Genome-Wide Gene–Diabetes and Gene–Obesity Interaction Scan in 8,255 Cases and 11,900 Controls from PanScan and PanC4 Consortia

Hongwei Tang[1], Lai Jiang[2], Rachael Z. Stolzenberg-Solomon[3], Alan A. Arslan[4,5,6], Laura E. Beane Freeman[3], Paige M. Bracci[7], Paul Brennan[8], Federico Canzian[9], Mengmeng Du[10], Steven Gallinger[11], Graham G. Giles[12,13,14], Phyllis J. Goodman[15], Charles Kooperberg[16], Loïc Le Marchand[17], Rachel E. Neale[18], Xiao-Ou Shu[19], Kala Visvanathan[20,21], Emily White[22], Wei Zheng[19], Demetrius Albanes[3], Gabriella Andreotti[3], Ana Babic[23], William R. Bamlet[24], Sonja I. Berndt[3], Amanda Blackford[21], Bas Bueno-de-Mesquita[25,26,27,28], Julie E. Buring[2,29], Daniele Campa[30], Stephen J. Chanock[3], Erica Childs[21], Eric J. Duell[31], Charles Fuchs[32,33,34], J. Michael Gaziano[29,35], Michael Goggins[36], Patricia Hartge[3], Manal H. Hassam[1], Elizabeth A. Holly[7], Robert N. Hoover[3], Rayjean J. Hung[11], Robert C. Kurtz[37], I.-Min Lee[2,29], Núria Malats[38], Roger L. Milne[12,13,14], Kimmie Ng[23], Ann L. Oberg[24], Irene Orlow[10], Ulrike Peters[22], Miquel Porta[39,40], Kari G. Rabe[24], Nathaniel Rothman[3], Ghislaine Scelo[8], Howard D. Sesso[2,29], Debra T. Silverman[3], Ian M. Thompson Jr[41], Anne Tjønneland[42], Antonia Trichopoulou[43], Jean Wactawski-Wende[44], Nicolas Wentzensen[3], Lynne R. Wilkens[17], Herbert Yu[17], Anne Zeleniuch-Jacquotte[5,6], Laufey T. Amundadottir[3], Eric J. Jacobs[42], Gloria M. Petersen[24], Brian M. Wolpin[23], Harvey A. Risch[45], Nilanjan Chatterjee[46], Alison P. Klein[21,36], Donghui Li[1], Peter Kraft[2,47], and Peng Wei[48]

ABSTRACT

Background: Obesity and diabetes are major modifiable risk factors for pancreatic cancer. Interactions between genetic variants and diabetes/obesity have not previously been comprehensively investigated in pancreatic cancer at the genome-wide level.

Methods: We conducted a gene–environment interaction (GxE) analysis including 8,255 cases and 11,900 controls from four pancreatic cancer genome-wide association study (GWAS) datasets (Pancreatic Cancer Cohort Consortium I–III and Pancreatic Cancer Case Control Consortium). Obesity (body mass index ≥30 kg/m^2) and diabetes (duration ≥3 years) were the environmental variables of interest. Approximately 870,000 SNPs (minor allele frequency ≥0.005, genotyped in at least one dataset) were analyzed. Case–control (CC), case-only (CO), and joint-effect test methods were used for SNP-level GxE analysis. As a complementary approach, gene-based GxE analysis was also performed. Age, sex, study site, and principal components accounting for population substructure were included as covariates. Meta-analysis was applied to combine individual GWAS summary statistics.

Results: No genome-wide significant interactions (departures from a log-additive odds model) with diabetes or obesity were detected at the SNP level by the CC or CO approaches. The joint-effect test detected numerous genome-wide significant GxE signals in the GWAS main effects top hit regions, but the significance diminished after adjusting for the GWAS top hits. In the gene-based analysis, a significant interaction of diabetes with variants in the *FAM63A* (family with sequence similarity 63 member A) gene (significance threshold $P < 1.25 \times 10^{-6}$) was observed in the meta-analysis ($P_{GxE} = 1.2 \times 10^{-6}$, $P_{Joint} = 4.2 \times 10^{-7}$).

Conclusions: This analysis did not find significant GxE interactions at the SNP level but found one significant interaction with diabetes at the gene level. A larger sample size might unveil additional genetic factors via GxE scans.

Impact: This study may contribute to discovering the mechanism of diabetes-associated pancreatic cancer.

[1]Department of Gastrointestinal Medical Oncology, The University of Texas MD Anderson Cancer Center, Houston, Texas. [2]Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, Massachusetts. [3]Division of Cancer Epidemiology and Genetics, NCI, NIH, Bethesda, Maryland. [4]Department of Obstetrics and Gynecology, New York University School of Medicine, New York, New York. [5]Department of Population Health, New York University School of Medicine, New York, New York. [6]Department of Environmental Medicine, New York University School of Medicine, New York, New York. [7]Department of Epidemiology and Biostatistics, University of California, San Francisco, San Francisco, California. [8]International Agency for Research on Cancer, Lyon, France. [9]Genomic Epidemiology Group, German Cancer Research Center (DKFZ), Heidelberg, Germany. [10]Department of Epidemiology and Biostatistics, Memorial Sloan Kettering Cancer Center, New York, New York. [11]Lunenfeld-Tanenbaum Research Institute, Sinai Health System and University of Toronto, Toronto, Ontario, Canada. [12]Division of Cancer Epidemiology, Cancer Council Victoria, Melbourne, Victoria, Australia. [13]Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health, The University of Melbourne, Melbourne, Victoria, Australia. [14]Precision Medicine, School of Clinical Sciences at Monash Health, Monash University, Clayton, Victoria, Australia. [15]SWOG Statistical Center, Fred Hutchinson Cancer Research Center, Seattle, Washington. [16]Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, Washington. [17]Cancer Epidemiology Program, University of Hawaii Cancer Center, Honolulu, Hawaii. [18]Population Health Department, QIMR Berghofer Medical Research Institute, Brisbane, Australia. [19]Division

Introduction

Pancreatic cancer is the third leading cause of cancer-related death, accounting for more than 47,000 deaths each year in the United States (1). It is a highly lethal disease with a 5-year survival rate of 9% (2). Epidemiologic studies have shown that 20%–25% of pancreatic cancer cases are attributable to cigarette smoking (3). However, the incidence of pancreatic cancer has been rising slightly each year in the United States since 2002; this is unexpected given the decreasing prevalence of cigarette smoking, and may be due to the rising prevalence of obesity and diabetes. Accumulating evidence suggests that obesity and long-term type II diabetes are associated with increased risk of pancreatic cancer. For example, a pooled analysis of 14 cohort studies of body mass index (BMI) has shown that obesity (BMI ≥30 kg/m^2) was associated with 47% [95% confidence interval (CI), 23%–75%] increased risk of pancreatic cancer (4). A meta-analysis of 23 cohort and case–control (CC) studies suggests that the association between BMI and pancreatic cancer is not linear (5). At least four meta-analyses of large datasets from cohort and CC studies have shown that long-term diabetes was associated with a 1.5- to 2-fold increased risk of pancreatic cancer (6–9). Because only a small portion of obese and diabetic individuals develop pancreatic cancer, understanding how genetic factors affect risk among those individuals could inform targeted interventions or screening. Identifying variants that are only associated with risk of cancer (or have stronger associations) among obese or diabetic individuals is of particular interest.

Genome-wide association studies (GWAS) conducted by the Pancreatic Cancer Cohort Consortium (PanScan) and Pancreatic Cancer Case Control Consortium (PanC4) have identified 21 genetic loci and chromosome regions significantly associated with the risk of pancreatic cancer (10–15). However, these findings explain limited heritability of the disease, that is, the established GWAS loci explain 2.1% of the heritability of pancreatic cancer in contrast to the estimated heritability of 36% from a large population-based twin study (13, 16). Beyond main effects, some genetic factors may contribute to the risk of pancreatic cancer only in the presence of specific risk factors for the disease such as obesity and diabetes, that is, gene–obesity/diabetes interaction, and broadly referred as gene–environment interaction (GxE) herein. Therefore, a genome-wide GxE scan may help find the missing heritability of pancreatic cancer. Several of the susceptibility genes identified by GWAS (*NR5A2, PDX1, HNF1B*, and *HNF4G*) are important for pancreas development (17). These genes are important components of the transcriptional networks governing embryonic pancreatic development and differentiation, as well as maintaining pancreatic homeostasis. Mutations in some of these genes are responsible for maturity onset diabetes of the young and common variants of these genes have been associated with BMI and risk of type 2 diabetes in GWAS (17). Therefore, in addition to their roles in regulating the development and function of the pancreas, these genes may contribute to pancreatic cancer, partially through an increased risk of obesity and diabetes. Whether these genes and other unidentified genes have an interactive action with obesity and diabetes in modifying the risk of pancreatic cancer is the focus of the current investigation.

We have previously performed GxE analyses at SNP/gene/pathway levels using GWAS data from 2,028 cases and 2,109 controls from PanScan I and II. No significant interactions at the SNP or gene levels were observed for diabetes or obesity. At the pathway level, NF-κB–mediated chemokine signaling and axonal guidance signaling pathway, respectively, were identified as the top pathways interacting with obesity and smoking in modifying the risk of pancreatic cancer (18, 19). These studies were limited by the small sample size, and underpowered for genome-wide GxE analysis (20). To address this limitation, we conducted the current analysis in a much larger combined dataset of PanScan I–III and PanC4 with 8,255 cases and 11,900 controls. We further leveraged recently developed, more powerful SNP-set/gene-based GxE tests (21, 22) to discover novel genetic variants that may modify the association between diabetics/obesity and pancreatic cancer.

of Epidemiology, Department of Medicine, Vanderbilt Epidemiology Center, Vanderbilt-Ingram Cancer Center, Vanderbilt University School of Medicine, Nashville, Tennessee. [20]Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland. [21]Department of Oncology, Sidney Kimmel Comprehensive Cancer Center, Johns Hopkins School of Medicine, Baltimore, Maryland. [22]Cancer Prevention Program, Fred Hutchinson Cancer Research Center, Seattle, Washington. [23]Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, Massachusetts. [24]Department of Health Sciences Research, Mayo Clinic College of Medicine, Rochester, Minnesota. [25]Department for Determinants of Chronic Diseases (DCD), National Institute for Public Health and the Environment (RIVM), Bilthoven, the Netherlands. [26]Department of Gastroenterology and Hepatology, University Medical Centre, Utrecht, the Netherlands. [27]Department of Epidemiology and Biostatistics, The School of Public Health, Imperial College London, London, United Kingdom. [28]Department of Social and Preventive Medicine, Faculty of Medicine, University of Malaya, Kuala Lumpur, Malaysia. [29]Division of Preventive Medicine, Brigham and Women's Hospital, Boston, Massachusetts. [30]Department of Biology, University of Pisa, Pisa, Italy. [31]Oncology Data Analytics Program, Catalan Institute of Oncology (ICO), Bellvitge Biomedical Research Institute (IDIBELL), L'Hospitalet de Llobregat, Barcelona, Spain. [32]Yale Cancer Center, New Haven, Connecticut. [33]Department of Medicine, Yale School of Medicine, New Haven, Connecticut. [34]Smilow Cancer Hospital, New Haven, Connecticut. [35]Boston Veteran Affairs Healthcare System, Boston, Massachusetts. [36]Department of Pathology, Sol Goldman Pancreatic Cancer Research Center, Johns Hopkins School of Medicine, Baltimore, Maryland. [37]Gastroenterology, Hepatology, and Nutrition Service, Memorial Sloan Kettering Cancer Center, New York, New York. [38]Genetic and Molecular Epidemiology Group, Spanish National Cancer Research Centre, Madrid, Spain. [39]CIBER Epidemiología y Salud Pública (CIBERESP), Barcelona, Spain. [40]Hospital del Mar Institute of Medical Research (IMIM), Universitat Autònoma de Barcelona, Barcelona, Spain. [41]CHRISTUS Santa Rosa Hospital – Medical Center, San Antonio, Texas. [42]Department of Public Health, University of Copenhagen and Danish Cancer Society Research Center Diet, Genes and Environment, Copenhagen, Denmark. [43]Hellenic Health Foundation, World Health Organization Collaborating Center of Nutrition, Medical School, University of Athens, Athens, Greece. [44]Department of Epidemiology and Environmental Health, University of Buffalo, Buffalo, New York. [45]Department of Chronic Disease Epidemiology, Yale School of Public Health, New Haven, Connecticut. [46]Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland. [47]Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, Massachusetts. [48]Department of Biostatistics, The University of Texas MD Anderson Cancer Center, Houston, Texas.

Note: Supplementary data for this article are available at Cancer Epidemiology, Biomarkers & Prevention Online (http://cebp.aacrjournals.org/).

H. Tang, L. Jiang, D. Li, P. Kraft, and P. Wei contributed equally to this article.

Corresponding Authors: Peng Wei, The University of Texas MD Anderson Cancer Center, 1400 Pressler Street, Unit 1411, Houston, TX 77030. Phone: 713-563-4285; E-mail: pwei2@mdanderson.org; Peter Kraft, Departments of Epidemiology and Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA 02115. E-mail: pkraft@hsph.harvard.edu; and Donghui Li, Department of Gastrointestinal Medical Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX 77030. E-mail: dli@mdanderson.org

Cancer Epidemiol Biomarkers Prev 2020;29:1784–91

doi: 10.1158/1055-9965.EPI-20-0275

Materials and Methods

Study population and datasets

This genome-wide GxE study includes 8,255 cases and 11,900 controls of European ancestry drawn from the PanScan and PanC4 consortia. Cases were patients with known or presumed primary pancreatic ductal adenocarcinoma (ICD-O-3 code C250–C259) and controls were free of pancreatic cancer. Individual studies were approved by the respective institutional review board following the institution's requirement. Written informed consent was obtained from each study participant. The approaches for data harmonization and meta-analysis were approved by the University of Texas MD Anderson Cancer Center Institutional Review Board (Houston, TX).

Genotype data were generated in four previously reported GWASs, that is, PanScan I, II, and III and PanC4, and the details of these studies have been described previously (10–13). Genotyping in PanScan I, II, and III was conducted at the Cancer Genomics Research Laboratory of the NCI of the National Institutes of Health (NIH) using the Illumina HumanHap550 Infinium II, Human 610-Quad, and OmniExpress series arrays, respectively. PanC4 employed the HumanOmniExpress-Exome-8v1 array. Because different genotyping platforms were used in these studies, missing genotypes were imputed using the University of Michigan imputation server (https://imputationserver.sph.umich.edu/index.html) with the Haplotype Reference Consortium (23) as the reference panel or IMPUTE2 with the 1000 Genomes Phase 3 as the reference panel (https://mathgen.stats.ox.ac.uk/impute/impute_v2.html). After imputation, SNPs that were identified by imputation only (not genotyped in any of the four GWASs), having minor allele frequency (MAF) ≤ 0.005, imputation quality score <0.3, or Hardy–Weinberg equilibrium test $P < 1 \times 10^{-6}$ in controls were excluded; a total of about 870,000 common SNPs to all four studies were included in this GxE analysis. The PanScan (I, II, and III) and PanC4 GWAS data are available through dbGaP (accession numbers phs000206.v5.p3 and phs000648.v1.p1, respectively).

Exposure variables

The exposure variables considered in this GxE analysis were obesity (BMI ≥30 kg/m^2 vs. <30 kg/m^2) and diabetes (diabetes with ≥3 years of duration vs. nondiabetes). Because diabetes could be a manifestation of occult pancreatic cancer, we excluded diabetes with a short duration (<3 years) for studies with diabetes duration information to control reverse causality. Covariates for adjustment included age (continuous), sex, study sites, and principal components accounting for population substructure. The distribution of demographics and risk factors of participants in each GWAS included in this analysis are summarized in Supplementary Table S1.

Statistical analyses

We applied CC, case-only (CO), and 2 degrees-of-freedom (2-df) joint-effect test (24) methods at the SNP level, and the "rareGE" method (21) at the gene level in the genome-wide GxE scan. The 2-df joint-effect test is more powerful in detecting a susceptible SNP in the presence of strong genetic main effect (SNP), strong interaction effect (SNPxE), or a combination of weak/moderate main and interaction effects (SNP + SNPxE). Thus, the joint-effect test is a useful complementary approach to CC, CO, and single-SNP marginal association analysis in identifying disease susceptible loci (20).

The PanScan I–III and PanC4 datasets were analyzed individually using the CC, CO, and joint-effect test at the SNP level. The "rareGE" method was used for gene-based GxE analysis. The summary statistics for each consortium were then subjected to meta-analysis.

SNP-level tests

To perform SNP-level analysis, we ran the logistic regression model as follows:

$$\text{Logit}(P(Y=1)) = \beta_0 + \beta_E E + \beta_G \text{SNP} + \beta_{GE}\text{SNP} * E + \beta_C C, \quad \text{(A)}$$

where Y is the disease status (1 for case; 0 for control); β_0 is the intercept; E is the exposure variable of interest (diabetes or obesity); SNP is the dosage of the genetic variant of interest, coded additively accounting for genotype imputation uncertainty (ranging from 0 to 2); and C is the vector of all covariates including age (continuous), sex, study indicators, principal components accounting for population substructures, and either diabetes or BMI [e.g., diabetes serves as the exposure of interest with BMI (continuous) included in the covariate vector]. For the CC study design, the null hypothesis to be tested H_0: $\beta_{GE} = 0$. $e^{\beta_{GE}}$ was referred as the interaction OR.

Joint-effect analysis of SNP and SNPxE were run using the approach by Aschard and colleagues (25) by testing the null hypothesis H_0: $\beta_G = \beta_{GE} = 0$, derived from model (A) with a 2-df χ^2 Wald test. For the CO study design, a logistic regression model was run in the case group only as follows:

$$\text{Logit}(P(E=1)) = \beta_0 + \beta_G \text{SNP} + \beta_C C, \quad \text{(B)}$$

where the coefficients in model (B) are denoted the same as those in model (A).

Gene-level tests

Gene regions were defined according to coordinates of the hg19 assembly, retrieved from the University of California, Santa Cruz (UCSC) Genome Browser (26). About 22,300 genes were downloaded from UCSC server, of which approximately 20,000 genes covering ≥2 GWAS genotyped SNPs were analyzed in this study.

We performed gene-based GxE analysis using the "rareGE" method (21) based on common SNPs (MAF ≥ 0.005, located within 20 kb upstream or downstream of a given gene). For a gene with p SNPs, the full model is as follows:

$$\text{Logit}(P(Y=1)) = \beta_0 + \beta_E E + \sum_{j=1}^{p} \beta_{Gj}\text{SNP}_j + \sum_{j=1}^{p} \beta_{GEj}\text{SNP}_j * \text{E} + \beta_C C, \quad \text{(C)}$$

where β_{Gj} and β_{GEj} are the regression coefficients for the genetic main effect and GxE effect for the jth SNP, respectively.

Two tests were implemented in the "rareGE" R package: GxE test with genetic main effects estimated as random effects (P_{Int}) under the null hypothesis of no GxE, that is, H_0: $\beta_{GE1} = \beta_{GE2} = \ldots = \beta_{GEp} = 0$, and a joint test of G and GxE (P_{Joint}) with H_0: $\beta_{G1} = \beta_{G2} = \ldots = \beta_{Gp} = 0$ and $\beta_{GE1} = \beta_{GE2} = \ldots = \beta_{GEp} = 0$, analogous to the 2-df SNP-level joint-effect test.

Meta-analyses

We applied a fixed-effects meta-analysis in METAL to combine SNP-level GxE results from the CC or CO method across individual consortia (27). Fisher meta-analysis was used to combine gene-level GxE P values from the rareGE method (28).

Statistical thresholds

All tests were two sided. We consider $P < 2.5 \times 10^{-8}$ and $P < 1.25 \times 10^{-6}$ as genome-wide significant at the SNP and gene level, respectively (29), for each individual study and each meta-analysis, adjusted for 1 million SNPs, 20,000 genes, and two exposures of interest by the

Table 1. Top SNPs interacting with diabetes and obesity (CC).

SNP	Chr.	Position	Gene[a]	Effect/ref allele	MAF[b]	Meta-analysis	
						OR (95% CI)	*P*
Diabetes							
rs7505930	18	4092001	*ROCK1P1-SLC35G4	G/A	0.35	1.60 (1.34–1.91)	1.9E-07
rs2777534	10	34109601	*GTPBP4-FGF8	A/G	0.12	2.04 (1.56–2.67)	2.3E-07
rs2812656	10	34116863	*GTPBP4-FGF8	G/A	0.12	0.50 (0.38–0.65)	2.4E-07
rs11086650	20	57183256	*APCDD1L_AS1*	C/T	0.32	0.61 (0.51–0.74)	5.8E-07
Obesity							
rs7802442	7	22736446	*COX19-SLC12A9	C/A	0.31	0.73 (0.65–0.83)	1.2E-06
rs4298423	7	151643909	*PRKAG2_AS1-GALNTL5**	A/G	0.34	1.34 (1.19–1.51)	2.3E-06
rs559449	11	55340379	*OR4C16*	A/G	0.45	1.31 (1.17–1.47)	3.6E-06
rs7608326	2	37903390	*GRHL1-CHST10	C/T	0.07	0.51 (0.38–0.68)	4.2E-06
rs759831	16	82863660	*CDH13*	A/C	0.31	1.32 (1.17–1.49)	5.5E-06
rs1476483	7	22731199	*COX19-SLC12A9	G/A	0.20	0.72 (0.62–0.83)	8.5E-06

Abbreviation: Chr., chromosome.
[a]Gene region was defined by the UCSC Genome Browser; *, the nearest gene to the SNP.
[b]Derived from the PanC4 dataset.

Bonferroni correction at family-wise error rate of 0.05. $P < 5.0 \times 10^{-2}$ was considered as nominally significant for all analyses.

Statistical power estimation

We used the QUANTO software (version 1.2.4; ref. 30) to perform power estimation for these GxE scans. With 8,255 cases and 11,900 controls, we had 80% power to detect an interaction OR of 1.5 and 1.6, respectively, for obesity (main effect OR = 1.2 with 20% prevalence in controls based on Supplementary Table S1) and diabetes (main effect OR = 1.7 with 10% prevalence in controls based on Supplementary Table S1) for an SNP with MAF of 20% at a significance level of 2.5×10^{-8} by the standard CC test.

Results

First, we examined the GxE (obesity and diabetes) interactions at the SNP level using the CC, CO, and joint tests in each individual GWAS, followed by meta-analysis of the summary statistics. Supplementary Fig. S1 shows the quantile–quantile (Q–Q) plots for the CC and CO meta-analyses. There was no discernable abnormal behavior in the Q–Q plots for CC and CO study designs (genomic control λ ranged from 0.942 to 1.023). Q–Q plots also performed well for meta-analysis of joint-effect tests (λs: 0.94–1.045).

CC and CO analyses

No signal at a genome-wide threshold of significance ($P < 2.5 \times 10^{-8}$) was detected in CC or CO analyses on interactions of genes with diabetes or obesity. Using the CC approach, four SNPs on chromosomes 10, 18, and 20 showed evidence of interactions with diabetes at near genome-wide significance ($P < 1 \times 10^{-6}$) and six SNPs on chromosomes 2, 7, 11, and 16 showed weaker evidence of interactions with obesity ($P < 1 \times 10^{-5}$; **Table 1**). By the CO approach, four SNPs on chromosomes 3 and 10 showed evidence of interactions with diabetes at near genome-wide significance ($P < 1 \times 10^{-6}$; **Table 2**). Of these, two SNPs (rs12255372 and rs7901695) were near *TCF7L2* and in linkage disequilibrium ($r^2 = 0.74$ and 0.87, respectively) with the lead SNP from a recent GWAS of type 2 diabetes (rs7903146; $P = 1 \times 10^{-347}$; ref. 31). Thus, the CO signals for these two SNPs likely reflect violations of the gene–environment independence assumption rather than evidence for GxE. In addition, five SNPs on chromosomes 4, 8, 14, and 17 had possible interactions with obesity at $P < 1 \times 10^{-5}$ (**Table 2**). Further, no significant across-consortium heterogeneity was

Table 2. Top interaction signals from CO analyses.

SNP	Chr.	Position	Gene[a]	Effect/ref allele	MAF[b]	P_{CO}
Diabetes						
rs608841	3	138764229	*PRR23C-BPESC1	G/A	0.24	1.6E-07
rs696638	3	138775377	*PRR23C-BPESC1	A/G	0.16	2.2E-07
rs12255372	10	114808902	*TCF7L2*	A/C	0.28	2.3E-07
exm-rs7903146	10	114758349	*TCF7L2*	A/G	0.29	4.9E-07
Obesity						
rs2018572	17	11599798	*BHLHA9-DNAH9**	G/A	0.19	1.3E-07
rs4791473	17	11574959	*BHLHA9-DNAH9**	G/T	0.18	1.9E-06
rs4413478	4	48491651	*SLC10A4-ZAR1**	A/G	0.25	2.8E-06
rs925611	8	9768690	*OR4F21-C8orf49**	T/G	0.097	3.2E-06
rs961044	14	87608094	*LOC283585-GALC	G/A	0.14	6.9E-06

Abbreviations: Chr., chromosome; P_{CO}, CO test *P* value.
[a]Gene region was defined by the UCSC Genome Browser; *, the nearest gene to the SNP.
[b]Derived from the PanC4 dataset.

Table 3. Top genes interacting with diabetes and obesity by rareGE method.

Gene	Chr.	Meta		PanScan I		PanScan II		PanScan III		PanC4	
		P_{Int}	P_{Joint}	P_{Int}	P_{Joint}	P_{Int}	P_{Joint}	P_{Int}	P_{Joint}	P_{Int}	P_{Joint}
Diabetes											
FAM63A	1q21.3	**1.2E-6**[a]	**4.2E-7**[a]	3.8E-2	6.8E-2	0.024	0.04	2.2E-4	8.8E-6	3.3E-3	8.1E-3
CLTCL1	22q11.21	1.5E-4	5.2E-4	0.85	0.98	4.9E-3	0.01	0.77	0.95	6.0E-5	1.0E-4
MIR561	2q32.1	4.1E-5	6.6E-4	9.3E-4	1.7E-3	0.043	8.5E-2	8.1E-3	3.5E-2	0.13	0.25
GNG2	14q22.1	3.4E-5	1.1E-3	0.76	0.66	3.4E-3	7.9E-3	2.5E-5	5.9E-4	0.51	0.76
ADA	20q13.12	6.8E-5	4.6E-4	0.14	0.27	1.8E-3	3.7E-3	0.47	0.6	0.97	0.38
TP53I3	2p23.2	7.0E-5	1.7E-3	0.31	0.52	0.28	0.36	2.0E-6	3.0E-5	0.46	0.71
SF3B14	2p23.3	6.9E-5	1.9E-3	0.31	0.52	0.28	0.36	2.0E-6	3.6E-5	0.45	0.7
DCAF6	1q24.1	2.7E-2	1.6E-5	2.4E-2	0.05	0.49	2.7E-5	0.21	7.3E-2	0.07	0.14
OR6K2	1q23.1	3.3E-6	4.0E-3	6.4E-2	0.12	0.037	7.3E-2	2.0E-6	2.9E-3	0.45	0.48
MIR4457	5p15.33	0.57	9.9E-6	0.14	2.9E-2	0.93	0.61	0.61	5.8E-4	0.44	7.6E-4
Obesity											
CDC42EP3	2p22.2	3.40E-04	2.10E-05	0.18	0.17	0.62	4.3E-3	5.50E-02	1.70E-01	9.10E-05	1.5E-4
FSD1L	9q31.2	6.50E-02	3.60E-05	4.2E-2	0.066	6.1E-2	0.13	8.80E-01	8.70E-06	2.80E-01	0.48
MIR4457	5p15.33	3.10E-01	5.10E-06	0.45	0.0045	0.82	0.32	6.20E-01	6.50E-03	4.00E-02	3.8E-4

Abbreviations: Chr., chromosome; P_{Int} and P_{Joint}, *P* values, respectively, derived from random-effect GxE interaction test and joint-effect test.
[a]Genome-wide significant *P* values (<1.25E-6).

found for the meta-analysis results in **Tables 1** and **2** (all heterogeneity test $P > 0.05$).

2-df joint-effect test

Meta-analysis of joint-effect tests for SNP and SNP × diabetes or SNP × obesity detected numerous genome-wide significant signals that are all located in the chromosome regions containing previously identified GWAS top hits (Supplementary Table S2). Conditional analysis adjusting for the GWAS top hits in each region resulted in null findings, indicating that joint-effect test signals were all driven by the strong main effects of the SNPs.

Gene-level GxE analysis

Possible interactions of nine genes with diabetes and three genes with obesity at a meta-analysis significance level of $P < 1 \times 10^{-4}$ in at least one of the interaction-only and joint tests are listed in **Table 3**. Among these genes, a significant ($P < 1.25 \times 10^{-6}$) interaction of diabetes with *FAM63A* gene was observed in the meta-analysis ($P_{Interaction} = 1.2 \times 10^{-6}$, $P_{Joint} = 4.2 \times 10^{-7}$; **Table 3**). The SNPs contributing to this gene are listed in Supplementary Table S3. No individual SNP of this gene showed a significant interaction with diabetes.

Discussion

In this genome-wide gene–obesity/diabetes interaction study of pancreatic cancer, no significant departures from a log-linear odds model at the SNP level were identified by the CC or CO approaches. In the gene-based analysis, a significant interaction between variants in the *FAM63A* gene and diabetes was observed.

FAM63A, also known as MINDY-1 (MINDY lysine 48 deubiquitinase 1) is a member of an evolutionarily conserved and structurally distinct family of deubiquitinating enzymes (32), which specifically cleaves K48-linked poly-ubiquitin chain to regulate protein degradation. This distinct deubiquitinase class localizes to DNA lesions, where it plays an important role in genome stability pathways, functioning to prevent spontaneous DNA damage and to promote cellular survival in response to exogenous DNA damage (33). Previous GWASs have associated *FAM63A* or *FAM63A* homolog gene variants with the risk of primary rhegmatogenous retinal detachment (34) and chronic renal disease (35). Genetic analysis of a diabetes-prone mouse strain has revealed gene regions homologous to *FAM63A* contributing to diabetes susceptibility (36). Although the role of *FAM63A* in pancreatic cancer is unknown at present, the observed interaction with diabetes deserves further investigation.

Genome-wide GxE analysis has unique challenges compared with genetic main effects analysis in GWAS. First, GxE analysis requires a much larger sample size to detect a realistic interaction OR than does a GWAS scan for a comparable main effect OR (20, 37), largely explaining why few positive findings have been reported in GxE studies (38–40). For example, this GxE scan with 8,255 cases and 11,900 controls, even though about four times as large as our previous gene–obesity/diabetes interaction analysis (18), had 80% power to detect an interaction OR of 1.5 and 1.6, respectively, for obesity and diabetes for an SNP with MAF of 20% at a significance level of 2.5×10^{-8} by the standard CC test; in contrast, the same sample size had 80% power to detect a genetic main effect OR of 1.18 at the same MAF and significance level. To boost the power for a given sample size, novel statistical and analytic methods have been proposed to leverage *a priori* biological knowledge in the form of genes, pathways, or other functional genomic annotations such as those derived from the ENCODE and NIH Epigenomics Roadmap projects (18, 19, 41). Second, exposure variability and measurement accuracy play a considerable role in determining the power and reproducibility of GxE studies (42, 43). Third, there is no single most powerful statistical method for either SNP or gene-level genome-wide GxE analysis due to the largely unknown patterns of GxE interaction signals and combinations of genetic main and GxE effects (20, 22). Therefore, we suggest that the GxE analysis should make use of multiple methods with complementary strengths, as used here and suggested by other investigators (44), to discover the missing heritability of pancreatic cancer (45).

This study identified a statistically significant interaction of diabetes with variants in *FAM63A* in gene-based GxE analysis, but no significant SNP-level GxE interactions with either diabetes or obesity. We note that the absence of interaction on the log-odds scale has

potentially important implications for risk modeling, as it typically implies presence of interaction on the risk difference scale, sometimes referred to as "public health interaction" (46). Developing and validating a multifactorial risk model is beyond the scope of this article, but we note that our results lend support to the common assumption of additive log odds when combining genetic, clinical, and environmental risk factors to predict risk (47, 48).

This study has strengths and limitations. This is by far the largest GxE analysis in pancreatic cancer. Quality control was strictly performed in steps of genotyping, population structure definition, exposure measurement, and harmonization. Diabetes was defined as disease with ≥3-year duration, avoiding reverse causality. Along the same line, because it is common for patients with pancreatic cancer to experience severe weight loss (43), we avoided using body weight at or close to cancer diagnosis for cases when calculating the BMI. Following the state-of-the-art analysis strategies in large consortium-based GxE scans (49, 50), we only adjusted for a "minimum" set of covariates, including age, sex, study sites, and principal components accounting for population substructure, in the regression analysis. As shown by the well-behaved Q–Q plots in Supplementary Fig. S1, there was no indication of uncontrolled confounding effects. Finally, genome-wide significant thresholds based on the Bonferroni correction were applied to reduce false-positive discovery. Nevertheless, relatively small sample sizes curbed the power of the genome-wide GxE scan from CC and CO study designs. Despite this, the current GxE analysis discovered a novel susceptibility locus for pancreatic cancer using a gene-based GxE test, and may contribute to discovering the mechanism of diabetes-associated pancreatic cancer.

Disclosure of Potential Conflicts of Interest

C. Fuchs reports other commercial research support from Agios, Bain Capital, Unum Therapeutics, CytomX Therapeutics, Daiichi Sankyo, Eli Lilly, Entrinsic Health, Evolveimmune Therapeutics, Genentech, Merck, and Taiho; has ownership interest (including patents) in CytomX Therapeutics, Entrinsic Health, and Evolveimmune Therapeutics; and reports other remuneration from Amylin Pharma. K. Ng reports receiving commercial research grants from Celgene and Revolution Medicines. No potential conflicts of interest were disclosed by other authors.

Disclaimer

The authors assume full responsibility for analyses and interpretation of these data. Where authors are identified as personnel of the International Agency for Research on Cancer/World Health Organization, the authors alone are responsible for the views expressed in this article and they do not necessarily represent the decisions, policy, or views of the International Agency for Research on Cancer/World Health Organization.

Authors' Contributions

Conception and design: H. Tang, R.Z. Stolzenberg-Solomon, G.G. Giles, B. Bueno-de-Mesquita, S.J. Chanock, J.M. Gaziano, P. Hartge, M.H. Hassam, E.A. Holly, A. Tjønneland, L.T. Amundadottir, G.M. Petersen, A.P. Klein, D. Li, P. Kraft, P. Wei

Development of methodology: B. Bueno-de-Mesquita, P. Hartge, M.H. Hassam, E.A. Holly, M. Porta, D. Li, P. Kraft, P. Wei

Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): R.Z. Stolzenberg-Solomon, A.A. Arslan, L.E.B. Freeman, P.M. Bracci, P. Brennan, F. Canzian, M. Du, S. Gallinger, G.G. Giles, C. Kooperberg, L.L. Marchand, R.E. Neale, X.-O. Shu, E. White, W. Zheng, D. Albanes, G. Andreotti, S.I. Berndt, B. Bueno-de-Mesquita, J.E. Buring, S.J. Chanock, C. Fuchs, J.M. Gaziano, M. Goggins, M.H. Hassam, E.A. Holly, R.N. Hoover, R.J. Hung, R.C. Kurtz, I.-M. Lee, N. Malats, R.L. Milne, I. Orlow, U. Peters, M. Porta, K.G. Rabe, N. Rothman, G. Scelo, H.D. Sesso, D.T. Silverman, I.M. Thompson Jr, A. Tjønneland, A. Trichopoulou, J. Wactawski-Wende, N. Wentzensen, L.R. Wilkens, H. Yu, A. Zeleniuch-Jacquotte, E.J. Jacobs, G.M. Petersen, B.M. Wolpin, H.A. Risch, A.P. Klein, D. Li

Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): H. Tang, L. Jiang, P. Brennan, E. White, G. Andreotti, A. Babic, A. Blackford, E. Childs, M.H. Hassam, R.N. Hoover, R.L. Milne, K. Ng, L.T. Amundadottir, G.M. Petersen, H.A. Risch, D. Li, P. Kraft, P. Wei

Writing, review, and/or revision of the manuscript: H. Tang, L. Jiang, R.Z. Stolzenberg-Solomon, A.A. Arslan, L.E.B. Freeman, P.M. Bracci, P. Brennan, G.G. Giles, P.J. Goodman, C. Kooperberg, L.L. Marchand, R.E. Neale, X.-O. Shu, K. Visvanathan, E. White, W. Zheng, D. Albanes, G. Andreotti, A. Babic, W.R. Bamlet, S.I. Berndt, A. Blackford, B. Bueno-de-Mesquita, J.E. Buring, D. Campa, S.J. Chanock, E.J. Duell, C. Fuchs, J.M. Gaziano, M. Goggins, P. Hartge, M.H. Hassam, E.A. Holly, R.N. Hoover, R.J. Hung, I.-M. Lee, N. Malats, R.L. Milne, K. Ng, A.L. Oberg, I. Orlow, U. Peters, M. Porta, K.G. Rabe, G. Scelo, H.D. Sesso, D.T. Silverman, I.M. Thompson Jr, A. Tjønneland, A. Trichopoulou, J. Wactawski-Wende, N. Wentzensen, L.R. Wilkens, H. Yu, A. Zeleniuch-Jacquotte, L.T. Amundadottir, E.J. Jacobs, G.M. Petersen, B.M. Wolpin, H.A. Risch, N. Chatterjee, A.P. Klein, D. Li, P. Kraft, P. Wei

Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): H. Tang, A.A. Arslan, L.E.B. Freeman, M. Du, S. Gallinger, G. Andreotti, W.R. Bamlet, A. Blackford, E.J. Duell, C. Fuchs, A.L. Oberg, G. Scelo, H.D. Sesso, I.M. Thompson Jr, N. Wentzensen, A.P. Klein, P. Kraft

Study supervision: M.H. Hassam, E.A. Holly, R.J. Hung, I.-M. Lee, J. Wactawski-Wende, D. Li, P. Kraft, P. Wei

Other (original principal investigator of one of the NIH/NCI-funded studies that is part of the consortium data that were used for this analysis): E.A. Holly

Acknowledgments

This work was supported by the NIH grants R01CA169122 (to P. Wei) and UH2CA191284 (to P. Kraft). The IARC/Central Europe study was supported by NIH grant R03 CA123546-02 and grants from the Ministry of Health of the Czech Republic (NR 9029-4/2006, NR9422-3, NR9998-3, and MH CZ-DRO-MMCI 00209805). The work at Johns Hopkins University was supported by the NCI grants P50CA062924 and R01CA97075. Additional support was provided by the Lustgarten Foundation, and Susan Wojcicki and Dennis Troper and the Sol Goldman Pancreas Cancer Research Center. This work was supported by an NCI grant R01 CA154823. The Mayo Clinic Biospecimen Resource for Pancreas Research study is supported by the Mayo Clinic SPORE in Pancreatic Cancer (P50 CA102701). The MD Anderson Cancer Center study was supported by the NIH grant R01CA98030 and a grant from the Khalifa Bin Zayed Al Nahyan Foundation. The Memorial Sloan Kettering Cancer Center Pancreatic Tumor Registry was supported by P30CA008748, the Geoffrey Beene Foundation, the Arnold and Arlene Goldstein Family Foundation, and the Society of MSKCC. The PACIFIC Study was supported by the NCI grant R01CA102765, and Kaiser Permanente and Group Health Cooperative. The Queensland Pancreatic Cancer Study was supported by a grant from the National Health and Medical Research Council of Australia (NHMRC; grant number 442302). R.E. Neale was supported by an NHMRC Senior Research Fellowship (#1060183). The UCSF pancreas study was supported by NIH-NCI grants (R01CA1009767, R01CA109767-S1, and R0CA059706) and the Joan Rombauer Pancreatic Cancer Fund. The Yale (CT) pancreas cancer study was supported by the NCI grant 5R01CA098870. The cooperation of 30 Connecticut hospitals, including Stamford Hospital, in allowing patient access, is gratefully acknowledged. The Connecticut Pancreas Cancer Study was approved by the State of Connecticut Department of Public Health Human Investigation Committee. Certain data used in that study were obtained from the Connecticut Tumor Registry in the Connecticut Department of Public Health. Studies included in PANDoRA were partly funded by: the Czech Science Foundation (No. P301/12/1734), the Internal Grant Agency of the Czech Ministry of Health (IGA NT 13 263); the Baden-Württemberg State Ministry of Research, Science and Arts (to Prof. H. Brenner), the Heidelberger EPZ-Pancobank (to Prof. M.W. Büchler and team: Prof. T. Hackert, Dr. N. A. Giese, Dr. Ch. Tjaden, E. Soyka, and M. Meinhardt; Heidelberger Stiftung Chirurgie and BMBF grant 01GS08114), the BMBH (to Prof. P. Schirmacher; BMBF grant 01EY1101), the "5 × 1000" voluntary contribution of the Italian Government, the Italian Ministry of Health (RC1203GA57, RC1303GA53, RC1303GA54, and RC1303GA50), the Italian Association for Research on Cancer (to Prof. A. Scarpa; AIRC n. 12182), the Italian Ministry of Research (to Prof. A. Scarpa; FIRB - RBAP10AHJB), the Italian FIMP-Ministry of Health (to Prof. A. Scarpa; 12 CUP_J33G13000210001), and by the National Institute for Health Research Liverpool Pancreas Biomedical Research Unit, United Kingdom. We would like to acknowledge the contribution of Dr Frederike Dijk and Prof. Oliver Busch (Academic Medical Center, Amsterdam, the Netherlands). The American Cancer Society funds the creation, maintenance, and updating of the Cancer Prevention Study II cohort. Cancer incidence data for CLUE were provided by the Maryland Cancer Registry, Center for Cancer Surveillance and Control, Department of Health and Mental Hygiene (Baltimore, MD; http://phpa.dhmh.maryland.gov/cancer, 410-767-4055). We acknowledge the State of Maryland, the Maryland Cigarette Restitution Fund, and

the National Program of Cancer Registries of the Centers for Disease Control and Prevention for the funds that support the collection and availability of the cancer registry data. We thank all the CLUE participants. The Melbourne Collaborative Cohort Study (MCCS) recruitment was funded by VicHealth and Cancer Council Victoria. The MCCS was further supported by Australian NHMRC grants 209057 and 396414 and by infrastructure provided by Cancer Council Victoria. Cases and their vital status were ascertained through the Victorian Cancer Registry and the Australian Institute of Health and Welfare, including the National Death Index and the Australian Cancer Database. The NYU study (to A Zeleniuch-Jacquotte and A.A. Arslan) was funded by NIH R01 CA098661, UM1 CA182934, and center grants P30 CA016087 and P30 ES000260. The PANKRAS II Study in Spain was supported by research grants from Instituto de Salud Carlos III-FEDER, Spain; Fondo de Investigaciones Sanitarias (FIS; #PI13/00082 and #PI15/01573) and Red Temática de Investigación Cooperativa en Cáncer, Spain (#RD12/0036/0050); European Cooperation in Science and Technology (COST Action #BM1204: EU_Pancreas), Ministerio de Ciencia y Tecnología (CICYT SAF 2000-0097), Fondo de Investigación Sanitaria (95/0017), Madrid, Spain; Generalitat de Catalunya (CIRIT - SGR); and "Red temática de investigación cooperativa de centros en Cáncer" (C03/10), "Red temática de investigación cooperativa de centros en Epidemiología y salud pública" (C03/09), and CIBER de Epidemiología (CIBERESP), Madrid, Spain. The Physicians' Health Study was supported by research grants CA-097193, CA-34944, CA-40360, HL-26490, and HL-34595 from the NIH (Bethesda, MD). The Women's Health Study was supported by research grants CA-047988, HL-043851, HL-080467, and HL-099355 from the NIH (Bethesda, MD). Health Professionals Follow-up Study was supported by NIH grant UM1 CA167552 from the NCI (Bethesda, MD). Nurses' Health Study was supported by NIH grants UM1 CA186107, P01 CA87969, and R01 CA49449 from the NCI (Bethesda, MD). Additional support was provided from the Hale Center for Pancreatic Cancer Research, U01 CA21017 from the NCI (Bethesda, MD), and the United States Department of Defense CA130288, Lustgarten Foundation, Pancreatic Cancer Action Network, Noble Effort Fund, Peter R. Leavitt Family Fund, Wexler Family Fund, and Promises for Purple (to B.M. Wolpin). The WHI program was funded by the National Heart, Lung, and Blood Institute, NIH, U.S. Department of Health and Human Services through contracts HHSN268201600018C, HHSN268201600001C, HHSN268201600002C, HHSN268201600003C, and HHSN268201600004C. The authors thank the WHI investigators and staff for their dedication, and the study participants for making the program possible. A full listing of WHI investigators can be found at: http://www.whi.org/researchers/Documents%20%20Write%20a%20Paper/WHI%20Investigator%20Long%20List.pdf.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received February 24, 2020; revised May 6, 2020; accepted June 9, 2020; published first June 16, 2020.

References

1. American Cancer Society. Cancer facts and figures 2019. Atlanta (GA): American Cancer Society; 2019.
2. Ward EM, Sherman RL, Henley SJ, Jemal A, Siegel DA, Feuer EJ, et al. Annual report to the nation on the status of cancer, featuring cancer in men and women age 20–49 years. J Natl Cancer Inst 2019;111:1279–97.
3. Bosetti C, Lucenteforte E, Silverman DT, Petersen G, Bracci PM, Ji BT, et al. Cigarette smoking and pancreatic cancer: an analysis from the International Pancreatic Cancer Case-Control Consortium (Panc4). Ann Oncol 2012;23:1880–8.
4. Genkinger JM, Spiegelman D, Anderson KE, Bernstein L, van den Brandt PA, Calle EE, et al. A pooled analysis of 14 cohort studies of anthropometric factors and pancreatic cancer risk. Int J Cancer 2011;129:1708–17.
5. Aune D, Greenwood DC, Chan DSM, Vieira R, Vieira AR, Navarro Rosenblatt DA, et al. Body mass index, abdominal fatness and pancreatic cancer risk: a systematic review and non-linear dose-response meta-analysis of prospective studies. Ann Oncol 2012;23:843–52.
6. Everhart J, Wright D. Diabetes mellitus as a risk factor for pancreatic cancer. A meta-analysis. JAMA 1995;273:1605–9.
7. Huxley R, Ansary-Moghaddam A, Berrington de Gonzalez A, Barzi F, Woodward M. Type-II diabetes and pancreatic cancer: a meta-analysis of 36 studies. Br J Cancer 2005;92:2076–83.
8. Elena JW, Steplowski E, Yu K, Hartge P, Tobias GS, Brotzman MJ, et al. Diabetes and risk of pancreatic cancer: a pooled analysis from the pancreatic cancer cohort consortium. Cancer Causes Control 2013;24:13–25.
9. Bosetti C, Rosato V, Li D, Silverman D, Petersen GM, Bracci PM, et al. Diabetes, antidiabetic medications, and pancreatic cancer risk: an analysis from the International Pancreatic Cancer Case-Control Consortium. Ann Oncol 2014;25:2065–72.
10. Amundadottir L, Kraft P, Stolzenberg-Solomon RZ, Fuchs CS, Petersen GM, Arslan AA, et al. Genome-wide association study identifies variants in the ABO locus associated with susceptibility to pancreatic cancer. Nat Genet 2009;41:986–90.
11. Petersen GM, Amundadottir L, Fuchs CS, Kraft P, Stolzenberg-Solomon RZ, Jacobs KB, et al. A genome-wide association study identifies pancreatic cancer susceptibility loci on chromosomes 13q22.1, 1q32.1 and 5p15.33. Nat Genet 2010;42:224–8.
12. Wolpin BM, Rizzato C, Kraft P, Kooperberg C, Petersen GM, Wang Z, et al. Genome-wide association study identifies multiple susceptibility loci for pancreatic cancer. Nat Genet 2014;46:994–1000.
13. Childs EJ, Mocci E, Campa D, Bracci PM, Gallinger S, Goggins M, et al. Common variation at 2p13.3, 3q29, 7p13 and 17q25.1 associated with susceptibility to pancreatic cancer. Nat Genet 2015;47:911–6.
14. Zhang M, Wang Z, Obazee O, Jia J, Childs EJ, Hoskins J, et al. Three new pancreatic cancer susceptibility signals identified on chromosomes 1q32.1, 5p15.33 and 8q24.21. Oncotarget 2016;7:66328–43.
15. Klein AP, Wolpin BM, Risch HA, Stolzenberg-Solomon RZ, Mocci E, Zhang M, et al. Genome-wide meta-analysis identifies five new susceptibility loci for pancreatic cancer. Nat Commun 2018;9:556.
16. Chen F, Childs EJ, Mocci E, Bracci P, Gallinger S, Li D, et al. Analysis of heritability and genetic architecture of pancreatic cancer: a PanC4 study. Cancer Epidemiol Biomarkers Prev 2019;28:1238–45.
17. Li D, Duell EJ, Yu K, Risch HA, Olson SH, Kooperberg C, et al. Pathway analysis of genome-wide association study data highlights pancreatic development genes as susceptibility factors for pancreatic cancer. Carcinogenesis 2012;33:1384–90.
18. Tang H, Wei P, Duell EJ, Risch HA, Olson SH, Bueno-de-Mesquita HB, et al. Genes-environment interactions in obesity- and diabetes-associated pancreatic cancer: a GWAS data analysis. Cancer Epidemiol Biomarkers Prev 2014;23:98–106.
19. Tang H, Wei P, Duell EJ, Risch HA, Olson SH, Bueno-de-Mesquita HB, et al. Axonal guidance signaling pathway interacting with smoking in modifying the risk of pancreatic cancer: a gene- and pathway-based interaction analysis of GWAS data. Carcinogenesis 2014;35:1039–45.
20. Gauderman WJ, Mukherjee B, Aschard H, Hsu L, Lewinger JP, Patel CJ, et al. Update on the state of the science for analytical methods for gene-environment interactions. Am J Epidemiol 2017;186:762–70.
21. Chen H, Meigs JB, Dupuis J. Incorporating gene-environment interaction in testing for association with rare genetic variants. Hum Hered 2014;78:81–90.
22. Yang T, Chen H, Tang H, Li D, Wei P. A powerful and data-adaptive test for rare-variant-based gene-environment interaction analysis. Stat Med 2019;38:1230–44.
23. Das S, Forer L, Schonherr S, Sidore C, Locke AE, Kwong A, et al. Next-generation genotype imputation service and methods. Nat Genet 2016;48:1284–7.
24. Kraft P, Yen YC, Stram DO, Morrison J, Gauderman WJ. Exploiting gene-environment interaction to detect genetic associations. Hum Hered 2007;63:111–9.
25. Aschard H, Hancock DB, London SJ, Kraft P. Genome-wide meta-analysis of joint tests for genetic and gene-environment interaction effects. Hum Hered 2010;70:292–300.
26. Petrovski S, Wang Q, Heinzen EL, Allen AS, Goldstein DB. Genic intolerance to functional variation and the interpretation of personal genomes. PLoS Genet 2013;9:e1003709.
27. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics 2010;26:2190–1.
28. Hedges LV, Olkin I. The inverse chi-square method. In: Statistical methods for meta-analysis. San Diego (CA): Academic Press; 1985. p. 37–9.
29. Barsh GS, Copenhaver GP, Gibson G, Williams SM. Guidelines for genome-wide association studies. PLoS Genet 2012;8:e1002812.
30. Gauderman WJ. Sample size requirements for matched case-control studies of gene-environment interaction. Stat Med 2002;21:35–50.

31. Xue A, Wu Y, Zhu Z, Zhang F, Kemper KE, Zheng Z, et al. Genome-wide association analyses identify 143 risk variants and putative regulatory mechanisms for type 2 diabetes. Nat Commun 2018;9:2941.
32. Abdul Rehman SA, Kristariyanto YA, Choi SY, Nkosi PJ, Weidlich S, Labib K, et al. MINDY-1 Is a member of an evolutionarily conserved and structurally distinct new family of deubiquitinating enzymes. Mol Cell 2016;63:146–55.
33. Kwasna D, Abdul Rehman SA, Natarajan J, Matthews S, Madden R, De Cesare V, et al. Discovery and characterization of ZUFSP/ZUP1, a distinct deubiquitinase class important for genome stability. Mol Cell 2018;70:150–64.
34. Kirin M, Chandra A, Charteris DG, Hayward C, Campbell S, Celap I, et al. Genome-wide association study identifies genetic risk underlying primary rhegmatogenous retinal detachment. Hum Mol Genet 2013;22:3174–85.
35. Kottgen A, Pattaro C, Boger CA, Fuchsberger C, Olden M, Glazer NL, et al. New loci associated with kidney function and chronic kidney disease. Nat Genet 2010; 42:376–84.
36. Mao HZ, Roussos ET, Peterfy M. Genetic analysis of the diabetes-prone C57BLKS/J mouse strain reveals genetic contribution from multiple strains. Biochim Biophys Acta 2006;1762:440–6.
37. McAllister K, Mechanic LE, Amos C, Aschard H, Blair IA, Chatterjee N, et al. Current challenges and new opportunities for gene-environment interaction studies of complex diseases. Am J Epidemiol 2017;186:753–61.
38. Gong J, Hutter CM, Newcomb PA, Ulrich CM, Bien SA, Campbell PT, et al. Genome-wide interaction analyses between genetic variants and alcohol consumption and smoking for risk of colorectal cancer. PLoS Genet 2016;12:e1006296.
39. Kantor ED, Hutter CM, Minnier J, Berndt SI, Brenner H, Caan BJ, et al. Gene-environment interaction involving recently identified colorectal cancer susceptibility loci. Cancer Epidemiol Biomarkers Prev 2014;23:1824–33.
40. Ritz BR, Chatterjee N, Garcia-Closas M, Gauderman WJ, Pierce BL, Kraft P, et al. Lessons learned from past gene-environment interaction successes. Am J Epidemiol 2017;186:778–86.
41. Ritchie MD, Davis JR, Aschard H, Battle A, Conti D, Du M, et al. Incorporation of biological knowledge into the study of gene-environment interactions. Am J Epidemiol 2017;186:771–7.
42. Kraft P, Aschard H. Finding the missing gene-environment interactions. Eur J Epidemiol 2015;30:353–5.
43. Wei P, Tang H, Li D. Functional logistic regression approach to detecting gene by longitudinal environmental exposure interaction in a case-control study. Genet Epidemiol 2014;38:638–51.
44. Hsu L, Jiao S, Dai JY, Hutter C, Peters U, Kooperberg C. Powerful cocktail methods for detecting genome-wide gene-environment interaction. Genet Epidemiol 2012;36:183–94.
45. Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, et al. Finding the missing heritability of complex diseases. Nature 2009;461: 747–53.
46. Siemiatycki J, Thomas DC. Biological models and statistical interactions: an example from multistage carcinogenesis. Int J Epidemiol 1981;10:383–7.
47. Kim J, Yuan C, Babic A, Bao Y, Clish C, Pollak MN, et al. Genetic and circulating biomarker data improve risk prediction for pancreatic cancer in the general population. Cancer Epidemiol Biomarkers Prev 2020;29:99–1008.
48. Klein AP, Lindstrom S, Mendelsohn JB, Steplowski E, Arslan AA, Bueno-de-Mesquita HB, et al. An absolute risk model to identify individuals at elevated risk for pancreatic cancer in the general population. PLoS One 2013;8: e72311.
49. Rao DC, Sung YJ, Winkler TW, Schwander K, Borecki I, Cupples LA, et al. Multiancestry study of gene-lifestyle interactions for cardiovascular traits in 610 475 individuals from 124 cohorts design and rationale. Circ Cardiovasc Genet 2017;10:e001649.
50. Nan HM, Hutter CM, Lin Y, Jacobs EJ, Ulrich CM, White E, et al. Association of aspirin and NSAID use with risk of colorectal cancer according to genetic variants. JAMA 2015;313:1133–42.

Genome-Wide Gene–Diabetes and Gene–Obesity Interaction Scan in 8,255 Cases and 11,900 Controls from PanScan and PanC4 Consortia

Hongwei Tang, Lai Jiang, Rachael Z. Stolzenberg-Solomon, et al.

Cancer Epidemiol Biomarkers Prev 2020;29:1784-1791. Published OnlineFirst June 16, 2020.