.1002/ijc.21523

Hassan, M. M., Bondy, M. L., Wolff, R. A., Abbruzzese, J. L., Vauthey, J. N., Pisters, P. W., . . . Li, D. (2007). Risk factors for pancreatic cancer: case-control study. *Am J Gastroenterol, 102*(12), 2696-2707. doi: 10.1111/j.1572-0241.2007.01510.x

Hruban, R. H., Canto, M. I., Goggins, M., Schulick, R., & Klein, A. P. (2010). Update on familial pancreatic cancer. *Adv Surg, 44*, 293-311.

Jang, J. H., Cotterchio, M., Borgida, A., Gallinger, S., & Cleary, S. P. (2012). Genetic variants in carcinogen-metabolizing enzymes, cigarette smoking and pancreatic cancer risk. *Carcinogenesis, 33*(4), 818-827. doi: 10.1093/carcin/bgs028

Jiao, L., Doll, M. A., Hein, D. W., Bondy, M. L., Hassan, M. M., Hixson, J. E., . . . Li, D. (2007). Haplotype of N-acetyltransferase 1 and 2 and risk of pancreatic cancer. *Cancer Epidemiol Biomarkers Prev, 16*(11), 2379-2386. doi: 10.1158/1055-9965.epi-06-0992

Li, D., Jiao, L., Li, Y., Doll, M. A., Hein, D. W., Bondy, M. L., . . . Hassan, M. M. (2006). Polymorphisms of cytochrome P4501A2 and N-acetyltransferase genes, smoking,

and risk of pancreatic cancer. *Carcinogenesis, 27*(1), 103-111. doi: 10.1093/carcin/bgi171

Lin, Y., Yagyu, K., Egawa, N., Ueno, M., Mori, M., Nakao, H., . . . Kikuchi, S. (2011). An overview of genetic polymorphisms and pancreatic cancer risk in molecular epidemiologic studies. *J Epidemiol, 21*(1), 2-12.

Liu, G., Ghadirian, P., Vesprini, D., Hamel, N., Paradis, A. J., Lal, G., . . . Foulkes, W. D. (2000). Polymorphisms in GSTM1, GSTT1 and CYP1A1 and risk of pancreatic adenocarcinoma. *Br J Cancer, 82*(10), 1646-1649. doi: 10.1054/bjoc.2000.1221

Lynch, S. M., Vrieling, A., Lubin, J. H., Kraft, P., Mendelsohn, J. B., Hartge, P., . . . Stolzenberg-Solomon, R. Z. (2009). Cigarette smoking and pancreatic cancer: a pooled analysis from the pancreatic cancer cohort consortium. *Am J Epidemiol, 170*(4), 403-413. doi: 10.1093/aje/kwp134

Michaud, D. S., Vrieling, A., Jiao, L., Mendelsohn, J. B., Steplowski, E., Lynch, S. M., . . . Stolzenberg-Solomon, R. Z. (2010). Alcohol intake and pancreatic cancer: a pooled analysis from the pancreatic cancer cohort consortium (PanScan). *Cancer Causes Control, 21*(8), 1213-1225. doi: 10.1007/s10552-010-9548-z

Onay, V. U., Briollais, L., Knight, J. A., Shi, E., Wang, Y., Wells, S., . . . Ozcelik, H. (2006). SNP-SNP interactions in breast cancer susceptibility. *BMC Cancer, 6*, 114. doi: 10.1186/1471-2407-6-114

Petersen, G. M., Amundadottir, L., Fuchs, C. S., Kraft, P., Stolzenberg-Solomon, R. Z., Jacobs, K. B., . . . Chanock, S. J. (2010). A genome-wide association study

identifies pancreatic cancer susceptibility loci on chromosomes 13q22.1, 1q32.1 and 5p15.33. *Nat Genet, 42*(3), 224-228. doi: 10.1038/ng.522

Siegel, R., Ma, J., Zou, Z., & Jemal, A. (2014). Cancer statistics, 2014. *CA Cancer J Clin, 64*(1), 9-29. doi: 10.3322/caac.21208

Vrana, D., Novotny, J., Holcatova, I., Hlavata, I., & Soucek, P. (2010). CYP1B1 gene polymorphism modifies pancreatic cancer risk but not survival. *Neoplasma, 57*(1), 15-19.

Wang, Y., Schrag, D., Brooks, G. A., & Dominici, F. (2014). National trends in pancreatic cancer outcomes and pattern of care among Medicare beneficiaries, 2000 through 2010. *Cancer, 120*(7), 1050-1058. doi: 10.1002/cncr.28537

Yamada, I., Matsuyama, M., Ozaka, M., Inoue, D., Muramatsu, Y., Ishii, H., . . . Lin, Y. (2014). Lack of associations between genetic polymorphisms in GSTM1, GSTT1 and GSTP1 and pancreatic cancer risk: a multi-institutional case-control study in Japan. *Asian Pac J Cancer Prev, 15*(1), 391-395.