**Carcinogenesis** vol.33 no.7 pp.1384–1390, 2012
doi:10.1093/carcin/bgs151
*Advance Access publication April 20, 2012*

# Pathway analysis of genome-wide association study data highlights pancreatic development genes as susceptibility factors for pancreatic cancer

**Donghui Li[†,‡], Eric J.Duell[1,†,‡], Kai Yu[2,†,‡], Harvey A.Risch[3,‡], Sara H.Olson[4,‡], Charles Kooperberg[5,‡], Brian M.Wolpin[6,7,‡], Li Jiao[8,‡], Xiaoqun Dong[†], Bill Wheeler[9,‡], Alan A.Arslan[10,11,12], H. Bas Bueno-de-Mesquita[13,14], Charles S.Fuchs[6,7], Steven Gallinger[15], Myron Gross[16], Patricia Hartge[2], Robert N.Hoover[2], Elizabeth A.Holly[17], Eric J.Jacobs[18], Alison P.Klein[19,20], Andrea LaCroix[5], Margaret T.Mandelson[5,21], Gloria Petersen[22], Wei Zheng[23], Ilir Agalliu[24], Demetrius Albanes[2], Marie-Christine Boutron-Ruault[25], Paige M.Bracci[17], Julie E.Buring[26,27], Federico Canzian[28], Kenneth Chang[29], Stephen J.Chanock[2,30,31], Michelle Cotterchio[32,33], J.Michael Gaziano[34], Edward L.Giovannucci[7,35,36], Michael Goggins[37], Göran Hallmans[38], Susan E.Hankinson[7,35], Judith A.Hoffman Bolton[39], David J.Hunter[7,35], Amy Hutchinson[2,31], Kevin B.Jacobs[2,31,40], Mazda Jenab[41], Kay-Tee Khaw[42], Peter Kraft[35,43], Vittorio Krogh[44], Robert C.Kurtz[45], Robert R.McWilliams[46], Julie B.Mendelsohn[2], Alpa V.Patel[18], Kari G.Rabe[22], Elio Riboli[47], Xiao-Ou Shu[23], Anne Tjønneland[48], Geoffrey S.Tobias[2], Dimitrios Trichopoulos[35,49], Jarmo Virtamo[50], Kala Visvanathan[37], Joanne Watters[51], Herbert Yu[3], Anne Zeleniuch-Jacquotte[11,12], Laufey Amundadottir[2,30,‡,§] and Rachael Z.Stolzenberg-Solomon[2,‡,§,*]**

Department of Gastrointestinal Medical Oncology, The University of Texas M. D. Anderson Cancer Center, Houston, TX, USA, [1]Catalan Institute of Oncology (ICO-IDIBELL), Barcelona, Spain, [2]Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Department of Health and Human Services, Bethesda, MD, USA, [3]Yale University School of Public Health, New Haven, CT, USA, [4]Department of Epidemiology and Biostatistics, Memorial Sloan-Kettering Cancer Center, New York, NY, USA, [5]Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA, USA, [6]Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, MA, USA, [7]Channing Laboratory, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA, [8]Department of Medicine, Baylor College of Medicine, Houston, TX, USA, [9]Information Management Services, Silver Spring, MD, USA, [10]Department of Obstetrics and Gynecology, New York University School of Medicine, New York, NY, USA, [11]Department of Environmental Medicine, New York University School of Medicine, New York, NY, USA, [12]New York University Cancer Institute, New York, NY, USA, [13]National Institute for Public Health and the Environment (RIVM), Bilthoven, The Netherlands, [14]Department of Gastroenterology and Hepatology, University Medical Center Utrecht, Utrecht, The Netherlands, [15]Samuel Lunenfeld Research Institute, Mount Sinai Hospital, University of Toronto, Toronto, Canada, [16]Department of Laboratory Medicine/Pathology, School of Medicine, University of Minnesota, Minneapolis, MN, USA, [17]Department of Epidemiology and Biostatistics, University of California San Francisco, San Francisco, CA, USA, [18]Epidemiology Research Program, American Cancer Society, Atlanta, GA, USA, [19]Department of Oncology, The Johns Hopkins University School of Medicine, Baltimore, MD, USA, [20]Department of Epidemiology, The Bloomberg School of Public Health, The Sol Goldman Pancreatic Research Center, The Johns Hopkins Medical Institutions, Baltimore, MD, USA, [21]Group Health Center for Health Studies, Seattle, WA, USA, [22]Department of Health Sciences Research, Mayo Clinic, Rochester, MN, USA, [23]Division of Epidemiology, Department of Medicine, Vanderbilt Epidemiology Center, and Vanderbilt-Ingram Cancer Center, Vanderbilt University, Nashville, TN, USA, [24]Department of Epidemiology and Population Health, Albert Einstein College of Medicine, Bronx, NY, USA, [25]Inserm, Paris-Sud University, Institut Gustave-Roussy, Villejuif, France, [26]Department of Ambulatory Care and Prevention, Harvard Medical School, Boston, MA, USA, [27]Divisions of Preventive Medicine and Aging, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA, [28]Division of Cancer Epidemiology, German Cancer Research Center (DKFZ), Heidelberg, Germany, [29]Comprehensive Digestive Disease Center, University of California, Irvine Medical Center, Orange, CA, USA, [30]Laboratory of Translational Genomics, Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Department of Health and Human Services, Bethesda, MD, USA, [31]Core Genotyping Facility, Advanced Technology Program, SAIC-Frederick Inc., NCI-Frederick, Frederick, MD, USA, [32]Dalla Lana School of Public Health, University of Toronto, Toronto, Ontario, Canada, [33]Prevention and Cancer Control, Cancer Care Ontario, Toronto, Ontario, Canada, [34]Physicians' Health Study, Divisions of Aging, Cardiovascular Medicine, and Preventive Medicine, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, and Massachusetts Veterans Epidemiology Research and Information Center, Veterans Affairs Boston Healthcare System, Boston, MA, USA, [35]Department of Epidemiology, Harvard School of Public Health, Boston, MA, USA, [36]Department of Nutrition, Harvard School of Public Health, Boston, MA, USA, [37]Departments of Oncology, Pathology and Medicine, The Johns Hopkins University School of Medicine, Baltimore, MD, USA, [38]Department of Public Health and Clinical Medicine, Nutritional Research, Umeå University, Umeå, Sweden, [39]Department of Epidemiology, The Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, USA, [40]Bioinformed Consulting Services, Gaithersburg, MD, USA, [41]International Agency for Research on Cancer (IARC/WHO), Lyon, France, [42]Department of Public Health and Primary Care, Clinical Gerontology, Addenbrooke's Hospital, University of Cambridge, Cambridge, UK, [43]Department of Biostatistics, Harvard School of Public Health, Boston, MA, USA, [44]Nutritional Epidemiology Unit, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy, [45]Department of Medicine, Memorial Sloan-Kettering Cancer Center, New York, NY, USA, [46]Division of Medical Oncology, Mayo Clinic, Rochester, MN, USA, [47]Division of Epidemiology and Biostatistics, School of Public Health, Imperial College London, London, UK, [48]Institute of Cancer Epidemiology, Danish Cancer Society, Copenhagen, Denmark, [49]Bureau of Epidemiologic Research, Academy of Athens, Athens, Greece, [50]Department of Chronic Disease Prevention, National Institute for Health and Welfare, Helsinki, Finland and [51]Division of Cancer Prevention and Population Control, National Cancer Institute, National Institutes of Health, Department of Health and Human Services, Bethesda, MD, USA

[*]To whom correspondence should be addressed. Rachael Stolzenberg-Solomon, Nutritional Epidemiology Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, 6120 Executive Boulevard, Room 3022, Rockville, MD 20852-7232, USA.
Tel: +1 301 594 2939; Fax: +1 301 496 6829;
Email: rs221z@nih.gov.
Correspondence may also be addressed to Laufey Amundadottir.
Tel: +1 301 594 8131; Fax: +1 301 402 3134;
Email: amundadottirl@mail.nih.gov

**Four loci have been associated with pancreatic cancer through genome-wide association studies (GWAS). Pathway-based analysis of GWAS data is a complementary approach to identify groups of genes or biological pathways enriched with disease-associated single-nucleotide polymorphisms (SNPs) whose individual effect sizes may be too small to be detected by standard single-locus methods. We used the adaptive rank truncated product method in a pathway-based analysis of GWAS data from 3851 pancreatic cancer cases and 3934 control participants pooled from 12 cohort studies and 8 case–control studies (PanScan). We compiled 23 biological pathways hypothesized to be relevant to pancreatic cancer and observed a nominal association between pancreatic cancer and five pathways ($P < 0.05$), i.e. pancreatic development,**

**Abbreviations:** ARTP, adaptive rank truncated product; GWAS, genome-wide association study; SNP, single-nucleotide polymorphism; TGF, transforming growth factor.

[†]Co-first authors.
[‡]Writing committee member.
[§]Co-senior authors.

***Helicobacter pylori*** **lacto/neolacto, hedgehog, Th1/Th2 immune response and apoptosis ($P = 2.0 \times 10^{-6}$, $1.6 \times 10^{-5}$, 0.0019, 0.019 and 0.023, respectively). After excluding previously identified genes from the original GWAS in three pathways (*NR5A2*, *ABO* and *SHH*), the pancreatic development pathway remained significant ($P = 8.3 \times 10^{-5}$), whereas the others did not. The most significant genes ($P < 0.01$) in the five pathways were *NR5A2*, *HNF1A*, *HNF4G* and *PDX1* for pancreatic development; *ABO* for *H.pylori* lacto/neolacto; *SHH* for hedgehog; *TGFBR2* and *CCL18* for Th1/Th2 immune response and *MAPK8* and *BCL2L11* for apoptosis. Our results provide a link between inherited variation in genes important for pancreatic development and cancer and show that pathway-based approaches to analysis of GWAS data can yield important insights into the collective role of genetic risk variants in cancer.**

## Introduction

Genome-wide association studies (GWAS) have become the standard for investigating the association between common inherited genetic variants across the genome and risk of complex diseases such as cancer. Two GWAS (PanScan 1 and PanScan 2) recently identified four susceptibility loci for pancreatic cancer at chromosomes: 9q34.2, 13q22.1, 1q32.1 and 5p15.33 (1,2). Despite these important findings, the statistical tests applied in GWAS are typically restricted to single markers; furthermore, some markers and genes may be missed because of the stringent statistical threshold necessary to minimize false-positive findings (genome-wide significance) (3,4). Pathway-based analysis of GWAS data is a complementary approach for identifying groups of genes or biological pathways enriched with disease-associated single-nucleotide polymorphisms (SNPs) whose individual effect sizes may be too small to be detected by standard single-locus methods.

The idea for pathway-based approaches stems from two concepts: (i) that a functional pathway represents a series of biochemical actions leading to an end point or cellular function such as an activated or inactivated enzyme or metabolite, an enhanced or repressed signaling cascade, a repaired DNA strand or a coordinated immune response and (ii) that small changes due to variation in the expression of genes involved in a functional pathway may lead to an outcome such as cancer (5).

We performed a comprehensive pathway-based analysis of the combined dataset of two pancreatic cancer GWAS, PanScan 1 and PanScan 2, using an adaptive combination of *P*-values in the adaptive rank truncated product (ARTP) method (6). Twenty-three biological pathways and groups of genes known or hypothesized from the literature to be important in pancreatic tumorigenesis were selected, including pancreas development, DNA repair, apoptosis, cell cycle signaling, immune function and inflammatory pathways, insulin resistance, PI3 kinase, Wnt, Notch, hedgehog and transforming growth factor (TGF)-β signaling. We confirmed the major results from the ARTP analysis with a logic regression analysis (7,8).

## Materials and methods

### *Study population*

The study population included 3851 pancreatic cancer cases and 3934 control participants from the previously conducted GWAS in the Pancreatic Cancer Cohort Consortium and the Pancreatic Cancer Case Control Consortium (PanC4) (1,2). Briefly, this collaborative project included 1528 incident cases and 1594 controls from nested case–control studies of 12 cohort studies and 2323 cases and 2340 controls from 8 case–control studies (1,2). Cases were defined as participants diagnosed with primary adenocarcinoma of the exocrine pancreas; controls were matched to cases according to birth year, sex and self-reported race/ethnicity and were free of pancreatic cancer at the time of recruitment (1,2). Genotyping was performed by the National Cancer Institute's Core Genotyping Facility using the Illumina HumanHap550 and HumanHap550-Duo SNP arrays (PanScan 1) and Illumina Human 610-Quad arrays (PanScan 2) (1,2). PanScan 1 and PanScan 2 were approved by the Institutional Review Board of each participating institution and National Cancer Institute's Special Studies Institutional Review Board (1,2).

### *Pathway selection*

Pathways were chosen on the basis of our current understanding of the etiology and molecular mechanisms of pancreatic cancer with the aim of constructing concise core pathways known to be important for pancreatic biology; 23 pathways or groups of genes were compiled (Table I) based on literature searches and online resources accessed between 2008 and 2010 (e.g. http://www.SNPs3D.org; http://sciencepark.mdanderson.org/labs/wood/DNA_Repair_Genes.html and http://www.genome.jp/kegg/pathway.html). These included pathways related to pancreatic organ development and differentiation, DNA repair, apoptosis, cell cycle regulation, immune response, *Helicobacter pylori* infection, inflammation, insulin resistance, PI3 kinase, Wnt, Notch, hedgehog and TGF-β signaling pathways. For example, the DNA repair pathway, including subpathways, were included in this analysis based upon results from previous candidate gene association studies (9–11). Diabetes mellitus (12,13), glucose intolerance (12,14,15) and chronic pancreatitis also appear to predispose individuals to pancreatic cancer (16). Allergies have been associated with reduced risk of pancreatic cancer in several studies (17) but little is known about the genetic basis of this association. Two pathways related to allergies were considered for this study: one including genes related to the balance between T-helper 1 and T-helper 2 cells (Th1/Th2) and the other including genes related to serum IgE levels (18). The pancreas develops from the endodermal epithelium of the foregut of the vertebrate embryo. A series of transcriptional regulators govern the development and cell type differentiation of the gland. We compiled a list of transcriptional regulators important for early pancreatic development by reviews of the literature and by searching GO and KEGG terms (19–22) Predisposing genetic factors for pancreatic cancer remain poorly understood; however, genetic variation in genes that influence the above risk factors are viable candidate genes for interrogation. The total number of genes was 577 (of which 4 were included in 3 pathways and 30 in 2 pathways).

### *Statistical analysis*

SNP association analysis was conducted with use of the logistic regression model using a boundary for each gene beginning 20 kb upstream of the transcriptional start site and ending 10 kb downstream of the transcriptional end site of the gene (including exons, introns and untranslated regions). This model was fit for genotype trend effects (1 d.f.) adjusted for study, age, sex, self-described ancestry and 10 principal components for the population

**Table I.** Association of pathways with risk of pancreatic cancer

| Pathway | No. of genes | No. of SNPs | *P*-value |
|---|---|---|---|
| Pancreas development | 22 | 271 | $2.0 \times 10^{-6}$ |
| *Helicobacter pylori* lacto/neolacto | 18 | 292 | $1.6 \times 10^{-5}$ |
| Hedgehog | 31 | 588 | 0.0019 |
| Th1/Th2 immune response | 32 | 448 | 0.019 |
| Apoptosis | 42 | 665 | 0.023 |
| Nucleotide excision repair | 25 | 285 | 0.078 |
| Cell cycle | 43 | 424 | 0.080 |
| DNA polymerase | 13 | 126 | 0.093 |
| Notch | 25 | 487 | 0.12 |
| TGF-β | 40 | 954 | 0.13 |
| PI3 kinase | 43 | 798 | 0.21 |
| Homologous recombination repair | 22 | 366 | 0.24 |
| Mismatch repair | 10 | 142 | 0.26 |
| DNA damage response | 16 | 198 | 0.27 |
| *H.pylori* protein metabolism | 20 | 188 | 0.30 |
| Insulin resistance | 23 | 537 | 0.35 |
| IgE | 18 | 214 | 0.41 |
| Wnt | 23 | 414 | 0.49 |
| *H.pylori* extracellular | 9 | 289 | 0.49 |
| *H.pylori* cytokine signaling | 27 | 302 | 0.53 |
| Base excision repair | 17 | 171 | 0.61 |
| Glycation | 45 | 542 | 0.80 |
| *H.pylori* other/unclassified | 51 | 495 | 0.94 |

Results from the ARTP pathway analysis. The analysis was adjusted for age in 10 years categories, sex, study arm and 10 principal components of population stratification.

stratification adjustment, which included the top 5 principal components identified in the cohort studies and the top 5 principal components identified in the case–control studies (1,2). *P*-values for individual SNPs were based on the 1 d.f. Wald test derived from the fitted logistic regression model.

First, we conducted a gene-based analysis to evaluate the association between a candidate gene/region and cancer risk. The test statistic used was the minP statistic, which was the minimum *P*-value among all *P*-values from the single SNP analysis conducted within the candidate gene. The *P*-value for the gene-based analysis (called gene *P*-value) can be evaluated through a bootstrap procedure. Second, we conducted the pathway analysis to evaluate the association between a set of candidate genes included in a pathway and cancer risk. The pathway analysis was based on the ARTP method (6) and was implemented in the R package ARTP (http://dceg.cancer.gov/bb/tools/artp). The ARTP method aims at maximizing the association signal by combining gene-level *P*-values from a set of selected genes within the pathway into the test statistic and uses a bootstrap procedure to estimate its *P*-value and has been shown to account properly for the type I error (6). The bootstrap procedure is used for the purpose of generating datasets under the null hypothesis while keeping the correlation among SNPs the same as that in the observed dataset. The *P*-value for both the gene-based and pathway analyses was initially estimated by 30 000 parametric bootstrap steps. We re-evaluated *P*-values for genes or pathways that had initially estimated *P*-values of $<0.05$ using 1 000 000 bootstrap steps.

As a complementary approach to the ARTP method, we used a logic regression model (7,8) to reanalyze several promising pathways identified by the ARTP method to determine whether those pathways were enriched with interactions. The ARTP method looks for marginal effects from individual SNPs but does not aim at detecting epistatic interactions among SNPs. In contrast, logic regression is an adaptive regression methodology that attempts to identify 'logic' (binary) combinations of predictors that are associated with a regression outcome. Each SNP is recoded as two binary predictors: one is based on whether at least one variant allele is present ('dominant coding') and the other is based on whether two variant alleles are present ('recessive coding'). We fit models using a simulated annealing algorithm. Model selection was conducted using cross-validation and permutation tests. A Bayesian approach to model selection was used to generate a list of possible candidates of predictors.

## Results

Of the 23 pathways analyzed (Table I), the most statistically significant association was seen for the pancreatic developmental pathway ($P = 2.0 \times 10^{-6}$) and the *H.pylori* lacto/neolacto pathway ($P = 1.6 \times 10^{-5}$). Three additional pathways were nominally significant: hedgehog signaling ($P = 0.0019$), Th1/Th2 immune response ($P = 0.019$) and apoptosis ($P = 0.023$). The top three pathways (pancreatic development, *H.pylori* lacto/neolacto and hedgehog) were significant after Bonferroni correction for the 23 pathways tested ($P < 0.002$). However, after excluding genes (i.e. removing all SNPs within the gene) previously identified by the initial GWAS (*NR5A2* from the pancreatic development pathway, *ABO* from the *H.pylori* lacto/neolacto and *SHH* from the hedgehog pathway), the pancreatic development pathway remained significant ($P = 8.3 \times 10^{-5}$), whereas the other two pathways became nonsignificant ($P > 0.05$).

We also computed gene-level *P*-values for the 577 genes included in the study; 46 genes had *P*-values of $<0.05$ (Table II). The major genes contributing to the significant pathways include *NR5A2*, *HNF1A*, *HNF4G*, *PDX1* and *HNF1B* for pancreatic development; *ABO* for *H.pylori* lacto/neolacto; *SHH*, *BTRC* and *HHIP* for hedgehog; *TGFBR2*, *CCL18* and *IL13RA2* for Th1/Th2 immune response and *MAPK8*, *BCL2L11*, *FAS*, *FASLG* and *CASP7* for the apoptosis pathway. For the other pathways analyzed, zero to four genes were nominally significant ($P < 0.05$) (Table II).

Individual SNPs that were significant at the $P < 0.001$ level for the five significant pathways are listed in Table III. The pancreatic development pathway showed 15 SNPs: 6 located in the *NR5A2* gene, 5 in *HNF1A*, 3 in *HNF4G* and 1 in *HNF1B*. Five SNPs in the *H.pylori* lacto/neolacto pathway were significant; however, they were all located within the *ABO* gene previously identified in the GWAS (1,2). Two SNPs in the hedgehog signaling pathway were significant at this level, located approximately 10–15 kb upstream of the *SHH* gene; again, both were identified in PanScan 1, but the association was not replicated in PanScan 2 (1,2). Two SNPs in the *TGFBR2* gene within the Th1/Th2 immune response pathway were significant at a threshold of $P < 0.001$; these SNPs were also included in the TGF-β pathway that was not significant overall. Finally, three SNPs in the apoptosis pathway were significant at the same *P*-value level: one in *MAPK8* and two in *BCL2L11*.

We also observed a significant association between the pancreatic development pathway and cancer risk using logic regression analysis. The SNPs that occurred most frequently in the models were rs2816939, rs3762399, rs2737621 (*NR5A2*), rs7310409, rs7953249 (*HNF1A*), rs2943547 (*HNF4G*) and rs2688 (*HNF1B*). The results of the Bayesian version of logic regression were compared 1000 times with a permuted response. The fit on the permuted data was always worse than the fit on the real data, thus providing strong evidence of an association between the pancreatic development pathway and pancreatic cancer. For the Th1/Th2 immune response pathway and apoptosis genes, logic regression also provided some evidence of associations with pancreatic cancer (data not shown).

## Discussion

Our pathway-based analysis of GWAS data has shown that common germ line variation in pancreatic developmental genes may be important susceptibility factors for pancreatic cancer. The genes that contributed to this significant association include *NR5A2*, *HNF1A*, *HNF4G*, *PDX1* and *HNF1B*. This association remains significant even after removing variants in the *NR5A2* gene shown previously to be associated with pancreatic cancer risk ($P < 0.001$). Four additional pathways showed nominally significant association with risk of pancreatic cancer ($P < 0.05$), i.e. *H.pylori* lacto/neolacto, hedgehog signaling, apoptosis and Th1/Th2 immune response, although genes previously implicated in pancreatic cancer risk may drive the association for the hedgehog (*SHH*) and *H.pylori* lacto/neolacto (*ABO*) pathways.

The five genes that contributed to the significant association with the pancreatic development pathway are important components of the transcriptional networks governing embryonic pancreatic development and differentiation as well as maintaining pancreatic homeostasis in adults (23,24). *PDX1* (pancreas-duodenal homeobox 1) regulates the very early steps of exocrine pancreas development (25). *NR5A2* is a direct downstream target of *PDX1* in this process (26). *HNF1A* and *HNF1B* encode hepatocyte nuclear factor 1 alpha and beta, also known as transcription factors 1 and 2 (*TCF1* and *TCF2*), respectively. These proteins belong to the homeobox family of DNA-binding proteins and regulate expression of a large number of genes. *HNF1A* primarily regulates the growth and function of islet β cells, and *HNF1B* plays an essential role in controlling pancreatic organogenesis and differentiation (23). Consistent with our observations, *HNF1A* was identified as the top hit for pancreatic cancer in a separate analysis of PanScan data by assessing markers previously identified in GWAS of phenotypes other than pancreatic cancer (27). Heterozygous compound knockout mouse models have shown that *PDX1*, *NR5A2*, *HNF1A* and *HNF1B* act in a tightly regulated feedback circuit in regulating pancreas development and differentiation (26,28). Therefore, even subtle differences in the relative activity of any of these genes may have profound consequences on overall network activity. Notably, the hedgehog signaling pathway, in particular the *SHH* gene, also plays an essential role during embryonic pancreatic development (29). Genes involved in organ development and differentiation contribute to the ability of tumor cells to proliferate and evade cell death, but they also often alter cell plasticity, i.e. reprogram cells to a state that may give rise to a tumor (29).

Mutations in *HNF1A*, *PDX1* and *HNF1B* are responsible for maturity onset diabetes of the young (MODY) types 3, 4 and 5, respectively (30,31). Both mutations and common variants in *HNF1A* and *HNF1B* have been associated with the risk of type II diabetes (32–34). Common variants in *NR5A2*, *HNF1B* and *HNF4G* (35) also have been associated with body mass index in recent GWAS. A recent study has reported a critical role of *NR5A2* in phosphatidylcholine signaling pathway regulating fatty acid and glucose homeostasis (36). Because

**Table II.** Genes associated with risk of pancreatic cancer at a $P < 0.05$

| Pathway | Gene | No. of SNPs | *P*-value | Most significant SNP |
|---|---|---|---|---|
| Pancreas development | *NR5A2* | 58 | $1.0 \times 10^{-6}$ | rs3790844 |
| | *HNF1A* | 15 | 0.00014 | rs7310409 |
| | *HNF4G* | 16 | 0.00048 | rs1805100 |
| | *PDX1* | 8 | 0.0079 | rs9554197 |
| | *HNF1B* | 29 | 0.019 | rs4794758 |
| *Helicobacter pylori* lacto/neolacto | *ABO* | 20 | $1.0 \times 10^{-6}$ | rs505922 |
| Hedgehog | *SHH* | 13 | $2.5 \times 10^{-5}$ | rs167020 |
| | *BTRC* | 19 | 0.016 | rs11191017 |
| | *HHIP* | 16 | 0.038 | rs17721701 |
| Th1/Th2 immune response | *TGFBR2* | 43 | 0.00062 | rs2043136 |
| | *CCL18* | 8 | 0.0063 | rs1719220 |
| | *IL13RA2* | 5 | 0.020 | rs638376 |
| Apoptosis | *MAPK8* | 7 | 0.0033 | rs1062225 |
| | *BCL2L11* | 14 | 0.0057 | rs13396983 |
| | *FAS* | 22 | 0.016 | rs4406737 |
| | *FASLG* | 8 | 0.038 | rs2021840 |
| | *CASP7* | 19 | 0.041 | rs7906704 |
| Nucleotide excision repair | *RPA1* | 23 | 0.0086 | rs2287321 |
| | *GTF2H3* | 7 | 0.014 | rs11572966 |
| | *LIG1* | 16 | 0.033 | rs3730913 |
| | *CDK7* | 6 | 0.049 | rs12651858 |
| Cell cycle | *TFDP1* | 15 | 0.0013 | rs6577059 |
| | *SKP1* | 7 | 0.013 | rs4958217 |
| | *CDK7* | 6 | 0.049 | rs12651858 |
| | *GADD45A* | 12 | 0.049 | rs647008 |
| DNA polymerase | *POLG* | 10 | 0.0092 | rs976072 |
| | *POLL* | 6 | 0.017 | rs3730477 |
| Notch | *MAML1* | 5 | 0.0025 | rs7734102 |
| | *RBBP8* | 10 | 0.044 | rs7234479 |
| TGF-β | *TGFBR2* | 43 | 0.0006 | rs2043136 |
| | *SMAD1* | 14 | 0.043 | rs7698944 |
| PI3 kinase | *PDPK1* | 1 | 0.0049 | rs13336495 |
| | *AKT3* | 29 | 0.015 | rs2125231 |
| Homologous recombination repair | *RAD52* | 17 | 0.0068 | rs1051669 |
| | *RBBP8* | 10 | 0.044 | rs7234479 |
| Mismatch repair | *PMS2* | 9 | 0.036 | rs2228006 |
| | *MLH3* | 2 | 0.039 | rs175057 |
| DNA damage response | *FANCI* | 15 | 0.011 | rs976072 |
| *H.pylori* protein metabolism | *GGCT* | 11 | 0.020 | rs38410 |
| Insulin resistance | *RETN* | 2 | 0.020 | rs1423096 |
| | *INSR* | 57 | 0.029 | rs2042902 |
| IgE | *IL13RA2* | 5 | 0.019 | rs638376 |
| Wnt | *AXIN1* | 24 | 0.042 | rs12719801 |
| *H.pylori*-cytokine signaling | *NOD1* | 16 | 0.024 | rs2529445 |
| Base excision repair | *MBD4* | 7 | 0.036 | rs140693 |
| Glycation | *APP* | 71 | 0.041 | rs375369 |

No significant genes were observed in the *H.pylori* extracellular and *H.pylori* other/unclassified pathways.

obesity and long-term type II diabetes are known risk factors for pancreatic cancer, it is possible that these genes may contribute to pancreatic cancer, partially through an increased risk of obesity and diabetes.

In addition to their roles in regulating the development and function of the pancreas, *HNF1A* and *HNF1B* also control terminal differentiation and cell fate commitment in the gut epithelium (37,38). Somatic mutations of the *HNF1A* gene have been reported in several types of human cancer, suggesting a tumor suppressor role (39–41). *HNF1A* silencing by small interfering RNA in hepatocellular carcinoma cells induces overexpression of several genes encoding growth factor receptors, components of the translational machinery, cell cycle and angiogenesis regulators, with, in particular, activation of the mammalian target of rapamycin pathway (42). Moreover, *HNF1A* has been recognized as a master regulator of plasma protein fucosylation (43) and plasma levels of C-reactive protein (44,45). This suggests that *HNF1A* may also contribute to pancreatic cancer via regulation of immunity, tumor progression and metastasis as well as through metabolic and inflammatory pathways. Overall, the pancreatic development pathway may have an impact on pancreatic cancer risk through multiple diversified mechanisms.

We also observed weaker associations of the Th1/Th2 immune response and apoptosis genes with pancreatic cancer. Genes in the Th1/Th2 pathway influence the balance of T-helper cells; individuals with

**Table III.** Highly significant SNPs ($P < 0.001$) in pathways that are associated with pancreatic cancer

| Pathway | Marker[a] | Chr[b] | Gene[c] | Alleles[d] | MAF[e] | *P*-value[f] | Allelic OR (95% CI) |
|---|---|---|---|---|---|---|---|
| Pancreas development | rs3790844[g] | 1q32.1 | *NR5A2* | T,C | 0.259/0.214 | $1.3 \times 10^{-10}$ | 0.77 (0.72–0.84) |
| | rs3790843[g] | 1q32.1 | *NR5A2* | G,A | 0.317/0.273 | $5.0 \times 10^{-9}$ | 0.81 (0.75–0.87) |
| | rs2821367 | 1q32.1 | *NR5A2* | T,C | 0.325/0.366 | $1.6 \times 10^{-6}$ | 1.18 (1.10–1.27) |
| | rs2816939 | 1q32.1 | *NR5A2* | T,C | 0.142/0.167 | $2.0 \times 10^{-5}$ | 1.22 (1.11–1.33) |
| | rs2821347 | 1q32.1 | *NR5A2* | G,A | 0.142/0.166 | $2.5 \times 10^{-5}$ | 1.21 (1.11–1.33) |
| | rs2737621 | 1q32.1 | *NR5A2* | T,C | 0.122/0.144 | $6.3 \times 10^{-5}$ | 1.22 (1.11–1.34) |
| | rs7310409 | 12q24.31 | *HNF1A* | G,A | 0.387/0.423 | $1.0 \times 10^{-5}$ | 1.16 (1.09–1.24) |
| | rs2464196 | 12q24.31 | *HNF1A* | C,T | 0.302/0.331 | $6.3 \times 10^{-5}$ | 1.15 (1.07–1.24) |
| | rs1169300 | 12q24.31 | *HNF1A* | G,A | 0.302/0.331 | $7.9 \times 10^{-5}$ | 1.15 (1.07–1.23) |
| | rs735396 | 12q24.31 | *HNF1A* | A,G | 0.358/0.386 | $1.0 \times 10^{-4}$ | 1.14 (1.07–1.22) |
| | rs7953249 | 12q24.31 | *HNF1A* | A,G | 0.428/0.456 | 0.00023 | 1.13 (1.06–1.21) |
| | rs1805100 | 8q21.11 | *HNF4G* | G,A | 0.472/0.506 | $3.2 \times 10^{-5}$ | 1.15 (1.08–1.23) |
| | rs2977926 | 8q21.11 | *HNF4G* | T,G | 0.421/0.390 | 0.00034 | 0.88 (0.82–0.94) |
| | rs2943547 | 8q21.11 | *HNF4G* | G,A | 0.470/0.440 | 0.00062 | 0.89 (0.84–0.95) |
| | rs4794758 | 17q12 | *HNF1B* | C,T | 0.266/0.244 | 0.00073 | 0.88 (0.82–0.95) |
| *Helicobacter pylori* lacto/neolacto | rs505922[g] | 9q34.2 | *ABO* | T,C | 0.350/0.395 | $2.0 \times 10^{-8}$ | 1.21 (1.13–1.30) |
| | rs657152[g] | 9q34.2 | *ABO* | G,T | 0.375/0.417 | $2.5 \times 10^{-7}$ | 1.19 (1.12–1.27) |
| | rs630014[g] | 9q34.2 | *ABO* | C,T | 0.475/0.436 | $1.3 \times 10^{-6}$ | 0.85 (0.80–0.91) |
| | rs2073828 | 9q34.2 | *ABO* | G,A | 0.414/0.370 | $1.6 \times 10^{-6}$ | 0.85 (0.79–0.91) |
| | rs495828 | 9q34.2 | *ABO* | G,T | 0.211/0.238 | $6.3 \times 10^{-5}$ | 1.18 (1.09–1.28) |
| Hedgehog | rs167020[g] | 7q36.3 | *SHH* | G,A | 0.258/0.292 | $2.5 \times 10^{-6}$ | 1.19 (1.11–1.28) |
| | rs172310[g] | 7q36.3 | *SHH* | C,A | 0.279/0.314 | $3.2 \times 10^{-6}$ | 1.19 (1.10–1.27) |
| Th1/Th2 immune response | rs2043136 | 3p24.1 | *TGFBR2* | T,C | 0.239/0.266 | $1.6 \times 10^{-5}$ | 1.18 (1.09–1.27) |
| | rs3773650 | 3p24.1 | *TGFBR2* | C,A | 0.182/0.205 | $7.9 \times 10^{-5}$ | 1.18 (1.09–1.28) |
| Apoptosis | rs1062225 | 10q11.22 | *MAPK8* | A,G | 0.129/0.111 | 0.00063 | 0.84 (0.76–0.93) |
| | rs13396983 | 2q13 | *BCL2L11* | G,A | 0.456/0.484 | 0.00068 | 1.12 (1.05–1.20) |
| | rs2015454 | 2q13 | *BCL2L11* | C,T | 0.451/0.477 | 0.00070 | 1.12 (1.05–1.20) |

The analysis was adjusted for age in 10 years categories, sex, study, arm, ancestry and five principal components of population stratification. CI, confidence interval; OR, odds ratio.
[a]NCBI dbSNP identifier.
[b]Chromosome.
[c]Gene name.
[d]Major allele, minor allele.
[e]Minor allele frequency in control and case participants.
[f]1 d.f. Wald test.
[g]Denotes SNPs reported in the PanScan 1 and 2 GWAS (1,2).

allergies, who are at lower risk of pancreatic cancer, have heightened Th2 (T-helper type 2) response. *TGFBR2* and *CCL18* contribute to the significance of the Th1/Th2 pathway. Although T-helper cells are mostly implicated in diseases associated with immune responses, such as allergy, asthma and infections, they may also play a role in immune surveillance of tumor cells (46). On the other hand, TGF-β is one of the core signaling pathways involved in pancreatic cancer (47), and the *TGFBR2* gene is mutated in 4% of pancreatic cancer cases (48). Chemokines such as CCL18 have been implicated in biological processes involving tumor growth including leukocyte migration, angiogenesis and metastasis (49); CCL18 is associated with some allergic conditions and is induced by Th2 cytokines. However, the role of CCL18 in pancreatic carcinogenesis is unknown. Defective apoptosis represents a contributory feature in the development and progression of cancer. Among the 42 apoptosis-related genes analyzed, *MAPK8* and *BCL2L11* were the most notable. Mitogen-activated protein kinases are involved in cell proliferation, differentiation, apoptosis, transcription regulation and development. MAPK8 (aka JNK1 or SAPK1) is a serine–threonine kinase that belongs to the stress-activated signaling cascade and has been shown to play a role in obesity and insulin resistance (50). *BCL2L11* is a member of the *BCL2* family and plays a role in neuronal and lymphocyte apoptosis.

In summary, our pathway-based association analysis of pancreatic cancer GWAS data has revealed a connection between pancreatic development and cancer risk by using sets of genes previously known to be important for pancreatic cancer through various processes and molecular functions. We use an ARTP method as our primary approach and confirmed the results for the developmental pathway with another approach, logic regression. Our selection of pathways incorporated databases (such as KEGG and GO), however, was narrowed to include only those genes central to each pathway, based on the literature. A more agnostic wider pathway based analysis might elucidate new pathways beyond that which is known. Our study is the largest to date to examine candidate pathways and genes associated with pancreatic cancer. A limitation to our study is that in order to maximize power, all available case–control pairs were used for the analysis. Replication efforts in independent studies are therefore needed to confirm our findings. These findings may open new research avenues in our understanding of the etiology of this deadly malignancy.

## Supplementary material

Supplementary Table 1 can be found at http://carcin.oxfordjournals.org/.

## Acknowledgements

This research was supported by the Intramural Research Program of the National Institutes of Health (NIH), Division of Cancer Epidemiology and Genetics, National Cancer Institute (NCI), National Institutes of Health, Department of Health and Human Services.

E.J.D. supported by Insituto de Salud Carlos III (RETICC, RD06/0020).

The NYU Women's Health Study is supported by research grant (R01CA098661) and center grant (P30CA016087) from the NCI and the center grant (ES000260) from the National Institute of Environmental Health Sciences.

The WHI program is funded by the National Heart, Lung, and Blood Institute, NIH, US Department of Health and Human Services through contracts

N01WH22110, 24152, 32100-2, 32105-6, 32108-9, 32111-13, 32115, 32118-32119, 32122, 42107-26, 42129-32 and 44221.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.whiscience.org/publications/WHI_investigators_shortlist.pdf.

The Mayo Clinic Molecular Epidemiology of Pancreatic Cancer study is supported by the Mayo Clinic SPORE in Pancreatic Cancer (P50 CA102701). The authors would like to acknowledge William Bamlet, Traci Hammer, Jodie Cogswell, Hugues Sicotte, Janet Olson, Martha Matsumoto, and Dennis Robinson.

The Yale University study was supported by grant number (5R01CA098870) from the NCI, NIH. The cooperation of 30 Connecticut hospitals, including Stamford Hospital, in allowing patient access, is gratefully acknowledged. This study was approved by the State of Connecticut Department of Public Health Human Investigation Committee. Certain data used in this study were obtained from the Connecticut Tumor Registry in the Connecticut Department of Public Health. The authors assume full responsibility for analyses and interpretation of these data.

The PHS, NHS, HPFS and WHS at Harvard were supported by the NCI, NIH (grants no. P01 CA87969, P01 CA55075, P50 CA127003, R01 CA124908, RO1 CA97193, RO1 CA34944, RO1 CA40360, RO1 HL26490, RO1 HL34595, RO1 CA047988, RO1 HL043851, RO1 HL080467).

The work at Johns Hopkins University was supported by the NCI (grants P50CA62924 and R01CA97075) and the Lustgarten Foundation for Pancreatic Cancer Research.

The Shanghai Men's Health Study was supported by the NCI extramural research grant (R01 CA82729). The Shanghai Women's Health Study was supported by the NCI extramural research grant (R37 CA70867) and, partially for biological sample collection, by the Intramural Research Program of NCI (Division of Cancer Epidemiology and Genetics). We are in debt to the contributions of Drs Yu-Tang Gao and Yong-Bing Xiang in these two cohort studies. The studies would not be possible without the continuing support and devotion from the study participants and staff of the SMHS and SWHS.

Pancreatic cancer research at Memorial Sloan-Kettering Cancer Center was supported by The Society of MSKCC and by the Geoffrey Beene Cancer Research Fund.

The work at M. D. Anderson was supported by NIH grant (RO1 CA98380).

The UCSF study was supported in part by NCI grants [CA59706, CA108370, CA109767, CA89726 (E.A.H., PI) and CA98889 (E.J.D., PI] and by the Rombauer Pancreatic Cancer Research Fund.

The University of Toronto study was supported by grants from the NIH (R01 CA97075, as part of the PACGENE consortium), The Lustgarten Foundation for Pancreatic Cancer Research and the Ontario Cancer Research Network. We acknowledge the Pancreatic Cancer Canada Foundation (www.pancreaticancercanada.ca) for their continued support of research into the early detection of pancreatic cancer and the Pancreas Cancer Screening Study at Mount Sinai Hospital and Princess Margaret Hospital. The authors acknowledge Ayelet Borgida and Heidi Rothenmund for their dedicated contributions toward data collection and study co-ordination.

PLCO was supported by individual contracts from the NCI to the University of Colorado Denver (NO1-CN-25514), Georgetown University (NO1-CN-25522), Pacific Health Research Institute (NO1-CN-25515), Henry Ford Health System (NO1-CN-25512), University of Minnesota (NO1-CN-25513), Washington University (NO1-CN-25516), University of Pittsburgh (NO1-CN-25511), University of Utah (NO1-CN-25524), Marshfield Clinic Research Foundation (NO1-CN-25518), University of Alabama at Birmingham (NO1-CN-75022), Westat, Inc. (NO1-CN-25476), University of California, Los Angeles (NO1-CN-25404).

The ATBC Study was supported by funding provided by the Intramural Research Program of the NCI, NIH and through US Public Health Service contracts (N01-CN-45165, N01-RC-45035, and N01-RC-37004) from the NCI.

For the EPIC cohorts, all coauthors coordinated the initial recruitment and management of the studies. All authors contributed to the final paper. The authors thank all the participants who took part in this research and the funders and support and technical staff who made this study possible. The work described in this paper was carried out with the support of the European Commission: Public Health and Consumer Protection Directorate 1993–2004; Research Directorate-General 2005–2008.; Ligue contre le Cancer, Societé 3M, Mutuelle Générale de l'Education Nationale, Institut National de la Santé et de la Recherche Médicale (INSERM) (France); German Cancer Aid, German Cancer Research Center, Federal Ministry of Education and Research (Germany); Danish Cancer Society (Denmark); ISCIII RETIC (RD06/0020) of the Spanish Ministry of Health, The participating regional governments and institutions (Spain); Cancer Research UK, Medical Research Council, Stroke Association, British Heart Foundation, Department of Health, Food Standards Agency, the Wellcome Trust (UK); Greek Ministry of Health and Social Solidarity, Hellenic Health Foundation and Stavros Niarchos Foundation (Greece); Italian Association for Research on Cancer (AIRC) (Italy); Dutch Ministry of Public Health, Welfare and Sports, Dutch Prevention Funds, LK Research Funds, Dutch ZON (Zorg Onderzoek Nederland), World Cancer Research Fund (WCRF) (The Netherlands); Swedish Cancer Society, Swedish Scientific Council, Regional Government of Skane and Västerbotten (Sweden).

CLUE II was supported by National Institute of Aging grant (5U01AG018033) and NCI grants (CA105069, CA73790). Cancer incidence data were provided by the Maryland Cancer Registry, Center for Cancer Surveillance and Control, Department of Health and Mental Hygiene, 201 W. Preston Street, Room 400, Baltimore, MD 21201, USA, www.fha.state.md.us/cancer/registry/, 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.

The Cancer Prevention Study II Nutrition Cohort is supported by the American Cancer Society. The authors thank all the men and women in the Cancer Prevention Study II Nutrition Cohort for their many years of dedicated participation in the study.

This project has been funded in whole or in part with federal funds from the NCI, NIH, under Contract No. HHSN261200800001E.

*Conflict of Interest Statement:* The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services nor does mention of trade names, commercial products or organizations imply endorsement by the US Government.

## References

1. Amundadottir,L. *et al*. (2009) Genome-wide association study identifies variants in the ABO locus associated with susceptibility to pancreatic cancer. *Nat. Genet.*, **41**, 986–990.
2. Petersen,G.M. *et al*. (2010) A genome-wide association study identifies pancreatic cancer susceptibility loci on chromosomes 13q22.1, 1q32.1 and 5p.15.33. *Nat. Genet.*, **42**, 224–228.
3. WTCCC. (2007) Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature*, **447**, 661–678.
4. Wang,K. *et al*. (2010) Analysing biological pathways in genome-wide association studies. *Nat. Rev. Genet.*, **11**, 843–854.
5. Cantor,R.M. *et al*. (2010) Prioritizing GWAS results: a review of statistical methods and recommendations for their application. *Am. J. Hum. Genet.*, **86**, 6–22.
6. Yu,K. *et al*. (2009) Pathway analysis by adaptive combination of P-values. *Genet. Epidemiol.*, **33**, 700–709.
7. Ruczinski,I. *et al*. (2003) Logic regression. *J. Comput. Graph. Stat.*, **12**, 475–511.
8. Kooperberg,C. *et al*. (2005) Identifying interacting SNPs using Monte Carlo logic regression. *Genet. Epidemiol.*, **28**, 157–170.
9. Duell,E.J. *et al*. (2002) A population-based study of the Arg399Gln polymorphism in X-ray repair cross- complementing group 1 (XRCC1) and risk of pancreatic adenocarcinoma. *Cancer Res.*, **62**, 4630–4636.
10. Jiao,L. *et al*. (2007) The XPD Asp312Asn and Lys751Gln polymorphisms, corresponding haplotype, and pancreatic cancer risk. *Cancer Lett.*, **245**, 61–68.
11. McWilliams,R.R. *et al*. (2009) Nucleotide excision repair pathway polymorphisms and pancreatic cancer risk: evidence for role of MMS19L. *Cancer Epidemiol. Biomarkers Prev.*, **18**, 1295–1302.
12. Huxley,R. *et al*. (2005) Type-II diabetes and pancreatic cancer: a meta-analysis of 36 studies. *Br. J. Cancer*, **92**, 2076–2083.
13. Stolzenberg-Solomon,R.Z. *et al*. (2005) Insulin, glucose, insulin resistance, and pancreatic cancer in male smokers. *JAMA*, **294**, 2872–2878.

14. Jee,S.H. *et al*. (2005) Fasting serum glucose level and cancer risk in Korean men and women. *JAMA*, **293**, 194–202.
15. Gapstur,S.M. *et al*. (2000) Abnormal glucose metabolism and pancreatic cancer mortality. *JAMA*, **283**, 2552–2558.
16. Lowenfels,A.B. *et al*. (1999) Chronic pancreatitis and other risk factors for pancreatic cancer. *Gastroenterol. Clin. North Am.*, **28**, 673–685.
17. Olson,S.H. (2012) Selected medical conditions and risk of pancreatic cancer. *Mol. Carcinog.*, **51**, 75–97.
18. Loza,M.J. *et al*. (2007) Assembly of inflammation-related genes for pathway-focused genetic analysis. *PLoS One*, **2**, e1035.
19. Zaret,K.S. (2008) Genetic programming of liver and pancreas progenitors: lessons for stem-cell differentiation. *Nat. Rev. Genet.*, **9**, 329–340.
20. Zaret,K.S. *et al*. (2008) Generation and regeneration of cells of the liver and pancreas. *Science*, **322**, 1490–1494.
21. Shih,D.Q. *et al*. (2001) Loss of HNF-1alpha function in mice leads to abnormal expression of genes involved in pancreatic islet development and metabolism. *Diabetes*, **50**, 2472–2480.
22. Fayard,E. *et al*. (2004) LRH-1: an orphan nuclear receptor involved in development, metabolism and steroidogenesis. *Trends Cell Biol.*, **14**, 250–260.
23. Maestro,M.A. *et al*. (2007) Distinct roles of HNF1beta, HNF1alpha, and HNF4alpha in regulating pancreas development, beta-cell function and growth. *Endocr. Dev.*, **12**, 33–45.
24. Martin,M. *et al*. (2007) Transcription factors in pancreatic development. Animal models. *Endocr. Dev.*, **12**, 24–32.
25. Kaneto,H. *et al*. (2008) PDX-1 functions as a master factor in the pancreas. *Front Biosci.*, **13**, 6406–6420.
26. Annicotte,J.S. *et al*. (2003) Pancreatic-duodenal homeobox 1 regulates expression of liver receptor homolog 1 during pancreas development. *Mol. Cell. Biol*, **23**, 6713–6724.
27. Pierce,B.L. *et al*. (2011) Genome-wide ''pleiotropy scan'' identifies HNF1A region as a novel pancreatic cancer susceptibility locus. *Cancer Res.*, **71**, 4352–4358.
28. Haumaitre,C. *et al*. (2005) Lack of TCF2/vHNF1 in mice leads to pancreas agenesis. *Proc. Natl Acad. Sci. USA*, **102**, 1490–1495.
29. Morris,J.P. *et al*. (2010) KRAS, Hedgehog, Wnt and the twisted developmental biology of pancreatic ductal adenocarcinoma. *Nat. Rev. Cancer*, **10**, 683–695.
30. Glucksmann,M.A. *et al*. (1997) Novel mutations and a mutational hotspot in the MODY3 gene. *Diabetes*, **46**, 1081–1086.
31. Carette,C. *et al*. (2007) Exonic duplication of the hepatocyte nuclear factor-1beta gene (transcription factor 2, hepatic) as a cause of maturity onset diabetes of the young type 5. *J. Clin. Endocrinol. Metab.*, **92**, 2844–2847.
32. Voight,B.F. *et al*. (2010) Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. *Nat. Genet.*, **42**, 579–589.
33. Furuta,H. *et al*. (2002) Nonsense and missense mutations in the human hepatocyte nuclear factor-1 beta gene (TCF2) and their relation to type 2 diabetes in Japanese. *J. Clin. Endocrinol. Metab.*, **87**, 3859–3863.
34. Holmkvist,J. *et al*. (2006) Common variants in HNF-1 alpha and risk of type 2 diabetes. *Diabetologia*, **49**, 2882–2891.
35. Speliotes,E.K. *et al*. (2010) Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. *Nat. Genet.*, **42**, 937–948.
36. Lee,J.M. *et al*. (2011) A nuclear-receptor-dependent phosphatidylcholine pathway with antidiabetic effects,. *Nature*, **474**, 506–510.
37. D'Angelo,A. *et al*. (2010) Hepatocyte nuclear factor 1alpha and beta control terminal differentiation and cell fate commitment in the gut epithelium. *Development*, **137**, 1573–1582.
38. Lussier,C.R. *et al*. (2010) Loss of hepatocyte-nuclear-factor-1alpha impacts on adult mouse intestinal epithelial cell growth and cell lineages differentiation. *PLoS One*, **5**, e12378.
39. Laurent-Puig,P. *et al*. (2003) Frequent mutations of hepatocyte nuclear factor 1 in colorectal cancer with microsatellite instability. *Gastroenterology*, **124**, 1311–1314.
40. Rebouissou,S. *et al*. (2004) Mutation of TCF1 encoding hepatocyte nuclear factor 1alpha in gynecological cancer. *Oncogene*, **23**, 7588–7592.
41. Bluteau,O. *et al*. (2002) Bi-allelic inactivation of TCF1 in hepatic adenomas. *Nat. Genet.*, **32**, 312–315.
42. Pelletier,L. *et al*. (2010) Loss of hepatocyte nuclear factor 1alpha function in human hepatocellular adenomas leads to aberrant activation of signaling pathways involved in tumorigenesis. *Hepatology*, **51**, 557–566.
43. Lauc,G. *et al*. (2010) Genomics meets glycomics-the first GWAS study of human N-Glycome identifies HNF1alpha as a master regulator of plasma protein fucosylation. *PLoS Genet.*, **6**, e1001256.
44. Kleber,M.E. *et al*. (2010) Effect of the rs2259816 polymorphism in the HNF1A gene on circulating levels of c-reactive protein and coronary artery disease (the ludwigshafen risk and cardiovascular health study). *BMC Med. Genet.*, **11**, 157.
45. Ley,S.H. *et al*. (2010) Assessing the association of the HNF1A G319S variant with C-reactive protein in Aboriginal Canadians: a population-based epidemiological study. *Cardiovasc. Diabetol.*, **9**, 39.
46. Swann,J.B. *et al*. (2007) Immune surveillance of tumors. *J. Clin. Invest.*, **117**, 1137–1146.
47. Jones,S. *et al*. (2008) Core signaling pathways in human pancreatic cancers revealed by global genomic analyses. *Science*, **321**, 1801–1806.
48. Goggins,M. *et al*. (1998) Genetic alterations of the transforming growth factor beta receptor genes in pancreatic and biliary adenocarcinomas. *Cancer Res.*, **58**, 5329–5332.
49. Bonecchi,R. *et al*. (2011) Chemokines and cancer: a fatal attraction. *Cancer Cell*, **19**, 434–435.
50. Sabio,G. *et al*. (2010) cJun NH2-terminal kinase 1 (JNK1): roles in metabolic regulation of insulin resistance. *Trends Biochem. Sci.*, **35**, 490–496.

*Received January 7, 2012; revised April 2, 2012; accepted March 9, 2012*