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/m²) 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 hits 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 $P < 1.25 \times 10^{-8}$) was observed in the meta-analysis ($P_{GxE} = 1.2 \times 10^{-8}$, $P_{dav} = 4.2 \times 10^{-7}$).

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.
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/m²) 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 (PanCa4) 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–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 (NRP5A2, 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-kB-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 PanCa4 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 diabetes/obesity and pancreatic cancer.

of Epidemiology, Department of Medicine, Vanderbilt Epidemiology Center, Vanderbilt-Ingram Cancer Center, Vanderbilt University School of Medicine, Nashville, Tennessee. 2Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland. 2Department of Oncology, Sidney Kimmel Comprehensive Cancer Center, Johns Hopkins School of Medicine, Baltimore, Maryland. 2Cancer Prevention Program, Fred Hutchinson Cancer Research Center, Seattle, Washington. 2Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, Massachusetts. 2Department of Health Sciences Research, Mayo Clinic College of Medicine, Rochester, Minnesota. 2Department for Determinants of Chronic Diseases (DCD), National Institute for Public Health and the Environment (RIVM), Bilthoven, the Netherlands. 2Department of Gastroenterology and Hepatology, University Medical Centre, Utrecht, the Netherlands. 2Department of Epidemiology and Biostatistics, The School of Public Health, Imperial College London, London, United Kingdom. 2Department of Social and Preventive Medicine, Faculty of Medicine, University of Malaya, Kuala Lumpur, Malaysia. 2Division of Preventive Medicine, Brigham and Women’s Hospital, Boston, Massachusetts. 2Department of Biology, University of Pisa, Pisa, Italy. 2Oncology Data Analytics Program, Catalan Institute of Oncology (ICO), Bellvitge Biomedical Research Institute (IDIBELL), L’Hospitale de Llobregat, Barcelona, Spain. 2Yale Cancer Center, New Haven, Connecticut. 2Department of Medicine, Yale School of Medicine, New Haven, Connecticut. 2Smilow Cancer Hospital, New Haven, Connecticut. 2Boston Veteran Affairs Healthcare System, Boston, Massachusetts. 2Department of Pathology, Sol Goldman Pancreatic Cancer Research Center, Johns Hopkins School of Medicine, Baltimore, Maryland. 2Gastroenterology, Hepatology, and Nutrition Service, Memorial Sloan Kettering Cancer Center, New York, New York. 2Genetic and Molecular Epidemiology Group, Spanish National Cancer Research Centre, Madrid, Spain. 2CIBER Epidemiología y Salud Pública (CIBER-ESP), Barcelona, Spain. 2Hospital del Mar Institute of Medical Research (IMIM), Universitat Autònoma de Barcelona, Barcelona, Spain. 2CHRISTUS Santa Rosa Hospital – Medical Center, San Antonio, Texas. 2Department of Public Health, University of Copenhagen and Danish Cancer Society Research Center Diet, Genes and Environment, Copenhagen, Denmark. 2Hellenic Health Foundation, World Health Organization Collaborating Center of Nutrition, Medical School, University of Athens, Athens, Greece. 2Department of Epidemiology and Environmental Health, University of Buffalo, Buffalo, New York. 2Department of Chronic Disease Epidemiology, Yale School of Public Health, New Haven, Connecticut. 2Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland. 2Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, Massachusetts. 2Department 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, The University of Texas MD Anderson Cancer Center, 1400 Pressler Street, Unit 1411, Houston, TX 77030. Phone: 713-563-4285; E-mail: pkraft1@hsp.harvard.edu; and Donghui Li, Department of Gastrointestinal Medical Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX 77030. E-mail: dlii@mdanderson.org

Cancer Epidemiol Biomarkers Prev 2020;29:1784–91

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

©2020 American Association for Cancer Research.

https://bloodcancerdiscover.aacrjournals.org at University of Washington on September 24, 2020. Copyright ©2020 American Association for Cancer Research.
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). 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 HumanOmnimExpress-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/m² vs. <30 kg/m²) 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 is 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:

\[
\logit(P(Y = 1)) = \beta_0 + \beta_1 E + \beta_2 \text{SNP} + \beta_3 \text{SNP} \times E + \beta_4 C, \quad (A)
\]

where \( Y \) is the disease status (1 for case; 0 for control); \( \beta_0 \) 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 \( \beta_{GE} = 0 \). \( e^{\beta_{GE}} \) 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 \( \beta_{GE} = \beta_{G} = \beta_{E} = 0 \), derived from model (A) with a 2-df \( \chi^2 \) Wald test. For the CO study design, a logistic regression model was run in the case group only as follows:

\[
\logit(P(Y = 1)) = \beta_0 + \beta_1 \text{SNP} + \beta_2 C, \quad (B)
\]

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:

\[
\logit(P(Y = 1)) = \beta_0 + \sum_{j=1}^{p} \beta_{0j} \text{SNP} + \sum_{j=1}^{p} \beta_{Gj} \text{SNP} \times E + \beta_{Ej} C, \quad (C)
\]

where \( \beta_{0j} \) and \( \beta_{Gj} \) are the regression coefficients for the genetic main effect and GxE effect for the \( j \)th SNP, respectively.

Two tests were implemented in the “rareGE” R package: GxE test with genetic main effects estimated as random effects (\( P_{(rand)} \) 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 (\( P_{(jnt)} \)) 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 \( P < 1.25 \times 10^{-8} \) 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 Bonferroni correction. 

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7645945/

CANCER EPIDEMIOLOGY, MARKERS & PREVENTION
Table 1. Top SNPs interacting with diabetes and obesity (CC).

<table>
<thead>
<tr>
<th>SNP</th>
<th>Chr.</th>
<th>Position</th>
<th>Genea</th>
<th>Effect/ref allele</th>
<th>MAFb</th>
<th>OR (95% CI)</th>
<th>Meta-analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diabetes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs7505830</td>
<td>18</td>
<td>4092001</td>
<td>*ROCK1P1-SLC35G4</td>
<td>G/A</td>
<td>0.35</td>
<td>1.60 (1.34-1.91)</td>
<td>1.9E-07</td>
</tr>
<tr>
<td>rs2777534</td>
<td>10</td>
<td>34109601</td>
<td>*GTBP2-4-FGF8</td>
<td>A/G</td>
<td>0.12</td>
<td>2.04 (1.56-2.67)</td>
<td>2.3E-07</td>
</tr>
<tr>
<td>rs2612656</td>
<td>10</td>
<td>34116863</td>
<td>*GTBP4-FGF8</td>
<td>G/A</td>
<td>0.12</td>
<td>0.50 (0.38-0.65)</td>
<td>2.4E-07</td>
</tr>
<tr>
<td>rs1008650</td>
<td>20</td>
<td>57183256</td>
<td>APCCDL_ASI</td>
<td>C/T</td>
<td>0.32</td>
<td>0.61 (0.51-0.74)</td>
<td>5.8E-07</td>
</tr>
</tbody>
</table>

| Obesity    |      |          |       |                   |      |             |              |
| rs7802442  | 7    | 22736466 | *COX19-SLC21A9  | C/A   | 0.31 | 0.73 (0.65-0.83) | 1.2E-06       |
| rs4298423  | 7    | 151643909| PRKAG2_ASI-GALNT5 | A/G  | 0.34 | 1.34 (1.19-1.51) | 2.3E-06       |
| rs559449   | 11   | 53340379 | 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-SLC21A9  | G/A   | 0.20 | 0.72 (0.62-0.83) | 8.5E-06       |

Abbreviation: Chr., chromosome.
aGene region was defined by the UCSC Genome Browser; * , the nearest gene to the SNP.
bDerived 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 \( \times 10^{-8} \) 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 lambda ranged from 0.942 to 1.023). Q–Q plots also performed well for meta-analysis of joint-effect tests (lambda: 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 observed.

Table 2. Top interaction signals from CO analyses.

<table>
<thead>
<tr>
<th>SNP</th>
<th>Chr.</th>
<th>Position</th>
<th>Genea</th>
<th>Effect/ref allele</th>
<th>MAFb</th>
<th>PCO</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diabetes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs608841</td>
<td>3</td>
<td>138764229</td>
<td>*PRR23C-BPESC1</td>
<td>G/A</td>
<td>0.24</td>
<td>1.6E-07</td>
</tr>
<tr>
<td>rs6966338</td>
<td>3</td>
<td>13875337</td>
<td>*PRR23-C-BPESC1</td>
<td>A/G</td>
<td>0.16</td>
<td>2.2E-07</td>
</tr>
<tr>
<td>rs12255372</td>
<td>10</td>
<td>114808902</td>
<td>TCF7L2</td>
<td>A/C</td>
<td>0.28</td>
<td>2.3E-07</td>
</tr>
<tr>
<td>exm-rs7903146</td>
<td>10</td>
<td>11475839</td>
<td>TCF7L2</td>
<td>A/G</td>
<td>0.29</td>
<td>4.9E-07</td>
</tr>
<tr>
<td>Obesity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs2018572</td>
<td>17</td>
<td>11599798</td>
<td>BHLHA9-DNAX9</td>
<td>G/A</td>
<td>0.19</td>
<td>1.3E-07</td>
</tr>
<tr>
<td>rs4794173</td>
<td>17</td>
<td>11574959</td>
<td>BHLHA9-DNAX9</td>
<td>G/T</td>
<td>0.18</td>
<td>1.9E-06</td>
</tr>
<tr>
<td>rs4413478</td>
<td>4</td>
<td>48491651</td>
<td>SLC1A4-ZAR1</td>
<td>A/G</td>
<td>0.25</td>
<td>2.8E-06</td>
</tr>
<tr>
<td>rs925611</td>
<td>8</td>
<td>9768690</td>
<td>OR4F21-CBOF49</td>
<td>T/G</td>
<td>0.097</td>
<td>3.2E-06</td>
</tr>
<tr>
<td>rs961044</td>
<td>14</td>
<td>87608094</td>
<td>*LOC28558S-LALC</td>
<td>G/A</td>
<td>0.14</td>
<td>6.9E-06</td>
</tr>
</tbody>
</table>

Abbreviations: Chr., chromosome; PCO, CO test \( P \) value.
aGene region was defined by the UCSC Genome Browser; * , the nearest gene to the SNP.
bDerived from the PanC4 dataset.
Table 3. Top genes interacting with diabetes and obesity by rareGE method.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Chr.</th>
<th>( P_{\text{meta}} )</th>
<th>( P_{\text{point}} )</th>
<th>( P_{\text{Int}} )</th>
<th>( P_{\text{Point}} )</th>
<th>( P_{\text{Int}} )</th>
<th>( P_{\text{Point}} )</th>
<th>( P_{\text{Int}} )</th>
<th>( P_{\text{Point}} )</th>
<th>( P_{\text{Int}} )</th>
<th>( P_{\text{Point}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diabetes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FAM63A</td>
<td>1q21.3</td>
<td>1.2E-6*</td>
<td>4.2E-7*</td>
<td>3.8E-2</td>
<td>6.8E-2</td>
<td>0.024</td>
<td>0.04</td>
<td>2.2E-4</td>
<td>8.8E-6</td>
<td>3.3E-3</td>
<td>8.1E-3</td>
</tr>
<tr>
<td>CLTCL1</td>
<td>22q12.21</td>
<td>1.5E-4</td>
<td>5.2E-4</td>
<td>0.85</td>
<td>0.98</td>
<td>4.9E-3</td>
<td>0.01</td>
<td>0.77</td>
<td>0.95</td>
<td>6.0E-5</td>
<td>1.0E-4</td>
</tr>
<tr>
<td>MIR561</td>
<td>2q32.1</td>
<td>4.1E-5</td>
<td>6.6E-4</td>
<td>9.3E-4</td>
<td>1.7E-3</td>
<td>0.043</td>
<td>8.5E-2</td>
<td>8.1E-3</td>
<td>3.5E-2</td>
<td>0.13</td>
<td>0.25</td>
</tr>
<tr>
<td>GNG2</td>
<td>1q22.1</td>
<td>3.4E-5</td>
<td>1.1E-3</td>
<td>0.76</td>
<td>0.66</td>
<td>3.4E-3</td>
<td>7.9E-3</td>
<td>2.5E-5</td>
<td>5.9E-4</td>
<td>0.51</td>
<td>0.76</td>
</tr>
<tr>
<td>ADA</td>
<td>20q13.12</td>
<td>6.8E-5</td>
<td>4.6E-4</td>
<td>0.14</td>
<td>0.27</td>
<td>1.8E-3</td>
<td>3.7E-3</td>
<td>0.47</td>
<td>0.6</td>
<td>0.97</td>
<td>0.38</td>
</tr>
<tr>
<td>TP53I3</td>
<td>2p23.2</td>
<td>7.0E-5</td>
<td>1.7E-3</td>
<td>0.31</td>
<td>0.52</td>
<td>0.28</td>
<td>0.36</td>
<td>2.0E-6</td>
<td>3.0E-5</td>
<td>0.46</td>
<td>0.71</td>
</tr>
<tr>
<td>SF3B14</td>
<td>2p23.3</td>
<td>6.9E-5</td>
<td>1.9E-3</td>
<td>0.31</td>
<td>0.52</td>
<td>0.28</td>
<td>0.36</td>
<td>2.0E-6</td>
<td>3.6E-5</td>
<td>0.45</td>
<td>0.7</td>
</tr>
<tr>
<td>DCACF6</td>
<td>1q24.1</td>
<td>2.7E-2</td>
<td>1.6E-5</td>
<td>2.4E-2</td>
<td>0.05</td>
<td>0.49</td>
<td>2.7E-5</td>
<td>0.21</td>
<td>7.3E-2</td>
<td>0.07</td>
<td>0.14</td>
</tr>
<tr>
<td>OR6K2</td>
<td>1q23.1</td>
<td>3.3E-6</td>
<td>4.0E-3</td>
<td>6.4E-2</td>
<td>0.12</td>
<td>0.037</td>
<td>7.3E-2</td>
<td>2.0E-6</td>
<td>2.9E-3</td>
<td>0.45</td>
<td>0.48</td>
</tr>
<tr>
<td>MIR4457</td>
<td>5p15.33</td>
<td>0.57</td>
<td>9.9E-6</td>
<td>0.14</td>
<td>2.9E-2</td>
<td>0.93</td>
<td>0.61</td>
<td>0.61</td>
<td>0.61</td>
<td>0.44</td>
<td>0.76</td>
</tr>
<tr>
<td>Obesity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CDC42EP3</td>
<td>2p22.2</td>
<td>3.4E-04</td>
<td>2.10E-05</td>
<td>0.18</td>
<td>0.17</td>
<td>0.62</td>
<td>4.3E-3</td>
<td>5.5E-02</td>
<td>1.70E-01</td>
<td>9.0E-05</td>
<td>1.5E-4</td>
</tr>
<tr>
<td>PSD2L</td>
<td>9q31.2</td>
<td>6.50E-02</td>
<td>3.60E-05</td>
<td>4.2E-2</td>
<td>0.066</td>
<td>6.1E-2</td>
<td>0.13</td>
<td>8.0E-01</td>
<td>8.70E-06</td>
<td>2.80E-01</td>
<td>0.48</td>
</tr>
<tr>
<td>MIR4457</td>
<td>5p15.33</td>
<td>3.10E-01</td>
<td>5.10E-06</td>
<td>0.45</td>
<td>0.0045</td>
<td>0.82</td>
<td>0.32</td>
<td>6.20E-03</td>
<td>6.50E-03</td>
<td>4.00E-02</td>
<td>3.8E-4</td>
</tr>
</tbody>
</table>

Abbreviations: Chr, chromosome; \( P_{\text{meta}} \) and \( P_{\text{point}} \) \( P \) values, respectively, derived from random-effect GxE interaction test and joint-effect test.

\(^*\)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 \( \times \) diabetes or SNP \( \times \) 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 \( P < 1 \times 10^{-6} \) 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 (\( P_{\text{interaction}} = 1.2 \times 10^{-6} \), \( P_{\text{point}} = 4.2 \times 10^{-7} \); 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-obsesity/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 response in reoxygenation of DNA damage (33). Previous GWASs have associated FAM63A or FAM63A homolog gene variants with the risk of primary rheumatogenous 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-obsesity/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 \( \times 10^{-8} \) 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). 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 novel GxE signals.
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, CytoX Therapeutics, Daichi Sankyo, Eli Lilly, Entirunic Health, Evolveimmune Therapeutics, Genentech, Merck, and Taiho; has ownership interest (including patents) in CytoX Therapeutics, Entirunic 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
Development of methodology: B. Bueno-de-Mesquita, P. Hartge, M.H. Hassam, E.A. Holly, M. Porta, D. Li, P. Kraft, P. Wei
Analysis and interpretation of data (e.g., statistical analysis, bioinformatics, 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. Rusch, D. Li, P. Kraft, P. Wei
Study supervision: M.H. Hassam, E.A. Holly, R.J. Hung, I.-M. Lee, J.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 LARC/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 02009005). The work at Johns Hopkins University was supported by the NCI grants P50CA062924 and R01CA89707. Additional support was provided by the Lustgarten Foundation, and Susan Wojcicki and Dennis Troser and the Sol Goldman Pancreas Cancer Research Center. This work was supported by an NCI grant R01 CA54823. The Mayo Clinic Biospecimen Resource for Research Pancarecancer study is supported by the Mayo Clinic SPOR in Pancreatic Cancer (P50 CA102701). The MD Anderson Cancer Center study was supported by the NIH grant RO1CA89061, a grant from the Khalifa Bin Zayed Al Nahyan Foundation. The Memorial Sloan Kettering Cancer Center Pancreatic Cancer Registry was supported by P30CA087487. The Geofrey Beene Foundation, the Arnold and Arlene Goldstein Family Foundation, and the Society of MSKCC. The PACIFIC Study was supported by the NCI grant CA102336. The PACIFIC study was supported by the NCI grant R01CA169076, and the Memorial Sloan Kettering Cancer Center 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 UCSFS pancreas study was supported by NIH-NCI grants (R01CA159767, R01CA159767, R01CA159767, and RO1CA055906) and R01CA059706). The UCSFS pancreas study was supported by the National Institutes of Health (NIH) and the American Cancer Society.

Gene–Obesity/Diabetes Interactions in Pancreatic Cancer


http://bloodcancerdiscoveries.aacrjournals.org/content/early/2020/09/24/bcd.2020.0178

1134-23071-46271

1789

1789

1789
the National Program of Cancer Registry 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 CA160687 and P30 ES00260. 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 (+RD12/0306/0050); 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 (CIBER - 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 U10 CA167552 from the NCI (Bethesda, MD). Nurses’ Health Study was supported by NIH grants U10 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, N.I.H, U.S. Department of Health and Human Services through contracts HHSN268201600018C, HHSN268201600010C, HHSN268201600012C, 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%20Written%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


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


Updated version
Access the most recent version of this article at:
doi: 10.1158/1055-9965.EPI-20-0275

Supplementary Material
Access the most recent supplemental material at:
http://cebp.aacrjournals.org/content/suppl/2020/06/16/1055-9965.EPI-20-0275.DC1

Cited articles
This article cites 48 articles, 4 of which you can access for free at:
http://cebp.aacrjournals.org/content/29/9/1784.full#ref-list-1

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, use this link
http://cebp.aacrjournals.org/content/29/9/1784.
Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.