HHS Public Access

Author manuscript

Cancer Epidemiol Biomarkers Prev. Author manuscript; available in PMC 2021 January 01.

Published in final edited form as:

Cancer Epidemiol Biomarkers Prev. 2020 July ; 29(7): 1501–1508. doi:10.1158/1055-9965.EPI-20-0091.

Associations between genetically predicted blood protein biomarkers and pancreatic cancer risk

Jingjing Zhu[1], **Xiang Shu**[2], **Xingyi Guo**[2], **Duo Liu**[1,3], **Jiandong Bao**[2], **Roger L Milne**[4,5,6], **Graham G Giles**[4,5,6], **Chong Wu**[7], **Mengmeng Du**[8], **Emily White**[9,10], **Harvey A Risch**[11], **Nuria Malats**[12], **Eric J. Duell**[13], **Phyllis J. Goodman**[14], **Donghui Li**[15], **Paige Bracci**[16], **Verena Katzke**[17], **Rachel E Neale**[18], **Steven Gallinger**[19], **Stephen K Van Den Eeden**[20], **Alan A Arslan**[21], **Federico Canzian**[22], **Charles Kooperberg**[9], **Laura E Beane-Freeman**[23], **Ghislaine Scelo**[24], **Kala Visvanathan**[25], **Christopher A. Haiman**[26], **Loïc Le Marchand**[1], **Herbert Yu**[1], **Gloria M Petersen**[27], **Rachael Stolzenberg-Solomon**[23], **Alison P Klein**[25,28], **Qiuyin Cai**[2], **Jirong Long**[2], **Xiao-Ou Shu**[2], **Wei Zheng**[2], **Lang Wu**[1]

[1.]Cancer Epidemiology Division, Population Sciences in the Pacific Program, University of Hawaii Cancer Center, University of Hawaii at Manoa, Honolulu, HI, USA

[2.]Division of Epidemiology, Department of Medicine, Vanderbilt Epidemiology Center, Vanderbilt-Ingram Cancer Center, Vanderbilt University Medical Center, Nashville, TN, USA

[3.]Department of Pharmacy, Harbin Medical University Cancer Hospital, Harbin, China

[4.]Cancer Epidemiology Division, Cancer Council Victoria, Melbourne, VIC, Australia

[5.]Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health, the University of Melbourne, VIC, Australia

[6.]Precision Medicine, School of Clinical Sciences at Monash Health, Monash University, Clayton, VIC, Australia

[7.]Department of Statistics, Florida State University, Tallahassee, FL, USA

[8.]Department of Epidemiology and Biostatistics, Memorial Sloan Kettering Cancer Center, New York, NY, USA

[9.]Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA, USA

[10.]Department of Epidemiology, University of Washington, Seattle, WA, USA

[11.]Department of Chronic Disease Epidemiology, Yale School of Public Health, New Haven, CT, USA

[12.]Spanish National Cancer Research Centre (CNIO) and CIBERONC, Madrid, Spain

[13.]Unit of Nutrition and Cancer, Cancer Epidemiology Research Program, Catalan Institute of Oncology (ICO-IDIBELL), L'Hospitalet de Llobregat, Spain

Corresponding to: Lang Wu, Cancer Epidemiology Division, Population Sciences in the Pacific Program, University of Hawaii Cancer Center, University of Hawaii at Manoa, Honolulu, HI, 96813, USA. lwu@cc.hawaii.edu. Phone: (808)564-5965.

Competing financial interests

The authors declare no competing financial interests.

[14.]SWOG Statistical Center, Fred Hutchinson Cancer Research Center, Seattle, WA, USA

[15.]Department of Gastrointestinal Medical Oncology, University of Texas MD Anderson Cancer Center, Houston, TX, USA

[16.]Department of Epidemiology and Biostatistics, University of California San Francisco, San Francisco, CA, USA

[17.]Division of Cancer Epidemiology, German Cancer Research Center (DKFZ), Heidelberg, Germany

[18.]Population Health Department, QIMR Berghofer Medical Research Institute, Brisbane, QLD, Australia

[19.]Lunenfeld-Tanenbaum Research Institute of Mount Sinai Hospital, Toronto, Ontario, Canada

[20.]Division of Research, Kaiser Permanente Northern California, Oakland, CA, USA

[21.]Department of Obstetrics and Gynecology, New York University School of Medicine, USA

[22.]Genomic Epidemiology Group, German Cancer Research Center (DKFZ), Heidelberg, Germany

[23.]Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Rockville, MD, USA

[24.]Genetic Epidemiology Group, Section of Genetics, International Agency for Research on Cancer, World Health Organization, Lyon, France

[25.]Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA

[26.]Center for Genetic Epidemiology, Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA, 90033, USA

[27.]Department of Health Sciences Research, Mayo Clinic College of Medicine, Rochester, MN, USA

[28.]Department of Oncology, Sidney Kimmel Comprehensive Cancer Center, Johns Hopkins School of Medicine, Baltimore, MD, USA

Abstract

Background—Pancreatic ductal adenocarcinoma (PDAC) is one of the most lethal malignancies with few known risk factors and biomarkers. Several blood protein biomarkers have been linked to PDAC in previous studies, but these studies have assessed only a limited number of biomarkers usually in small samples. In this study, we evaluated associations of circulating protein levels and PDAC risk using genetic instruments.

Methods—To identify novel circulating protein biomarkers of PDAC, we studied 8,280 cases and 6,728 controls of European descent from the Pancreatic Cancer Cohort Consortium and the Pancreatic Cancer Case-Control Consortium, using genetic instruments of protein quantitative trait loci (pQTL).

Results—We observed associations between predicted concentrations of 38 proteins and PDAC risk at a false discovery rate of < 0.05, including 23 of those proteins that showed an association even after Bonferroni correction. These include the protein encoded by *ABO*, which has been implicated as a potential target gene of PDAC risk variant. Eight of the identified proteins (LMA2L, TM11D, IP-10, ADH1B, STOM, TENC1, DOCK9, and CRBB2) were associated with PDAC risk after adjusting for previously reported PDAC risk variants (odds ratio ranged from 0.79 to 1.52). Pathway enrichment analysis showed that the encoding genes for implicated proteins were significantly enriched in cancer-related pathways, such as STAT3 and IL-15 production.

Conclusions—We identified 38 candidates of protein biomarkers for PDAC risk.

Impact—This study identifies novel protein biomarker candidates for PDAC, which if validated by additional studies, may contribute to the etiological understanding of PDAC development.

Keywords

Biomarkers; epidemiology; genetics; pancreatic cancer; risk

Introduction

Pancreatic cancer, 95% of which is pancreatic ductal adenocarcinoma (PDAC), is the second most commonly diagnosed gastrointestinal malignancy and the third leading cause of cancer-related death in the United States (US) (1). With a five-year survival of 8%, the incidence of pancreatic cancer keeps increasing in the US (2). Because pancreatic cancer is typically asymptomatic in early stages, most patients are diagnosed at an advanced stage, which precludes the possible application of curative surgery. Therefore, identifying biomarkers that would contribute to screening or early diagnosis in high-risk populations may improve pancreatic cancer outcomes. Serum CA 19–9 is currently the only biomarker for pancreatic cancer used in clinical settings. However, it is mainly used for diagnosing symptomatic patients, and monitoring disease prognosis and response to treatment (3). Besides CA 19–9, several other blood circulating proteins have been reported to be potentially associated with pancreatic cancer risk, such as CA242, PIVKA-II, PAM4, S100A6, OPN, RBM6, EphA2 and OPG (4–7), but the results in those studies are inconsistent. For example, those studies often only involved a small sample size and evaluated a few candidate proteins, and were often limited by a lack of external validation. Additionally, due to the observational study design, they were potentially subject to selection bias and residual and unmeasured confounding.

Mendelian randomization (MR) analysis is a widely applied design using genetic variants as instruments to evaluate the potential causal relationship between exposure and outcome (8–12). The nature of random assortment of alleles from parents to offspring during gamete formation makes such a design using genetic instruments to be less susceptible to biases encountered by conventional epidemiological studies (13,14).

In the current study, we aimed to use genetic variants as an instrument to study blood concentrations of proteins and to assess their associations with PDAC risk. Genome-wide association studies (GWAS) have identified hundreds of protein quantitative trait loci

(pQTL) (15,16), many of which can serve as strong instrumental variables. To our knowledge, this is the first large-scale study to comprehensively evaluate the associations between genetically predicted blood concentrations of a wide range of proteins and PDAC risk. We used data for 8,280 cases and 6,728 controls of European descent from the Pancreatic Cancer Cohort Consortium (PanScan) and the Pancreatic Cancer Case-Control Consortium (PanC4).

Methods

We conducted an extensive literature search to identify studies examining the associations between genetic variants at genome-wide scale and blood protein concentrations and based our analysis on a recently published comprehensive study (17). Focusing on a total of 3,301 healthy European descent individuals (2,481 and 820 in each of two sub-cohorts) in the INTERVAL study, Sun and colleagues identified 1,927 associations between 1,478 proteins and 764 genomic loci. In brief, 3,622 proteins in plasma were quantified by an aptamer-based multiplex protein assay (SOMAscan). Genotyping was performed using the Affymetrix Axiom UK Biobank genotyping array, with subsequent imputation based on a combined 1000 Genomes Phase 3-UK10K reference panel. After quality control, pQTL analyses for 3,283 SOMAmers were conducted separately for each sub-cohort with adjustment for age, sex, duration between blood draw and processing, and the first three principal components. The results from these two sub-cohorts were combined by fixed-effects inverse-variance meta-analysis. The estimated associations between genetic variants and protein concentrations were considered significant only if they meet all three criteria: 1) $P < 1.5 \times 10^{-11}$ in the meta-analysis (5×10^{-8}/3,283 aptamers tested); 2) $P < 0.05$ in both sub-cohorts; and 3) consistent effect across sub-cohorts. The pQTLs identified in this study were used to generate the instrumental variables for evaluating the associations between genetically predicted proteins concentrations in blood and pancreatic cancer risk. When protein concentrations were associated with more than one pQTL variant located at the same chromosome, the correlations between these SNPs were estimated using the Pairwise LD function of SNiPA (http://snipa.helmholtz-muenchen.de/snipa/index.php?task=pairwise_ld). Only independent SNPs ($R^2 < 0.1$ based on 1000 Genomes Project Phase 3 version 5 data for European descendants) were included to create a single instrument for each protein.

In the present study, we used data from GWAS conducted in the PanScan and PanC4 consortia downloaded from the database of Genotypes and Phenotypes (dbGaP), including 8,280 PDAC cases and 6,728 controls of European ancestry. Detailed information on GWAS from PanScan and PanC4 can be found elsewhere (18–23). In brief, four GWAS studies including PanScan I, PanScan II, PanScan III, and PanC4 were genotyped using the Illumina HumanHap550, 610-Quad, OmniExpress, and OmniExpressExome arrays, respectively. Standard quality control (QC) was performed according to the guidelines of each consortium (21). We excluded study participants who were related to each other, had gender discordance, had genetic ancestry other than European, had a low call rate (less than 98% and 94% in PanC4 and PanScan, respectively), or had missing information on age or sex. We removed duplicated SNPs, and those with a high missing call rate (at least 2% and 6% in PanC4 and PanScan, respectively) or with violations of Hardy-Weinberg equilibrium (HWE) ($P < 1\times10^{-4}$ and $P < 1\times10^{-7}$ in PanC4 and PanScan, respectively). For SNP data from

PanC4, we additionally excluded those with minor allele frequency < 0.005, with more than two discordant calls in duplicate samples, with more than one Mendelian error in HapMap control trios, and those with sex difference in allele frequency > 0.2 or in heterozygosity > 0.3 for autosomes/XY in European descendants. Genotype imputation was conducted using Minimac3 after prephasing with SHAPEIT from a reference panel of the Haplotype Reference Consortium (r1.1 2016) (24–26). Imputed SNPs with an imputation quality of at least 0.3 were retained. We then assessed associations between individual variants and PDAC risk after adjustment of age, sex and top ten principal components (Supplementary material; Supplementary Table 1).

Based on the summary statistics from the above-mentioned pQTL study (17) and the analyses of PanScan/PanC4 GWAS datasets, we used the inverse variance weights (IVW) method to assess the association between genetically predicted blood protein concentrations and PDAC risk (27,28). The beta coefficient of the association between each protein and PDAC risk was estimated using the formula of $\sum_i \beta_{i,GX} * \beta_{i,GY} * \sigma_{i,GY}^{-2} / \left(\sum_i \beta_{i,GX}^2 * \sigma_{i,GY}^{-2} \right)$, and its corresponding standard error was calculated by $1/(\sum_i \beta_{i,GX}^2 * \sigma_{i,GY}^{-2})^{0.5}$. Here, $\beta_{i,GX}$ represents the beta coefficient adopted from the pQTL study for the association between the ith SNP and concentration of the protein of interest; $\beta_{i,GY}$ and $\sigma_{i,GY}$ represent the estimated beta coefficient and standard error of the association between the ith SNP and PDAC risk in PanScan/PanC4 GWAS. We further computed odds ratios (ORs) and confidence intervals (CIs) by exponentiation of the beta coefficients. A Benjamini-Hochberg false discovery rate (FDR) of < 0.05 was used to define statistical significance. We also performed the analyses using individual level data. For this analysis, first we generated the predicted protein concentration for each subject in PanScan/PanC4 GWAS based on the individual-level genetic data and the beta coefficient from the pQTL study for the association between pQTL SNP and protein of interest. We then assessed the associations between predicted protein concentrations and PDAC risk. We further conducted conditional analysis with adjustments for previously identified risk variants to assess whether the observed associations between genetically predicted protein concentrations and PDAC risk in our main analyses were independent of the risk variants identified in GWAS studies. Previously reported PDAC risk SNPs that are available in the current dataset (rs2816938, rs3790844, rs1486134, rs2736098, rs35226131, rs401681, rs17688601, rs78417682, rs6971499, rs2941471, rs10094872, rs1561927, rs505922, rs9581943, rs9543325, rs4795218, rs11655237, rs1517037) were adjusted for in the conditional analysis. Additionally, we performed sensitivity analyses using data from different subgroups by consortium to assess the robustness of the significant associations.

For the proteins that were associated with PDAC risk, we performed an enrichment analysis of the genes encoding these proteins to examine whether they are enriched in specific pathways, functions or networks, by using Ingenuity Pathway Analysis (IPA) software. Detailed information of the methods has been described by the tool developer (29). In brief, the level of enrichment was estimated by assessing the overlap of the observed tested gene sets and the predicted regulated gene sets using Fisher's exact test.

Results

We were able to assess associations between genetically predicted protein levels andPDAC risk for 1,226 proteins using pQTLs as instruments. Using the IVW method, we identified 38 proteins for which the genetically predicted concentrations showed associations with PDAC risk at a false discovery rate of < 0.05 (23 proteins after Bonferroni-correction) (Tables 1 and 2); eight that remained significant after adjusting for known PDAC risk variants identified in previous GWAS (Table 1). Positive associations were observed for seven of these proteins, including Beta-crystallin B2 (CRBB2), Dedicator of cytokinesis protein 9 (DOCK9), VIP36-like protein (LMAN2L), Erythrocyte band 7 integral membrane protein (STOM), Tensin-2 (TENC1), Transmembrane protease serine 11D (TM11D), and Alcohol dehydrogenase 1B (ADH1B) (ORs ranging from 1.17 to 1.52) (Table 1). We observed a negative association between predicted protein concentration of C-X-C motif chemokine 10 (IP-10) and PDAC risk (OR per one standard deviation increase in genetically predicted protein =0.79, 95% CI: 0.69–0.91; *P*-value=1.19×10^{-3}) (Table 1).

The associations for the other 30 proteins were substantially attenuated after adjusting for previously identified PDAC risk variants, potentially due to 1) the previously identified associations of risk SNPs with PDAC at these loci may be mediated through these proteins identified in the current study, or 2) confounding effects. Of these 30 proteins, 14 were positively associated with PDAC risk, including Histo-blood group ABO system transferase (BGAT), C1GALT1-specific chaperone 1 (C1GLC), Cadherin-5, Platelet glycoprotein 4 (CD36-ANTIGEN), Desmoglein-2, Protein FAM3B, CD209 Antigen (DC-SIGN), GDNF family receptor alpha-like (GFRAL), D-glucuronyl C5-epimerase (GLCE), Neurogenic locus notch homolog protein 1 (Notch1), Tolloid-like protein 1 (TLL1), N-acetyllactosaminide beta-1,3-N- acetylglucosaminyltransferase 2 (B3GN2), Carbohydrate sulfotransferase 11 (CHSTB), and Angiopoietin-1 receptor, soluble (sTie-2) (ORs ranging from 1.12 to 3.62) (Table 2). Conversely, an inverse association between predicted protein concentrations and PDAC risk was identified for P-Selectin, Intestinal-type alkaline phosphatase, Endoglin, Insulin-like growth factor 1 receptor (IGF-IR), Interleukin-3 receptor subunit alpha (IL-3Ra), Insulin receptor (IR), Protein jagged-1 (JAG1), Leukemia inhibitory factor receptor (LIF-sR), Hepatocyte growth factor receptor (Met), E-selectin (sE-Selectin), Carbohydrate sulfotransferase 15 (CHST15), Thrombospondin type-1 domain-containing protein 1 (THSD1), Adhesion G protein-coupled receptor F5 (GP116), Interleukin-6 receptor subunit beta (gp130, soluble), Vascular endothelial growth factor receptor 2 (VEGF sR2), and Protein FAM177A1 (F177A) (ORs ranging from 0.38 to 0.86) (Table 2).

Based on subgroup analyses, the associations of the identified 38 proteins, in general, were robust across the GWAS subsets (PanScan I, II, and III; PanScan I and II; PanC4 and PanScan I and II; and PanC4) (Supplementary Table 2).

The IPA analysis showed enrichment in several cancer-related function pathways for the genes encoding the proteins identified by our study. The top canonical pathways identified included IL-15 production (*P*-value = 2.71×10^{-6}) and STAT3 (*P*-value = 5.25×10^{-6}) (Table 3).

Discussion

This is the first study with a large sample size to systematically evaluate the associations between genetically predicted circulating protein concentrations and PDAC risk using pQTLs as study instruments. Overall, we identified 38 proteins that were significantly associated with PDAC risk after FDR correction, including eight that showed an association with PDAC risk independently from the previously identified PDAC risk variants. If confirmed, our data suggest new knowledge on the etiology of PDAC, and provide a list of proteins as candidate blood biomarkers for assessing risk of PDAC, a malignancy with universally high case fatality.

Previous studies have suggested blood concentrations of CA242, PIVKA-II, PAM4, S100A6, OPN, RBM6, EphA2 and OPG to be associated with pancreatic cancer risk (4–7). However, with the exception of S100A6 and OPG, a pQTL was not identified for these proteins (17). Using the corresponding pQTL rs62143206 of S100A6 as an instrumental variable, we did not observe evidence of association for S100A6 (OR=1.01, 95% CI: 0.91–1.13; *P*-value=0.86) with PDAC. For OPG, by using the corresponding pQTL rs570618 as an instrumental variable, we observed an association OR=1.35, 95% CI: 1.04–1.76, *P*-value=0.03, although this was not significant after correcting for multiple comparisons. Nevertheless, the direction of the association is consistent with that identified in previous work. Our inconsistent finding with previous studies for S100A6 might be explained by either the weak instrument used in our study or potential biases in previous studies that used a conventional observational design.

In this large study, we identified eight PDAC-associated proteins that are independent of PDAC risk variants previously identified in GWAS. Compared with GWAS, which aim to identify novel susceptibility variants by assessing the association between each genetic variant and disease risk across the genome, the current study has improved statistical power by aggregating the effects of several SNPs into one continuous testing unit, the genetically predicted blood concentration of protein, when applicable. In the current study, we used both *cis* and *trans* pQTL as genetic instruments whenever possible (Tables 1–2). Previous research has supported a potential role for some of the novel proteins identified in this study in pancreatic tumorigenesis. Based on an immunohistochemical analysis, significantly higher expression of tensin-2 was observed in pancreatic tumor tissues than in adjacent normal tissues (30). In the same study, there were also positive associations of tensin-2 with glucose metabolism related insulin receptor substrate 1 and glucose transporter type 4, the proliferation marker ki-67, the angiogenesis marker CD31, and the mesenchymal markers N-cadherin and fibronectin, suggesting a potential role of tensin-2 in pancreatic cancer metabolism, proliferation, angiogenesis and epithelial-mesenchymal transition process (30). Protein TM11D, encoded by gene *TMPRSS11D*, serves as an efficient activator of macrophage stimulating protein (MSP). MSP can further stimulate the activation of its receptor, RON, which has been suggested to be overexpressed early in the progression of pancreatic malignancy (31,32).

For the other 30 proteins identified in this study, for which associations with PDAC risk were mainly explained by previously reported PDAC risk variants, some were also

suggested to play a role in pancreatic cancer development based on *in vitro*/*in vivo* or human studies. For example, GWAS has identified the *ABO* gene as a susceptibility locus for PDAC risk (21). The protective T allele of rs505922, the instrument SNP for the protein encoded by *ABO,* is in linkage disequilibrium with a single base pair deletion that encodes the O antigen. Genotype-inferred O blood type was shown to be associated with a reduced risk of PDAC compared with other blood types, which was suggested to be possibly attributed to altered inflammation state, glycosyltransferase activity, or differentiated expression of blood group antigens (33,34). Based on *in vitro* experiments, knockdown of *C1GALT1C1*, the encoding gene for protein C1GLC, promoted migration and survival but inhibited proliferation of pancreatic cancer cells (35). In contrast, for some of the proteins identified, it is worth noting that the directions of the observed associations are not consistent with those suggested in the literature. For example, CHST15 is an enzyme that biosynthesizes Chondroitin sulfate, which is known to be able to promote tumor invasion and metastasis. *CHST15* mRNA was found to be highly expressed in pancreatic cancer cell lines (36). Pancreatic tumor growth was inhibited after CHST15 protein blood concentrations were reduced in both mice and humans (37). In the current study, however, we found that a low level of genetically determined CHST15 concentration was associated with an increased risk of pancreatic cancer. Possible explanations for this inconsistency may include that the focus of the current study is the genetically regulated circulating protein concentrations, whereas the measured protein concentrations in previous studies may be influenced by both inherent and extrinsic factors. Additional well-designed studies with directly measured protein concentrations are warranted to better understand the relationship between the identified proteins and pancreatic cancer risk.

The strengths of our study include its large sample size for the main association analyses, providing high statistical power to detect proteins associated with PDAC risk. The use of genetic instruments potentially minimized several biases that are commonly encountered in conventional observational studies. However, several limitations of the current work need to be recognized. First, our results may be susceptible to potential pleiotropic effects. For example, rs3197999, the instrument for proteins CRBB2, DOCK9, TENC1, and TM11D, has also been associated with several other traits, including primary sclerosing cholangitis, Crohn's disease, and ulcerative colitis (38–40). Similarly, rs2519093, which was the instrument for proteins IL-3Ra and sE-Selectin, as well as one of the variants constituting the instrument for P-Selectin, C1GLC, FAM3B, GLCE and THSD1, was shown to be associated with coronary artery disease, allergy and venous thromboembolism (41–43). Although most of these traits do not appear to be strongly related to pancreatic carcinogenesis, allergy is known to be potentially associated with pancreatic cancer risk (44,45), and previous studies have linked Crohn's disease and ulcerative colitis with pancreatic cancer risk (46,47). Results of our MR-Egger regression analyses for protein FAM3B (*P*-value=0.55) and P-Selectin (*P*-value=0.73), which involved three variants as instrument, suggested that their associations were less likely to be influenced by potential directional pleiotropic effects (48). Second, in this study we were only able to capture the genetically regulated components of circulating protein concentrations, so that their utility of as a biomarker is unclear due to the impact of environmental factors. Further prospective studies with measured circulating protein concentrations in pre-disease blood samples are

warranted to validate the potential predicting role of our identified proteins in pancreatic cancer. Third, our analysis largely relies on the pQTLs identified by previous GWAS of circulating protein concentrations; thus our ability to evaluate candidate protein biomarkers for pancreatic cancer was limited by whether a pQTL had been identified for some of these proteins. We expect that additional protein biomarkers can be identified when new knowledge is generated regarding the pQTL for additional proteins. Fourth, research has suggested that specific variables, such as smoking and body weight, are related to protein levels in blood (49,50). Ideally for our study the instrument pQTL SNPs would be identified in analyses with adjustment of relevant variables; however, this is not the case for the INTERVAL study. Further research is needed to validate our findings.

In summary, in this large study, we identified multiple novel protein biomarkers, for which the genetically predicted circulating concentrations were associated with PDAC risk. Our study may serve as a basis for future investigation of these proteins to better understand the underlying mechanisms of PDAC and to advance the development of effective biomarker panels for risk assessment of PDAC.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

The datasets used for the analyses described in this manuscript were obtained from dbGaP at http://www.ncbi.nlm.nih.gov/sites/entrez?db=gap through dbGaP accession phs000206.v5.p3 and phs000648.v1.p1. The authors thank Laufey Amundadottir, Eric Jacobs, and Idan Ben-Barak for their help for this manuscript. The authors also would like to thank all of the individuals for their participation in the parent studies and all the researchers, clinicians, technicians and administrative staff for their contribution to the studies. Lang Wu is supported by NCI R00CA218892. Duo Liu is supported by the Harbin Medical University Cancer Hospital. The PanScan study was funded in whole or in part with federal funds from the National Cancer Institute (NCI), US National Institutes of Health (NIH) under contract number HHSN261200800001E. Additional support was received from NIH/NCI K07 CA140790, the American Society of Clinical Oncology Conquer Cancer Foundation, the Howard Hughes Medical Institute, the Lustgarten Foundation, the Robert T. and Judith B. Hale Fund for Pancreatic Cancer Research and Promises for Purple. A full list of acknowledgments for each participating study is provided in the Supplementary Note of the manuscript with PubMed ID: 25086665. For the PanC4 GWAS study, the patients and controls were derived from the following PANC4 studies: Johns Hopkins National Familial Pancreas Tumor Registry, Mayo Clinic Biospecimen Resource for Pancreas Research, Ontario Pancreas Cancer Study (OPCS), Yale University, MD Anderson Case Control Study, Queensland Pancreatic Cancer Study, University of California San Francisco Molecular Epidemiology of Pancreatic Cancer Study, International Agency of Cancer Research and Memorial Sloan Kettering Cancer Center. This work is supported by NCI R01CA154823 Genotyping services were provided by the Center for Inherited Disease Research (CIDR). CIDR is fully funded through a federal contract from the National Institutes of Health to The Johns Hopkins University, contract number HHSN2682011000111. The WHI program is funded by the National Heart, Lung, and Blood Institute, National Institutes of Health, U.S. Department of Health and Human Services through contracts, HHSN268201600018C, HHSN268201600001C, HHSN268201600002C, HHSN268201600003C, and HHSN268201600004C. This manuscript was prepared in collaboration with investigators of the WHI, and has been reviewed and/or approved by the Women's Health Initiative (WHI). WHI investigators are listed at https://www.whi.org/researchers/Propose a Paper/Write a Paper-Resources/Acknowledgement Lists:Short Lists. SELECT study is supported by National Institutes of Health grant award number U10 CA37429 (CD Blanke), and UM1 CA182883 (CM Tangen/IM Thompson). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Abbreviations list

PDAC	Pancreatic ductal adenocarcinoma

Pqtl	protein quantitative trait loci
MR	Mendelian randomization
GWAS	Genome-wide association studies
PanScan	the Pancreatic Cancer Cohort Consortium
PanC4	the Pancreatic Cancer Case-Control Consortium
QC	quality control
HWE	Hardy-Weinberg equilibrium
IVW	inverse variance weights
ORs	odds ratios
Cis	confidence intervals
FDR	false discovery rate
IPA	Ingenuity Pathway Analysis

References

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2018. CA Cancer J Clin. 2018;68(1):7–30. [PubMed: 29313949]
2. Raimondi S, Maisonneuve P, Lowenfels AB. Epidemiology of pancreatic cancer: an overview. Nat Rev Gastroenterol Hepatol. 2009 12;6(12):699–708. [PubMed: 19806144]
3. Ballehaninna UK, Chamberlain RS. Serum CA 19–9 as a Biomarker for Pancreatic Cancer-A Comprehensive Review. Indian J Surg Oncol. 2011 6;2(2):88–100. [PubMed: 22693400]
4. Koshikawa N, Minegishi T, Kiyokawa H, Seiki M. Specific detection of soluble EphA2 fragments in blood as a new biomarker for pancreatic cancer. Cell Death Dis. 2017 26;8(10):e3134. [PubMed: 29072678]
5. Loosen SH, Neumann UP, Trautwein C, Roderburg C, Luedde T. Current and future biomarkers for pancreatic adenocarcinoma. Tumour Biol J Int Soc Oncodevelopmental Biol Med. 2017 6;39(6):1010428317692231.
6. Duan B, Hu X, Fan M, Xiong X, Han L, Wang Z, et al. RNA-Binding Motif Protein 6 is a Candidate Serum Biomarker for Pancreatic Cancer. Proteomics Clin Appl. 2019 6 17;e1900048. [PubMed: 31207145]
7. Tartaglione S, Pecorella I, Zarrillo SR, Granato T, Viggiani V, Manganaro L, et al. Protein Induced by Vitamin K Absence II (PIVKA-II) as a potential serological biomarker in pancreatic cancer: a pilot study. Biochem Medica. 2019 6 15;29(2):020707.
8. Lawlor DA, Harbord RM, Sterne JAC, Timpson N, Davey Smith G. Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. Stat Med. 2008 4 15;27(8):1133–63. [PubMed: 17886233]
9. Jia J, Dou P, Gao M, Kong X, Li C, Liu Z, et al. Assessment of Causal Direction Between Gut Microbiota-Dependent Metabolites and Cardiometabolic Health: A Bidirectional Mendelian Randomization Analysis. Diabetes. 2019;68(9):1747–55. [PubMed: 31167879]
10. Shu X, Wu L, Khankari NK, Shu X-O, Wang TJ, Michailidou K, et al. Associations of obesity and circulating insulin and glucose with breast cancer risk: a Mendelian randomization analysis. Int J Epidemiol. 2019 01;48(3):795–806. [PubMed: 30277539]

11. Carreras-Torres R, Johansson M, Gaborieau V, Haycock PC, Wade KH, Relton CL, et al. The Role of Obesity, Type 2 Diabetes, and Metabolic Factors in Pancreatic Cancer: A Mendelian Randomization Study. J Natl Cancer Inst. 2017 01;109(9).
12. Langdon RJ, Richmond RC, Hemani G, Zheng J, Wade KH, Carreras-Torres R, et al. A Phenome-Wide Mendelian Randomization Study of Pancreatic Cancer Using Summary Genetic Data. Cancer Epidemiol Biomark Prev Publ Am Assoc Cancer Res Cosponsored Am Soc Prev Oncol. 2019 12;28(12):2070–8.
13. Didelez V, Sheehan N. Mendelian randomization as an instrumental variable approach to causal inference. Stat Methods Med Res. 2007 8;16(4):309–30. [PubMed: 17715159]
14. Lawlor DA, Harbord RM, Sterne JAC, Timpson N, Davey Smith G. Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. Stat Med. 2008 4 15;27(8):1133–63. [PubMed: 17886233]
15. Suhre K, Arnold M, Bhagwat AM, Cotton RJ, Engelke R, Raffler J, et al. Connecting genetic risk to disease end points through the human blood plasma proteome. Nat Commun. 2017 27;8:14357. [PubMed: 28240269]
16. Enroth S, Johansson A, Enroth SB, Gyllensten U. Strong effects of genetic and lifestyle factors on biomarker variation and use of personalized cutoffs. Nat Commun. 2014 8 22;5:4684. [PubMed: 25147954]
17. Sun BB, Maranville JC, Peters JE, Stacey D, Staley JR, Blackshaw J, et al. Genomic atlas of the human plasma proteome. Nature. 2018;558(7708):73–9. [PubMed: 29875488]
18. 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 9;46(9):994–1000. [PubMed: 25086665]
19. 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 3;42(3):224–8. [PubMed: 20101243]
20. 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 9;41(9):986–90. [PubMed: 19648918]
21. 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 08;9(1):556. [PubMed: 29422604]
22. 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 8;47(8):911–6. [PubMed: 26098869]
23. 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 10 11;7(41):66328–43. [PubMed: 27579533]
24. Delaneau O, Marchini J, Zagury J-F. A linear complexity phasing method for thousands of genomes. Nat Methods. 2012 2;9(2):179–81.
25. Howie BN, Donnelly P, Marchini J. A Flexible and Accurate Genotype Imputation Method for the Next Generation of Genome-Wide Association Studies. Schork NJ, editor. PLoS Genet. 2009 6 19;5(6):e1000529. [PubMed: 19543373]
26. the Haplotype Reference Consortium. A reference panel of 64,976 haplotypes for genotype imputation. Nat Genet. 2016 10;48(10):1279–83. [PubMed: 27548312]
27. Shu X, Bao J, Wu L, Long J, Shu X-O, Guo X, et al. Evaluation of Associations between Genetically Predicted Circulating Protein Biomarkers and Breast Cancer Risk. Int J Cancer. 2019 7 2;
28. Wu L, Shu X, Bao J, Guo X, Kote-Jarai Z, Haiman CA, et al. Analysis of Over 140,000 European Descendants Identifies Genetically Predicted Blood Protein Biomarkers Associated with Prostate Cancer Risk. Cancer Res. 2019 9 15;79(18):4592–8. [PubMed: 31337649]
29. Krämer A, Green J, Pollard J, Tugendreich S. Causal analysis approaches in Ingenuity Pathway Analysis. Bioinforma Oxf Engl. 2014 2 15;30(4):523–30.

30. Cheng L-C, Chen Y-L, Cheng A-N, Lee AY-L, Cho C-Y, Huang J-S, et al. AXL phosphorylates and up-regulates TNS2 and its implications in IRS-1-associated metabolism in cancer cells. J Biomed Sci. 2018 11 12;25(1):80. [PubMed: 30419905]

31. Kawaguchi M, Kataoka H. MST1 (macrophage stimulating 1 (hepatocyte growth factor-like)). Atlas Genet Cytogenet Oncol Haematol [Internet]. 2013 11 [cited 2019 Jul 31];(12). Available from: http://hdl.handle.net/2042/51872

32. Camp ER, Yang A, Gray MJ, Fan F, Hamilton SR, Evans DB, et al. Tyrosine kinase receptor RON in human pancreatic cancer: expression, function, and validation as a target. Cancer. 2007 3 15;109(6):1030–9. [PubMed: 17311308]

33. Wolpin BM, Kraft P, Xu M, Steplowski E, Olsson ML, Arslan AA, et al. Variant ABO blood group alleles, secretor status, and risk of pancreatic cancer: results from the pancreatic cancer cohort consortium. Cancer Epidemiol Biomark Prev Publ Am Assoc Cancer Res Cosponsored Am Soc Prev Oncol. 2010 12;19(12):3140–9.

34. Wolpin BM, Chan AT, Hartge P, Chanock SJ, Kraft P, Hunter DJ, et al. ABO blood group and the risk of pancreatic cancer. J Natl Cancer Inst. 2009 3 18;101(6):424–31. [PubMed: 19276450]

35. Hofmann BT, Schlüter L, Lange P, Mercanoglu B, Ewald F, Fölster A, et al. COSMC knockdown mediated aberrant O-glycosylation promotes oncogenic properties in pancreatic cancer. Mol Cancer. 2015 5 29;14:109. [PubMed: 26021314]

36. Takakura K, Shibazaki Y, Yoneyama H, Fujii M, Hashiguchi T, Ito Z, et al. Inhibition of Cell Proliferation and Growth of Pancreatic Cancer by Silencing of Carbohydrate Sulfotransferase 15 In Vitro and in a Xenograft Model. Batra SK, editor. PLOS ONE. 2015 12 7;10(12):e0142981. [PubMed: 26642349]

37. Nishimura M, Yahagi N, Itoi T, Ochiai Y, Matsuda Y. A translational study to investigate the role of carbohydrate sulfotransferase 15 for pancreatic cancer biology from in vitro to first-in-human clinical research. J Clin Oncol. 2015 5 20;33(15_suppl):e22201–e22201.

38. Melum E, Franke A, Schramm C, Weismüller TJ, Gotthardt DN, Offner FA, et al. Genome-wide association analysis in primary sclerosing cholangitis identifies two non-HLA susceptibility loci. Nat Genet. 2011 1;43(1):17–9. [PubMed: 21151127]

39. Julià A, Domènech E, Chaparro M, García-Sánchez V, Gomollón F, Panés J, et al. A genome-wide association study identifies a novel locus at 6q22.1 associated with ulcerative colitis. Hum Mol Genet. 2014 12 20;23(25):6927–34. [PubMed: 25082827]

40. Franke A, McGovern DPB, Barrett JC, Wang K, Radford-Smith GL, Ahmad T, et al. Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci. Nat Genet. 2010 12;42(12):1118–25. [PubMed: 21102463]

41. Klarin D, Emdin CA, Natarajan P, Conrad MF, INVENT Consortium, Kathiresan S Genetic Analysis of Venous Thromboembolism in UK Biobank Identifies the ZFPM2 Locus and Implicates Obesity as a Causal Risk Factor. Circ Cardiovasc Genet. 2017 4;10(2).

42. Pickrell JK, Berisa T, Liu JZ, Ségurel L, Tung JY, Hinds DA. Detection and interpretation of shared genetic influences on 42 human traits. Nat Genet. 2016;48(7):709–17. [PubMed: 27182965]

43. van der Harst P, Verweij N. Identification of 64 Novel Genetic Loci Provides an Expanded View on the Genetic Architecture of Coronary Artery Disease. Circ Res. 2018 02;122(3):433–43. [PubMed: 29212778]

44. Gandini S, Lowenfels AB, Jaffee EM, Armstrong TD, Maisonneuve P. Allergies and the risk of pancreatic cancer: a meta-analysis with review of epidemiology and biological mechanisms. Cancer Epidemiol Biomark Prev Publ Am Assoc Cancer Res Cosponsored Am Soc Prev Oncol. 2005 8;14(8):1908–16.

45. Olson SH, Hsu M, Satagopan JM, Maisonneuve P, Silverman DT, Lucenteforte E, et al. Allergies and risk of pancreatic cancer: a pooled analysis from the Pancreatic Cancer Case-Control Consortium. Am J Epidemiol. 2013 9 1;178(5):691–700. [PubMed: 23820785]

46. Hemminki K, Li X, Sundquist J, Sundquist K. Cancer risks in ulcerative colitis patients. Int J Cancer. 2008 9 15;123(6):1417–21. [PubMed: 18561319]

47. Hemminki K, Li X, Sundquist J, Sundquist K. Cancer risks in Crohn disease patients. Ann Oncol Off J Eur Soc Med Oncol. 2009 3;20(3):574–80.

48. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. Int J Epidemiol. 2015 4;44(2):512–25. [PubMed: 26050253]

49. Wingerd J, Sponzilli EE. Concentrations of serum protein fractions in white women: effects of age, weight, smoking, tonsillectomy, and other factors; Clin Chem. 1977 7;23(7):1310–7. [PubMed: 559554]

50. Yang M, Kohler M, Heyder T, Forsslund H, Garberg HK, Karimi R, et al. Long-term smoking alters abundance of over half of the proteome in bronchoalveolar lavage cell in smokers with normal spirometry, with effects on molecular pathways associated with COPD. Respir Res. 2018 08;19(1):40. [PubMed: 29514648]

Table 1.

Genetically predicted protein concentrations that are independently associated with pancreatic cancer risk after adjustment for previously identified risk SNPs

Protein	Protein full name	Protein-encoding gene	Region for protein encoding gene	Instrument variants	Type of pQTL	OR[a]	Lower bound 95% CI[a]	Upper bound 95% CI[a]	*P*-value	FDR *P*-value[b]	*P*-value after adjusting for risk SNPs[c]
LMA2L	VIP36-like protein	*LMAN2L*	2q11.2	rs2271893	*cis*	1.39	1.15	1.68	6.47×10^{-4}	3.17×10^{-2}	7.72×10^{-4}
TM11D	Transmembrane protease serine 11D	*TMPRSS11D*	4q13.2	rs3197999	*trans*	1.17	1.06	1.29	1.11×10^{-3}	3.78×10^{-2}	2.44×10^{-3}
IP-10	C-X-C motif chemokine 10	*CXCL10*	4q21.1	rs11548618	*cis*	0.79	0.69	0.91	1.19×10^{-3}	3.93×10^{-2}	9.71×10^{-4}
ADH1B	Alcohol dehydrogenase 1B	*ADH1B*	4q23	rs13085791	*trans*	1.22	1.08	1.37	1.28×10^{-3}	4.14×10^{-2}	2.81×10^{-3}
STOM	Erythrocyte band 7 integral membrane protein	*STOM*	9q33.2	rs6770670	*trans*	1.19	1.07	1.33	1.05×10^{-3}	3.78×10^{-2}	2.27×10^{-3}
TENC1	Tensin-2	*TNS2*	12q13.13	rs3197999	*trans*	1.25	1.09	1.42	1.11×10^{-3}	3.78×10^{-2}	2.44×10^{-3}
DOCK9	Dedicator of cytokinesis protein 9	*DOCK9*	13q32.3	rs3197999	*trans*	1.32	1.12	1.56	1.11×10^{-3}	3.78×10^{-2}	2.44×10^{-3}
CRBB2	Beta-crystallin B2	*CRYBB2*	22q11.23	rs3197999	*trans*	1.52	1.18	1.95	1.11×10^{-3}	3.78×10^{-2}	2.44×10^{-3}

[a] OR (odds ratio) and CI (confidence interval) per one standard deviation increase in genetically predicted protein after adjustment for age, sex, and top 10 principle components.

[b] FDR *P*-value: false discovery rate (FDR) adjusted *P*-value; associations with a FDR p≤0.05 considered statistically significant

[c] Associations were adjusted for risk SNPs include: rs2816938, rs3790844, rs1486134, rs2736098, rs35226131, rs401681, rs17688601, rs78417682, rs6971499, rs2941471, rs10094872, rs1561927, rs505922, rs9581943, rs9543325, rs4795218, rs11655237, and rs1517037

Table 2.

Genetically predicted protein concentrations in association with pancreatic cancer risk that are potentially influenced by previously identified risk SNPs.

Protein	Protein full name	Protein-encoding gene	Region for protein encoding gene	Instrument variants	Type of pQTL	OR[a]	Lower bound 95% CI[a]	Upper bound 95% CI[a]	*P*-value	FDR *P*-value[b]	*P*-value after adjusting for risk SNPs[c]
P-Selectin	P-Selectin	*SELP*	1q24.2	rs74227709 rs6136 rs2519093	*trans* *cis* *trans*	0.86	0.80	0.92	2.67×10^{-5}	1.49×10^{-3}	0.32
sE-Selectin	E-selectin	*SELE*	1q24.2	rs2519093	*trans*	0.84	0.80	0.88	4.80×10^{-13}	3.68×10^{-11}	0.13
B3GN2	N-acetyllactosaminide beta-1,3-N-acetylglucosaminyltransferase 2	*B3GNT2*	2p15	rs2519093	*trans*	1.97	1.64	2.37	4.80×10^{-13}	3.68×10^{-11}	0.13
Alkaline phosphatase, intestine	Intestinal-type alkaline phosphatase	*ALPI*	2q37.1	rs550057	*trans*	0.43	0.35	0.53	1.91×10^{-15}	4.68×10^{-13}	0.56
VEGF sR2	Vascular endothelial growth factor receptor 2	*KDR*	4q12	rs34231037 rs635634	*cis* *trans*	0.80	0.74	0.87	4.86×10^{-8}	3.31×10^{-6}	0.55
TLL1	Tolloid-like protein 1	*TLL1*	4q32.3	rs8176747	*trans*	1.30	1.11	1.52	1.10×10^{-3}	3.78×10^{-2}	0.60
LIF-sR	Leukemia inhibitory factor receptor	*LIFR*	5p13.1	rs635634	*trans*	0.49	0.41	0.59	2.10×10^{-13}	2.57×10^{-11}	0.10
gp130, soluble	Interleukin-6 receptor subunit beta	*IL6ST*	5q11.2	rs635634 rs11574765	*trans* *cis*	0.73	0.63	0.84	2.94×10^{-5}	1.57×10^{-3}	0.70
GFRAL	GDNF family receptor alpha-like	*GFRAL*	6p12.1	rs72975088 rs8176672	*trans* *cis*	1.33	1.13	1.57	6.15×10^{-4}	3.14×10^{-2}	0.47
GP116	Adhesion G protein-coupled receptor F5	*ADGRF5*	6p12.3	rs2519093	*trans*	0.76	0.71	0.82	4.80×10^{-13}	3.68×10^{-11}	0.13
CD36-ANTIGEN	Platelet glycoprotein 4	*CD36*	7q21.11	rs8176693	*trans*	1.27	1.10	1.46	1.01×10^{-3}	3.78×10^{-2}	0.61
Met	Hepatocyte growth factor receptor	*MET*	7q31	rs635634	*trans*	0.57	0.49	0.66	2.10×10^{-13}	2.57×10^{-11}	0.10
sTie-2	Angiopoietin-1 receptor, soluble	*TEK*	9p21.2	rs8176693	*trans*	1.28	1.10	1.48	1.01×10^{-3}	3.78×10^{-2}	0.61
Endoglin	Endoglin	*ENG*	9q34.11	rs635634	*trans*	0.41	0.32	0.52	2.10×10^{-13}	2.57×10^{-11}	0.10

Protein	Protein full name	Protein-encoding gene	Region for protein encoding gene	Instrument variants	Type of pQTL	OR[a]	Lower bound 95% CI[a]	Upper bound 95% CI[a]	*P*-value	FDR *P*-value[b]	*P*-value after adjusting for risk SNPs[c]
BGAT	Histo-blood group ABO system transferase	*ABO*	9q34.2	rs505922	*cis*	1.20	1.15	1.24	5.74×10^{-21}	2.35×10^{-18}	NA[d]
Notch1	Neurogenic locus notch homolog protein 1	*NOTCH1*	9q34.3	rs8176743	*trans*	1.46	1.16	1.83	1.10×10^{-3}	3.78×10^{-2}	0.60
CHST15	Carbohydrate sulfotransferase 15	*CHST15*	10q26.13	rs550057	*trans*	0.52	0.44	0.61	1.91×10^{-15}	4.68×10^{-13}	0.56
CHSTB	Carbohydrate sulfotransferase 11	*CHST11*	12q23.3	rs687621	*trans*	3.62	2.77	4.74	5.57×10^{-21}	2.35×10^{-18}	0.46
THSD1	Thrombospondin type-1 domain-containing protein 1	*THSD1*	13q14.3	rs41292808 rs2519093	*trans* *cis*	0.74	0.65	0.83	5.85×10^{-7}	3.77×10^{-5}	0.13
F177A	Protein FAM177A1	*FAM177A1*	14q13.2	rs550057 rs679574	*Trans* *trans*	0.63	0.54	0.73	1.54×10^{-9}	1.11×10^{-7}	0.45
GLCE	D-glucuronyl C5-epimerase	*GLCE*	15q23	rs11854180 rs2519093	*trans* *cis*	1.13	1.07	1.19	1.85×10^{-5}	1.08×10^{-3}	0.44
IGF-IR	Insulin-like growth factor 1 receptor	*IGF1R*	15q26.3	rs635634	*trans*	0.38	0.29	0.49	2.10×10^{-13}	3.22×10^{-11}	0.10
Cadherin-5	Cadherin-5	*CDH5*	16q21	rs8176746	*trans*	1.12	1.05	1.20	1.10×10^{-3}	3.78×10^{-2}	0.60
Desmoglein-2	Desmoglein-2	*DSG2*	18q12.1	rs2704050 rs687621	*trans* *cis*	1.88	1.59	2.23	3.31×10^{-13}	3.67×10^{-11}	0.99
IR	Insulin receptor	*INSR*	19p13.2	rs507666	*trans*	0.69	0.63	0.77	3.59×10^{-13}	4.40×10^{-11}	0.12
DC-SIGN	CD209 antigen	*CD209*	19p13.2	rs505922	*trans*	1.32	1.25	1.40	5.74×10^{-21}	2.35×10^{-18}	NA[d]
JAG1	Protein jagged-1	*JAG1*	20p12.2	rs7041 rs550057	*trans* *trans*	0.70	0.59	0.82	9.08×10^{-6}	7.95×10^{-4}	0.41
FAM3B	Protein FAM3B	*FAM3B*	21q22.3	rs2608894 rs73226194 rs2519093	*cis* *trans* *trans*	1.15	1.06	1.25	8.01×10^{-4}	3.78×10^{-2}	0.64
IL-3Ra	Interleukin-3 receptor subunit alpha	*IL3RA*	Xp22.3	rs2519093	*trans*	0.78	0.73	0.84	4.80×10^{-13}	4.90×10^{-11}	0.13
C1GLC	C1GALT1-specific chaperone 1	*C1GALT1C1*	Xq24	rs7787942 rs2519093	*trans* *trans*	1.39	1.28	1.50	5.34×10^{-15}	1.09×10^{-12}	3.02×10^{-3}

[a]OR (odds ratio) and CI (confidence interval) per one standard deviation increase in genetically predicted protein after adjustment for age, sex, and top 10 principle components.

[b] FDR *P*-value: false discovery rate (FDR) adjusted *P*-value; associations with a FDR p≤0.05 considered statistically significant.

[c] Associations were adjusted for risk SNPs include: rs2816938, rs3790844, rs1486134, rs2736098, rs35226131, rs401681, rs17688601, rs78417682, rs6971499, rs2941471, rs10094872, rs1561927, rs505922, rs9581943, rs9543325, rs4795218, rs11655237, rs1517037.

[d] Instrument SNP itself is a known PC risk SNP.

Cancer Epidemiol Biomarkers Prev. Author manuscript; available in PMC 2021 January 01.

Table 3.

Canonical pathways, diseases, bio functions and networks associated with the genes encoding identified pancreatic cancer risk associated proteins.

Top canonical pathways	Top diseases and disorders	Molecular and cellular functions	Top networks
IL-15 Production; STAT3 Pathway; Sperm motility; Heparan Sulfate Biosynthesis (Late Stages) Granulocyte Adhesion and Diapedesis	Cancer; Organismal Injury and Abnormalities; Dermatological Diseases and Conditions; Tumor Morphology; Inflammatory Response;	Cell-To-Cell Signaling and Interaction; Carbohydrate Metabolism; Cellular Development; Cellular Function and Maintenance; Cellular Growth and Proliferation;	Cardiovascular System Development and Function, Organismal Development, Cellular Movement; Cell Signaling, Cell-To-Cell Signaling and Interaction, Cancer