Three new pancreatic cancer susceptibility signals identified on chromosomes 1q32.1, 5p15.33 and 8q24.21

The Harvard community has made this article openly available. Please share how this access benefits you. Your story matters.

Citation
Zhang, M., Z. Wang, O. Obazee, J. Jia, E. J. Childs, J. Hoskins, G. Figlioli, et al. 2016. “Three new pancreatic cancer susceptibility signals identified on chromosomes 1q32.1, 5p15.33 and 8q24.21.” Oncotarget 7 (41): 66328-66343. doi:10.18632/oncotarget.11041. http://dx.doi.org/10.18632/oncotarget.11041.

Published Version
doi:10.18632/oncotarget.11041

Citable link
http://nrs.harvard.edu/urn-3:HUL.InstRepos:32072245

Terms of Use
This article was downloaded from Harvard University’s DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA
Three new pancreatic cancer susceptibility signals identified on chromosomes 1q32.1, 5p15.33 and 8q24.21

Mingfeng Zhang1, Zhaoming Wang2,3,100, Ofure Obazee4, Jinping Jia1, Erica J. Childs5, Jason Hoskins1, Gisella Figlioli4, Evelina Mocci5, Irene Collins1, Charles C. Chung2,3, Christopher Hauptman1, Alan A. Arslan6,7,8, Laura Beane-Freeman2, Paige M. Bracci9, Julie Buring10,11, Eric J. Duell12, Steven Gallinger13, Graham G. Giles14,15,16, Gary E. Goodman17, Phyllis J. Goodman18, Aruna Kamineni19, Laurence N. Kolonel20, Matthew H. Kulke21, Núria Malats22, Sara H. Olson23, Howard D. Sesso24,10,11, Kala Visvanathan25, Emily White17,26, Wei Zheng27,28, Christian C. Abnet2, Demetrius Albanes2, Gabriella Andreotti2, Lauren Brais21, H. Bas Bueno-de-Mesquita29,30,31, Daniela Basso32, Sonja I. Berndt2, Marie-Christine Boutron-Ruault33,34,35, Maarten F. Bijlsma36, Hermann Brenner37,38,39, Laurie Burdette2,3, Daniele Campa40, Neil E. Caporaso2, Gabriele Capurso41, Giulia Martina Cavestro42, Michelle Cotterchio43,44, Eithne Costello45, Joanne Elena46, Ugo Boggi47, J. Michael Gaziano10,11,48, Maria Gazouli49, Edward L. Giovannucci50,51, Michael Goggins52,53,54, Myron Grose55, Christopher A. Haiman56, Manal Hassan57, Kathy J. Helzlsouer46, Nan Hu2, David J. Hunter58,59,60, Elizbieta Iskierka-Jazdzewska61, Mazda Jenab62, Rudolf Kaaks63, Timothy J. Key64, Kay-Tee Khaw65, Eric A. Klein66, Manolis Kogevinas67,68,69, Vittorio Krogh70, Juozas Kucinskas71, Robert C. Kurtz72, Maria T. Landi46, Stefano Landi73, Loic Le Marchand20, Andrea Mambrini73, Satu Mannisto74, Roger L. Milne14,15, Rachel E. Neale75, Ann L. Oberg76, Salvatore Panico77, Alpa V. Patel78, Petra H. M. Peeters79,80, Ulrike Peters17,26, Raffaele Pezzilli80, Miquel Porta68,81,82, Mark Purdue2, J. Ramón Quiros83, Elio Riboli30, Nathaniel Rothman2, Aldo Scarpa84, Ghislaine Scelo62, Xiao-Ou Shu27,28, Debra T. Silverman2, Pavel Soucek85, Oliver Strobel86, Malin Sund87, Ewa Malecka-Panas88, Philip R. Taylor2, Francesca Tavano89, Ruth C. Travis64, Mark Thorquist17, Anne Tjønneland90, Geoffrey S. Tobias2, Dimitrios Trichopoulos24,91,92, Yogesh Vashist93, Pavel Vodicca94, Jan Wactawski-Wende95, Nicolas Wentzensen2, Herbert Yu20, Kai Yu2, Anne Zeleniuch-Jacquotte7,8, Charles Kooperberg17, Harvey A. Risch96, Eric J. Jacobs78, Donghui Li97, Charles Fuchs21,50, Robert Hoover2, Patricia Hartge2, Stephen J. Chanock2, Gloria M. Petersen97, Rachael S. Stolzenberg-Solomon2, Brian M. Wolpin21, Peter Kraft24,98, Alison P. Klein5,99, Federico Canzian1 and Laufey T. Amundadottir1

1 Laboratory of Translational Genomics, Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, Maryland, USA
2 Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, Maryland, USA
3 Cancer Genomics Research Laboratory, National Cancer Institute, Division of Cancer Epidemiology and Genetics, Leidos Biomedical Research, Inc., Frederick National Laboratory for Cancer Research, Frederick, Maryland, USA
4 Genomic Epidemiology Group, German Cancer Research Center (DKFZ), Heidelberg, Germany
5 Department of Oncology, the Johns Hopkins University School of Medicine, Baltimore, Maryland, USA
6 Department of Obstetrics and Gynecology, New York University School of Medicine, New York, New York, USA
7 Department of Environmental Medicine, New York University School of Medicine, New York, New York, USA
8 New York University Cancer Institute, New York, New York, USA
9 Department of Epidemiology and Biostatistics, University of California San Francisco, San Francisco, California, USA
10 Division of Preventive Medicine, Department of Medicine, Brigham and Women’s Hospital and Harvard Medical School, Boston, Massachusetts, USA
Department of Basic Medical Sciences, Laboratory of Biology, Medical School, National and Kapodistrian University of Athens, Athens, Greece

Channing Division of Network Medicine, Department of Medicine, Brigham and Women’s Hospital, and Harvard Medical School, Boston, Massachusetts, USA

Department of Nutrition, Harvard School of Public Health, Boston, Massachusetts, USA

Department of Pathology, Sidney Kimmel Cancer Center and Johns Hopkins University, Baltimore, Maryland, USA

Department of Medicine, Sidney Kimmel Cancer Center and Johns Hopkins University, Baltimore, Maryland, USA

Department of Oncology, Sidney Kimmel Cancer Center and Johns Hopkins University, Baltimore, Maryland, USA

Laboratory of Medicine and Pathology, University of Minnesota, Minneapolis, Minnesota, USA

Preventive Medicine, University of Southern California, Los Angeles, California, USA

Department of Gastrointestinal Medical Oncology, University of Texas M.D. Anderson Cancer Center, Houston, Texas, USA

Department of Medicine, Brigham and Women’s Hospital and Harvard Medical School, Boston, Massachusetts, USA

Harvard School of Public Health, Boston, Massachusetts, USA

Harvard Medical School, Boston, Massachusetts, USA

Department of Hematology, Medical University of Łodz, Łodz, Poland

International Agency for Research on Cancer (IARC), Lyon, France

Division of Cancer Epidemiology, German Cancer Research Center (DKFZ), Heidelberg, Germany

Cancer Epidemiology Unit, University of Oxford, Oxford, United Kingdom

School of Clinical Medicine, University of Cambridge, Cambridge, United Kingdom

Glickman Urological and Kidney Institute, Cleveland Clinic, Cleveland, Ohio, USA

Centre de Recerca en Epidemiologia Ambiental (CREAL), CIBER Epidemiología y Salud Pública (CIBERESP), Spain

Hospital del Mar Institute of Medical Research (IMIM), Barcelona, Spain

National School of Public Health, Athens, Greece

Epidemiology and Prevention Unit, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy

Department of Gastroenterology, Lithuanian University of Health Sciences, Kaunas, Lithuania

Department of Medicine, Memorial Sloan-Kettering Cancer Center, New York, New York, USA

Oncology Department, ASL1 Massa Carrara, Massa Carrara, Italy

National Institute for Health and Welfare, Department of Chronic Disease Prevention, Helsinki, Finland

Department of Population Health, QIMR Berghofer Medical Research Institute, Brisbane, Queensland, Australia

Division of Biomedical Statistics and Informatics, Department of Health Sciences Research, Mayo Clinic, Rochester, Minnesota, USA

Dipartimento di Medicina Clinica E Chirurgia, Federico II University, Naples, Italy

Epidemiology Research Program, American Cancer Society, Atlanta, Georgia, USA

Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, The Netherlands

Pancreas Unit, Department of Digestive Diseases and Internal Medicine, Sant’Orsola-Malpighi Hospital, Bologna, Italy

School of Medicine, Universitat Autònoma de Barcelona, Barcelona, Spain

CIBER de Epidemiología y Salud Pública (CIBERESP), Madrid, Spain

Public Health and Participation Directorate, Asturias, Spain

ARC-NET: Centre for Applied Research on Cancer, University and Hospital Trust of Verona, Verona, Italy

Laboratory of Pharmacogenomics, Biomedical Center, Faculty of Medicine in Pilsen, Charles University in Prague, Pilsen, Czech Republic

Department of General Surgery, University Hospital Heidelberg, Heidelberg, Germany

Department of Surgical and Per operative Sciences, Umeå University, Umeå, Sweden

Department of Digestive Tract Diseases, Medical University of Łodz, Łodz, Poland

Division of Gastroenterology and Research Laboratory, IRCCS Scientific Institute and Regional General Hospital “Casa Sollievo della Sofferenza”, San Giovanni Rotondo, Italy

Institute of Cancer Epidemiology, Danish Cancer Society, Copenhagen, Denmark

Bureau of Epidemiologic Research, Academy of Athens, Athens, Greece
INTRODUCTION

Although relatively rare, pancreatic tumors are highly lethal. Over 80% of patients present with advanced disease at the time of diagnosis and the five year survival is only 7% [1, 2]. This disease is currently the third leading cause of cancer deaths in the United States, sixth in Europe and seventh worldwide [3-5]. In contrast to most other cancers, mortality rates for pancreatic cancer are not improving [6, 7]. In the U.S., it is predicted to become the second leading cause of cancer-related deaths by 2030 [7]. Pancreatic cancer risk has been associated with smoking, obesity, diabetes and pancreatitis [8]. A small fraction of the familial aggregation of pancreatic cancer can be accounted for by rare, moderately or highly penetrant mutations [9]. Furthermore, genome-wide association studies (GWAS) have identified common variants at 13 loci associated with risk of pancreatic cancer in European populations and at 5 loci in Asian populations (at the GWAS threshold of \( P < 5.0 \times 10^{-8} \)), or a total of 18 loci [10-15].

ABSTRACT

Genome-wide association studies (GWAS) have identified common pancreatic cancer susceptibility variants at 13 chromosomal loci in individuals of European descent. To identify new susceptibility variants, we performed imputation based on 1000 Genomes (1000G) Project data and association analysis using 5,107 case and 8,845 control subjects from 27 cohort and case-control studies that participated in the PanScan I-III GWAS. This analysis, in combination with a two-staged replication in an additional 6,076 case and 7,555 control subjects from the PANDoRA and Pancreatic Cancer Case-Control (PanC4) Consortia uncovered 3 new pancreatic cancer risk signals marked by single nucleotide polymorphisms (SNPs) rs2816938 at chromosome 1q32.1 (per allele odds ratio (OR) = 1.20, \( P = 4.88 \times 10^{-15} \)), rs10094872 at 8q24.21 (OR = 1.15, \( P = 3.22 \times 10^{-9} \)) and rs35226131 at 5p15.33 (OR = 0.71, \( P = 1.70 \times 10^{-8} \)). These SNPs represent independent risk variants at previously identified pancreatic cancer risk loci on chr1q32.1 (NR5A2), chr8q24.21 (MYC) and chr5p15.33 (CLPTM1L-TERT) as per analyses conditioned on previously reported susceptibility variants. We assessed expression of candidate genes at the three risk loci in histologically normal (n = 10) and tumor (n = 8) derived pancreatic tissue samples and observed a marked reduction of NR5A2 expression (chr1q32.1) in the tumors (fold change -7.6, \( P = 5.7 \times 10^{-8} \)). This finding was validated in a second set of paired (n = 20) histologically normal and tumor derived pancreatic tissue samples (average fold change for three NR5A2 isoforms -31.3 to -95.7, \( P = 7.5 \times 10^{-4} - 2.0 \times 10^{-3} \)). Our study has identified new susceptibility variants independently conferring pancreatic cancer risk that merit functional follow-up to identify target genes and explain the underlying biology.
risk loci [16]. To discover additional pancreatic cancer susceptibility loci for individuals of European ancestry, we imputed three GWAS datasets including a total of 5,107 cases and 8,845 controls (PanScan I-III, Stage I) [12]. For replication of promising signals, we first genotyped an additional 1,912 cases and 3,763 controls (PANDoRA; Replication 1), and then further assessed promising signals in a second set of 4,164 cases and 3,792 controls (PanC4; Replication 2). We identified three new susceptibility signals that achieved genome-wide significance for pancreatic cancer risk.

RESULTS

We conducted imputation of three published pancreatic cancer GWAS datasets performed in individuals of European ancestry, PanScan I, II and III [10-12] using the 1000G (Phase 1, version 3) reference dataset [17]. We included 9,132,527 genotyped or imputed SNPs with an imputation information (INFO) score > 0.5 and minor allele frequency (MAF) > 0.01, and performed a fixed effects meta-analysis to combine association results for a total of 5,107 pancreatic cancer cases and 8,845 control subjects [10-12]. Little evidence of systematic inflation due to population stratification was observed (θ = 1.02 for PanScan I+II and θ = 1.07 for PanScan III). We attempted replication of promising findings in two stages. In the first replication stage, we genotyped 15 promising variants in 1,912 pancreatic cancer cases and 3,763 control subjects from the PANcreatic Disease ReseArch (PANDoRA) consortium, a case-control consortium including studies from eight European countries [18]. In the second replication stage, we assessed the three most significant variants based on the meta-analyses results for PanScan I+II, PanScan III and PANDoRA using 4,164 pancreatic cancer cases and 3,792 controls from the Pancreatic Cancer Case-Control Consortium (PanC4), including studies from the U.S., Canada, Europe and Australia [15]. In total, the discovery and replication stages included 11,183 cases and 16,400 controls (Supplementary Table 1).

In the meta-analysis of PanScan I-III (Stage I), two new variants were identified at genome-wide significance (P < 5.0x10^-8), one on chromosome 1q32.1 (rs2816938: P = 1.71x10^-10, OR = 1.23 95% CI 1.15-1.31) and one on 8q24.21 (rs10094872, P = 3.55x10^-4, OR = 1.18 95% CI 1.11-1.25 (Table 1, Supplementary Table 2). After adjusting the analysis on 1q32.1 for the previously reported GWAS SNP rs3790844 (r^2 = 0.097 in 1000G EUR population) [11], the association for rs2816938 remained statistically significant (P = 3.06x10^-4, OR = 1.17). This was also true for the signal at 8q24.21, marked by rs10094872, after adjusting for the GWAS SNP rs1561927 (r^2 = 0.01 in 1000G EUR) [12] (P = 1.09x10^-7, OR = 1.16). The signal at 1q32.1 is located ~11 kb upstream of NR5A2, a gene that encodes a nuclear transcription factor known to play important roles in multiple aspects of pancreatic development and function [19, 20]. The SNP at 8q24.21 is located ~28 kb upstream of MYC, in a susceptibility locus previously reported for bladder cancer (tagged by rs9642880; r^2 = 0.64 in 1000G EUR) [21-23] and ~850 kb upstream of a previously reported pancreatic cancer susceptibility locus [12].

A total of 15 promising variants (P < 5.0x10^-8) were selected for replication in 1,912 pancreatic cancer cases and 3,763 control subjects from the European PANDoRA case-control consortium [18]. After a meta-analysis of PanScan I, II and III and PANDoRA results, the three most promising variants (Supplemental Table 2) were carried forward to replication in PanC4 [15]. The meta-analysis of PanScan I-III with PANDoRA and PanC4 confirmed the signals on chr1q32.1 (P = 4.88x10^-10, OR = 1.20 95% CI 1.15-1.25) and chr8q24.21 (P = 3.22x10^-8, OR = 1.15 95% CI 1.10-1.20). In addition, a new signal in the multicancer region on chr5p15.33, upstream of TERT, was detected (rs35226131, P = 1.70x10^-8, OR = 0.71 95% CI 0.63-0.80) (Table 1). After conditioning the analysis in PanScan I, II and III on the two reported pancreatic cancer susceptibility loci at 5p15.33, rs36115365 [24] (tagging the fine-mapped signal for rs401681 [11] in CLPTM1L) and rs2736098 [12] (tagging the signal in TERT), the signal was still near GWAS significant (P = 1.80x10^-7, OR = 0.64; P = 3.43x10^-5, OR = 0.66). This SNP (rs35226131) is located ~200bps upstream of the transcriptional start site (TSS) of TERT, and marks the least common of the three loci with a MAF of 0.036 in the 1000G EUR populations. The LD between rs35226131 and the previously reported signals is low (r^2 = 0.009 for rs36115365 and r^2 = 0.011 for rs2736098) in 1000G EUR.

Bioinformatic analysis of susceptibility alleles and differential expression analysis

In order to take the first steps towards understanding the functional ramifications at these three new susceptibility signals, we conducted in silico bioinformatic analyses using HaploReg and RegulomeDB [25, 26]. Supporting evidence for putative regulatory function on gene expression was seen for the three loci, particularly for chr1q32.1 and 5p15.33, with open chromatin, modified histones and transcription factor binding in multiple tissues, including those derived from the pancreas and other gastrointestinal tissues (Supplemental Table 3). At chr5p15.33, one of the four variants highly correlated with rs35226131 is a missense variant in the second exon of TERT (rs61748181: r^2 = 1, D’ = 1 in 1000G EUR) whereby the minor allele, associated with reduced risk of pancreatic cancer, changes amino acid 279 from alanine to threonine (A279T).

We assessed expression quantitative trait loci (eQTL) for the three variants in GTEx [27] and the
nearest genes (NR5A2 for 1q32.1; TERT and CLPTM1L for 5p15.33; MYC and PVT1 for 8q24.21; Supplementary Table 4) in histologically normal post-mortem pancreatic tissue samples (n = 149). The minor allele of the SNP that marks the signal on 8q24.21 (rs10094872) and was associated with increased risk of pancreatic cancer, was associated with decreased PVT1 expression (β = -0.23, P = 0.0053) (Supplemental Figure 1). Nominally significant eQTLs (P < 0.05) were not seen for MYC (P = 0.29) or for the SNPs on chr5p15.33 (P = 0.91 for CLPTM1L; TERT was not expressed) or 1q32.1 (P = 0.68 for NR5A2). We furthermore assessed differential expression of the same genes in pancreatic cell lines (n = 9) and pancreatic ductal adenocarcinoma (PDAC) samples (n = 8) as compared with histologically normal pancreatic tissue samples (n = 10) by RNAseq [28] (Supplemental Tables 5-7). The most notable differences were seen for NR5A2 on chr1q32.1 where mRNA expression was markedly lower in pancreatic tumor samples (average fold change -7.6, P = 5.7x10^-14) and cell lines (average fold change -32.7, P = 1.5x10^-34) than in histologically normal pancreatic tissue samples (Supplemental Table 7). We further validated this finding in an independent set of paired histologically normal and tumor derived (PDAC) pancreatic tissue samples from 20 individuals by RT-qPCR for three NR5A2 isoforms and noted decreased expression in tumors for the majority of pairs (average fold change for paired samples was -78.5 for isoform 1, P = 2.0x10^-3; -95.7 fold for isoform 2, P = 7.5x10^-4; -31.3 fold for isoform 3, P = 1.5x10^-3) (Figure 2).

Results are shown from an unconditional logistic regression of the genotypes generated in PanScan I, II and III as well as the two replication studies. aClosest RefSeq gene(s). bPosition of SNP in NCBI genome build 37 (Hg19). cMinor and major alleles. dQuality of imputation metric. eMinor allele frequencies (MAF) are not listed for meta-analysis results. f1 d.f. score test; Chr: chromosome and band; OR, per-allele OR for the minor allele adjusted for for age, sex, study, arm and significant principal components for PanScan I+II; per-allele OR adjusted for age, sex, geographic region and significant principal components for PanScan III; per-allele OR adjusted for age, sex and study for PANDoRA; per-allele OR adjusted for age, sex, study and significant principal components for PanC4. The number of case and control subjects in the combined set of PanScan I, II, III, PANDoRA and PanC4 were: rs2816938 (11,158/16,343), rs10094872 (9,269/12,635) and rs35226131 (11,143/16,308). Text in bold indicates the combined meta-analysis results. NA: Note that the TaqMan assay for rs10094872 on chr8q24.21 failed manufacturing and was therefore not attempted in the PANDoRA samples.

To assess imputation quality, we performed TaqMan genotyping in 678 samples from PanScan I and III (see Materials and Methods). The correlation (r^2) between the imputed genotypes and those measured by TaqMan was 0.98 for rs2816938 (1q32.1), 0.90 for rs10094872 (8q24.21) and 0.37 for rs35226131 (5p15.33). Due to the lower correlation between imputed and directly assayed genotypes for rs35226131, we performed a second validation in an additional 875 samples, including both rs35226131 as well as the perfectly correlated coding SNP on 5p15.33 mentioned above, rs61748181. The imputed-genotyped r^2 for rs35226131 improved to 0.44 in the second validation set, and was 0.66 (0.61-0.72) for rs61748181. Genotype concordance for the most likely imputed genotypes and directly assayed genotypes (see Materials and Methods) for rs35226131 improved from 86.4% in the first validation set to 94.4% in the second set, and was 86.4% in the first validation set to 94.4% in the second set, and was

Table 1: Association results for three new pancreatic cancer susceptibility signals

| Chr | Nearest gene(s) | SNP  | Position(a) | Alleles(c) | INFO(d) | Stage | Allelic OR (95% CI) | Minor allele frequency(e) | p(f) | Conditional analysis |
|-----|-----------------|------|-------------|------------|---------|-------|---------------------|--------------------------|-----|---------------------|
| 1q32.1 | NR5A2 | rs2816938 | 199,985,368 | A | T | 1.00 PanScan I-II | 1.25 (1.16-1.35) | 0.216 0.255 | 1.57 x 10^-9 | 3.06 x 10^-8 | 1.17 (1.09-1.25) |
| 1.00 PanScan III | 1.19 (0.96-1.43) | 0.231 | 2.21 x 10^-10 | |
| Stage I | 1.23 (1.15-1.31) | |
| PANDoRA | 1.18 (1.07-1.30) | 0.204 | 0.226 | 1.12 x 10^-6 |
| 0.78 PanC4 | 1.16 (1.07-1.26) | 0.167 | 0.189 | 5.77 x 10^-7 |
| Combined | 1.20 (1.15-1.25) | 4.88 x 10^-10 |

8q24.21 | MYC | rs10094872 | 128719884 | T | A | 0.93 PanScan I-II | 1.19 (1.11-1.28) | 0.350 | 0.389 | 9.49 x 10^-6 |
| 0.94 PanScan III | 1.14 (1.03-1.27) | 0.345 | 0.367 | 2.97 x 10^-5 |
| Stage I | 1.18 (1.11-1.25) | |
| PANDoRA | NA | NA | NA | 3.55 x 10^-4 |
| 0.96 PanC4 | 1.10 (1.03-1.18) | 0.259 | 0.382 | 3.87 x 10^-7 |
| Combined | 1.15 (1.10-1.20) | 3.22 x 10^-9 |

5p15.33 | TERT, CLPTM1L | rs35226131 | 1295373 | T | C | 0.73 PanScan I-II | 0.63 (0.52-0.78) | 0.641 | 0.029 | 1.61 x 10^-3 |
| 0.79 PanScan III | 0.65 (0.49-0.87) | 0.636 | 0.028 | 3.27 x 10^-7 |
| Stage I | 0.64 (0.54-0.76) | |
| PANDoRA | 0.83 (0.65-1.05) | 0.633 | 0.026 | 0.125 |
| 0.87 PanC4 | 0.75 (0.60-0.93) | 0.623 | 0.017 | 1.01 x 10^-3 |
| Combined | 0.71 (0.63-0.80) | 1.70 x 10^-4 |

Technical validation of imputed SNPs

www.impactjournals.com/oncotarget 66333 Oncotarget
96.2% for rs61748181 in the second set. Since rs61748181 was directly genotyped in one of the replication studies (PanC4), we performed a meta-analysis of PanScan I-III (OR = 0.62 95% CI 0.52-0.75, \( P = 5.37 \times 10^{-7} \)) and PanC4 (OR = 0.67 95% CI 0.55-0.85, \( P = 1.91 \times 10^{-4} \)) data. This revealed a highly significant association for rs61748181 (OR = 0.64 95% CI 0.56-0.74, \( P = 1.28 \times 10^{-10} \)) that was stronger than that seen for rs35262131 (OR = 0.68 95% CI 0.59-0.79, \( P = 1.91 \times 10^{-4} \)). These results suggest that the association results at this potentially new pancreatic cancer risk locus are reliable.

**DISCUSSION**

In this study, we performed imputation across three pancreatic cancer GWAS datasets, namely PanScan I, II and III [10-12], using 1000G reference data. Through replication of promising variants in individuals from two independent pancreatic cancer case-control consortia, PANDoRA and PanC4, we identified three new GWAS significant risk signals for pancreatic cancer. They are independent signals in previously established pancreatic cancer risk loci on chromosomes 1q32.1, 5p15.33 and 8q24.21, as per conditional analysis, supporting their importance for pancreatic cancer risk.

The signal on 1q32.1 is located in NR5A2, a gene that encodes nuclear receptor subfamily 5 group A member 2 (NR5A2), a transcription factor important for pancreatic development and adult function in the pancreas, liver, intestine and ovary, where it regulates cholesterol synthesis, bile acid homeostasis and steroidogenesis [19, 20]. NR5A2 is an important regulator of exocrine function in the adult pancreas where it maintains homeostasis and promotes regeneration of acinar cells after inflammation caused by chemically induced pancreatitis, and protects

---

Figure 1: Regional plots for three signals associated with pancreatic cancer risk. The \(-\log_{10}(P\ value)\) (Y left axis) for Stage I (PanScan I-III, in gray), PANDoRA (purple) and PanC4 (light blue) was plotted on the genomic coordinates (X axis; NCBI genome build 37). Superimposed blue lines depicts likelihood ratio statistics (right Y axis) estimating putative recombination hotspots in the region. This was performed 5 times, each time utilizing 100 random EUR samples from the 1000 Genomes population (EUR, \( n = 503 \)) (Y right axis). The combined data for Stage I, PANDoRA and PanC4 for three regions: panel A, rs2816938 (1q32.1), panel B, rs10094872 (8q24.21), and panel C: rs3226131 (5p13.33) are shown in red.
the pancreas from \textit{KRAS} driven pre-neoplastic changes \cite{29-31}. Other studies have indicated a growth inducing role for \textit{NR5A2} in pancreatic cancer \cite{32, 33}. Highly correlated variants (\(r^2>0.7\)) span \(\sim 25\) kb on chr1q32.1 from \(\sim 11\) kb upstream of the TSS to within the second intron of the gene. We observed significantly lower mRNA expression of \textit{NR5A2} in the majority of pancreatic tumors and cell lines tested compared with histologically normal pancreatic tissue samples, indicating a possible role for reduced \textit{NR5A2} expression in pancreatic cancer. Although an expression QTL was not observed in GTEx data, the relationship between the two currently known pancreatic cancer risk loci on 1q32.1 and \textit{NR5A2} expression remains to be studied in greater detail.

The tag SNP on 8q24.21 is located \(\sim 28\) kb upstream of \textit{MYC} at an established bladder cancer risk locus \cite{21-23} that is \(\sim 850\) kb upstream of our previously reported pancreatic cancer susceptibility locus \cite{12}. Multiple independent susceptibility loci on 8q24.21, distributed over a \(2\) Mb region, are known to influence risk of bladder, breast, prostate, colorectal, lung, ovarian, pancreatic, renal cancer, glioma and chronic lymphocytic leukemia (CLL) \cite{34-38}. Deregulated expression of \textit{MYC}, a transcription factor that regulates multiple aspects of cell growth and proliferation, occurs in a broad range of human tumors \cite{39}. Although the proximity of rs10094872 to \textit{MYC} indicates that it may be the most likely target gene, 8q24.21 is known for long range chromosomal interactions, and additional candidate genes, including \textit{PVT1} (183 kb), \textit{POUSF1B} (290 kb), \textit{CCAT2} (305 kb) and \textit{MIR1205-MIR1208} (253-442 kb), could be involved \cite{40-44}. Several of the 8q24.21 risk loci interact with the \textit{MYC} and \textit{PVT1} promoters through long range chromosomal interaction, and allele-specific effects on gene expression have been reported for both genes \cite{42, 45}. An expression QTL for \textit{MYC} has been described for the bladder cancer susceptibility locus 8q24.21.

![Figure 2: Expression of three NR5A2 isoforms in paired histologically normal and tumor derived pancreatic tissue samples. Blue bars indicate expression in histologically normal samples and red in tumor derived samples. A. NR5A2 isoform 1, B. NR5A2 isoform 2 and C. NR5A2 isoform 3. Note that no expression was seen for isoform 1 in either the normal or tumor derived sample from Subject 3. Error bars represent standard deviation from four replicates.](image)
risk locus in histologically normal bladder samples from Chinese subjects, albeit from a very small set [46], but not in adipose or blood tissue samples from European subjects [22]. We noted an eQTL for rs10094872 and \textit{PVT1} expression in pancreatic tissue samples in GTEx, indicating that \textit{PVT1} may be a target gene for this locus. Replication of these findings is required in independent sample sets. \textit{PVT1} encodes a long noncoding RNA that is often amplified and upregulated along with \textit{MYC} across multiple cancers. Recently, it has been shown to increase \textit{MYC} protein levels and potentiate its activity [47]. In pancreatic cancer, \textit{PVT1} expression is associated with gemcitabine sensitivity in human pancreatic cancer cells and may be associated with poor prognosis [47-49].

The signal on chr5p15.33 lies in another multicancer susceptibility region reported by GWAS for bladder cancer, breast cancer, chronic lymphocytic leukemia, glioma, lung cancer, melanoma, non-melanoma skin cancer, ovarian cancer, pancreatic cancer, prostate cancer and testicular germ cell cancer [11, 12, 23, 37, 50-60]. For the 6 independent susceptibility loci that have been identified in the \textit{TERT-CLPTM1L} gene region, the same alleles are associated with an increased risk for some cancers but decreased risk of others [24, 60, 61]. Two independent pancreatic cancer susceptibility loci have previously been identified on chr5p15.33 through GWAS [11, 12, 24]. The first one, described in PanScan II [11] was marked by an intronic SNP (rs401681) in \textit{CLPTM1L} that has since been fine-mapped to rs451360 (and a set of highly correlated variants including rs36115365) [24]. A second independent signal on 5p15.33 was identified in PanScan III, tagged by a synonymous SNP (rs2736098) in the second exon of \textit{TERT} [12]. Recently, a third risk locus, marked by rs2853677, was identified in this genomic region through a candidate gene analysis of the \textit{TERT} and \textit{TERC} genes [62]; however this variant did not attain GWAS significance in our study (PanScan I-III, \(P = 4.2\times10^{-4}\)). The \textit{TERT} gene encodes the catalytic subunit of telomerase, known for its critical role in maintaining telomere ends and the increased telomerase activity frequently seen in human cancers [63-65]. Telomere-independent functions for TERT include regulation of gene expression, cell survival, epithelial to mesenchymal transition (EMT) and mitochondrial function [66]. The neighboring gene encodes cleft lip and palate associated transmembrane 1 like (CLPTM1L) protein that promotes growth and survival in pancreatic and lung cancer, respectively, and is overexpressed in some cancers [67-69]. The SNP (rs35226131) that marks the new risk signal on 5p15.33 reported here, and highly correlated variants, are located in the \textit{TERT} promoter (~200-500 bp upstream of the TSS) and could potentially influence its expression. Additionally, it is perfectly correlated with a nonsynonomous variant in \textit{TERT} (rs61748181, A279T) that was recently reported as a novel lung adenocarcinoma risk locus by deep sequencing and direct genotyping of 5,164 cases and 5,716 controls of European ancestry [70]. The threonine substitution at this amino acid in TERT negatively influences telomere length and proliferation in esophageal cancer cell lines compared with alanine, and leads to reduced Wnt signaling through destabilization of complexes containing TERT, transcription activator BRG-1 and \(\beta\)-catenin [71]. As the TERT-279T variant is protective for pancreatic cancer in our study, and for lung cancer [70], the underlying mechanism at this locus may relate to increased TERT activity via canonical and/or non-canonical TERT pathways. This hypothesis needs to be formally investigated by future molecular studies.

In conclusion, through imputation of three existing GWAS datasets and replication in two independent case-control consortia, we identified three new susceptibility signals for pancreatic cancer in populations of European ancestry. They are located in genomic regions previously reported by GWAS of pancreatic cancer, further supporting their importance for pancreatic cancer risk. Further work is required to identify target genes and explain the underlying biological mechanisms.

**MATERIALS AND METHODS**

**Study participants**

Participants were drawn from the Pancreatic Cancer Cohort Consortium and the Pancreatic Cancer Case-Control Consortium (PanC4) and include individuals from 17 cohort and 11 case-control studies genotyped in three previous GWAS phases, namely PanScan I, PanScan II and PanScan III [10-12]. Two replication cohorts were included, the PANDoRA consortium [18] (Replication I) and the Pancreatic Cancer Case-Control Consortium (PanC4) [15] (Replication 2). Cases were defined as individuals diagnosed with adenocarcinoma of the pancreas.

Each study obtained informed consent from study participants and approval from its Institutional Review Board (IRB) including IRB certification permitting data sharing in accordance with the NIH Policy for Sharing of Data Obtained in NIH Supported or Conducted Genome-Wide Association Studies (GWAS). The PanScan and PanC4 GWAS data are available through dbGAP (accession numbers phs000206.v5.p3 and phs0006548.v1.p1, respectively).

**Genotyping, imputation and association analysis**

GWAS genotyping was performed at the Cancer Genomics Research Laboratory (CGR) of the National Cancer Institute (NCI) of the National Institutes of Health (NIH) using the Illumina HumanHap series arrays (Illumina HumanHap550 Infinium II, Human 610-Quad)
for PanScan I-II, and the Illumina Omni series arrays (OmniExpress, Omni1M, Omni2.5 and Omni5M) for PanScan III [10-12]. The 1000 Genomes (1000G) Phase 1, Release 3 [17] reference dataset was used to impute the PanScan I-III GWAS datasets using IMPUTE2 [72] as previously described [12, 24]. Due to the large overlap of variants on genotyping arrays for PanScan I and II, these datasets were imputed and analyzed together. The PanScan III data was imputed and analyzed separately. For quality control, variants were excluded based on: 1) completion rate < 90%; 2) MAF < 0.01; 3) Hardy-Weinberg Proportion P value < 1x10^-6; 4) low quality imputation score (IMPUTE 2 INFO score < 0.5). After quality control, 9,132,527 SNPs in 5,107 pancreatic cancer cases and 8,845 controls of European ancestry were included in the analysis. The association analysis was performed using SNPTST [73] based on probabilistic genotypes from IMPUTE2 [72] using the same adjustments for study, geographical region, age, sex and population substructure as were used in PanScan [10-12]. The score test of the log additive genetic effect was used. A meta-analysis of data from PanScan I & II with PanScan III was performed using the fixed-effects inverse-variance method based on β estimates and standard errors. Heterogeneity was not observed for the SNPs identified as GWAS significant or suggestive in the combined study (P heterogeneity ≥0.30).

The estimated inflation of the test statistic, λ, was 1.02 for PanScan I+II and 1.07 for PanScan III, respectively (using variants with MAF>0.01 and INFO>0.5) [74].

**Replication**

Fifteen variants giving promising signals (P < 5.0x10^-5) were selected for replication in the PANDoRA consortium (Replication 1) [18]. Genotyping was performed by custom TaqMan genotyping assays (Applied Biosystems) at the German Cancer Research Center (DKFZ) in Heidelberg, Germany in 3,343 pancreatic cancer cases and 4,998 controls, of which 2,820 cases and 3,909 controls had complete demographic and clinical data and did not overlap with other study samples. Duplicate quality control samples (n = 541 pairs) showed 99.67% genotype concordance. Samples on a few plates were not genotyped for all variants. Unfortunately these plates contained more cases than controls. We excluded 908 cases and 146 controls, either with low genotyping completion rate (< 80%) or not genotyped, resulting in a total of 1,912 cases and 3,763 controls in the final analyses. The association analysis for PANDoRA was adjusted for age, gender and study in the same manner as previously described [12].

Three variants from the meta-analysis of PanScan and PANDoRA were then selected for a second replication in the Pancreatic Cancer Case-Control Consortium (PanC4) [15] (Replication 2). Genotyping for PanC4 had previously been performed at the Johns Hopkins Center for Inherited Disease Research (CIDR) using the IlluminaHumanOmniExpressExome-8v1 array followed by imputation using 100G Phase 3, version 1 [75] and IMPUTE2. Association analysis was performed in 4,164 pancreatic cancer cases and 3,792 control subjects of European ancestry as previously described [15]. Variants at 3 chromosomal locations were extracted from the results and meta-analyses conducted as described above. Heterogeneity between studies was assessed using the Cochran’s Q-test. IMPUTE2 information scores were 0.78 (rs2816938), 0.96 (rs10094872) and 0.87 (rs35226131) for the three reported variants.

Recombination hotspots for association plots were generated as previously described using SequenceLDhot (10-12). The recombination hotspot inference was performed 5 times, each time utilizing a hundred random sampled EUR from the 1000 Genomes (EUR, n = 503). The linkage disequilibrium heatmap was prepared using the 1000 Genomes Phase 3 EUR data (n = 505, CEU+FIN+GBR+IBS+TSI), and snp.plotter R software package [76].

**Validation of imputation accuracy**

Imputation accuracy was assessed by direct TaqMan genotyping or Sanger sequencing. TaqMan genotyping assays (ABI, Foster City, CA) were optimized for three SNPs (rs2816938 on 1q32.1, rs35226131 on 5p15.33 and rs10094872 on 8q24.21) in the independent regions. In an analysis of 678 samples from PanScan I and III [10, 12, 77], the allelic r^2 measured between imputed and assayed genotypes [78] were 0.98, 0.37 and 0.90, respectively. A second validation in an additional 875 samples from PanScan I included two perfectly correlated SNPs on 5p15.33, rs35226131 and rs61748181; the allelic r^2 in this set was 0.44 and 0.55, respectively. We also assessed concordance between the most likely imputed genotypes and directly measured genotypes as follows: samples with imputed allelic dosage ranging from 0-0.5 were designated as being of the homozygous common genotype; samples ranging from 0.51-1.5 as being of the heterozygous genotype and samples ranging from 1.51-2.0 as being of the rare homozygous genotype.

**Bioinformatic analysis of functional potential**

HaploReg v4.1 (http://www.broadinstitute.org/mammals/haploreg/haploreg.php) was used to annotate functional and regulatory potential of the most significant germline variants at each locus as well as highly correlated variants (r^2>0.7) that mark each of the regions identified [25]. RegulomeDB (http://regulomedb.
Analysis of gene expression

Gene expression was assessed for five genes that are closest to the reported variants on chromosomes 1q32.1 (NR5A2), 5p15.33 (TERT and CLPTM1L), and 8q24.21 (MYC and PVT1). We first assessed differential expression of these genes in pancreatic tumor samples (PDAC, n = 8), histologically normal (non-malignant) pancreatic tissue samples (n = 10), and pancreatic cell lines (n = 9) by RNA-sequencing as described previously [28]. We compared gene expression in tumors and cell lines to histologically normal pancreatic tissue samples by EdgeR analysis [28]. P-values represent an exact test of the differential expression of each gene in histologically normal and tumor derived samples using normalized read counts in EdgeR.

We also assessed the expression of one of these genes (NR5A2) in a second independent tissue sample set that included 20 fresh frozen paired histologically normal pancreatic samples (adjacent to tumor) and pancreatic ductal adenocarcinoma (PDAC) tumor samples. RNA samples were isolated from fresh frozen tissues and reverse transcribed to cDNA as previously described [28]. Three NR5A2 isoforms (isoform 1: NM_205860, isoform 2: NM_003822, and isoform 3: uc009wzh.3) were tested using TaqMan gene expression assays (Thermo Fisher Scientific, isoform 1: Hs00892375_m1, isoform 2: Hs00892375_m1 and a custom assay, for isoform 3 forward primer: 5’CTTTTTGCAGGAGTTGAAT3’, reverse primer: 5’GTCCCGGAAGCCACAGCA3’, probe: 5’ CTGTGCTGCCCGTGTCC3’) and a 7900HT system (ABI). Each reaction was run in triplicate and analyzed according to the ΔΔCt method using GAPDH (Hs99999907_m1), G3PDH (Hs99999905_m1) and PPIA (Hs99999904_m1) as housekeeping genes. P-values represent two-sided T-tests of the difference in expression between histologically normal and tumor derived samples.

All tissue samples were obtained from the Mayo Clinic in Rochester, MN. The project was approved by the Institutional Review Boards of Mayo Clinic in Rochester, MN. The project was approved by the Mayo Clinic and the NIH. Nine pancreatic cancer cell lines (AsPC-1, BxPC-3, Hs766T, SU.86.86, SW1990, CFPAC-1, Capan-1, PANC-1, MIA PaCa-2) were purchased from ATCC and cultured as recommended (http://www.ATCC.com). The cell lines were tested for authentication with a panel of short tandem repeats (STR) using the Identifiler kit (Life Technologies) and compared with the ATCC and the DSMZ (German Collection of Microorganisms and Cell Cultures) STR Profile Databases. All cell lines matched the listed profiles.

ACKNOWLEDGMENTS

The authors acknowledge the contribution of the staff of the Cancer Genomics Research Laboratory (CGR) at the National Cancer Institute, NIH, for their help throughout the project. This work was supported by the Intramural Research Program of the US National Institutes of Health (NIH), National Cancer Institute. 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 U.S. Government. Additional acknowledgements for individual participating studies are listed in the Supplemental Materials.

CONFLICTS OF INTEREST

The authors do not have any potential conflicts of interest.

REFERENCES

1. Society AC. Cancer Facts and Figures. Int J BiolSci. 2016; 12:273-282.
2. Vincent A, Herman J, Schulick R, Hruban RH and Goggins M. Pancreatic cancer. Lancet. 2011; 378:607-620.
3. Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J and Jemal A. Global cancer statistics, 2012. CA Cancer J Clin. 2015; 65:87-108.
4. Ferlay J, Steliarova-Foucher E, Lortet-Tieulent J, Rosso S, Coebergh JW, Comber H, Forman D and Bray F. Cancer incidence and mortality patterns in Europe: estimates for 40 countries in 2012. Eur J Cancer. 2013; 49:1374-1403.
5. Siegel RL, Miller KD and Jemal A. Cancer statistics, 2016. CA Cancer J Clin. 2016; 66:7-30.
6. Malvezzi M, Bertuccio P, Levi F, La Vecchia C and Negri E. European cancer mortality predictions for the year 2014. Annals of oncology. 2014; 25:1650-1656.
7. Rahib L, Smith BD, Aizensberg R, Rosenzweig AB, Fleshman JM and Matrisian LM. Projecting cancer incidence and deaths to 2030: the unexpected burden of thyroid, liver, and pancreas cancers in the United States. Cancer Res. 2014; 74:2913-2921.
8. Stolzenberg-Solomon RZ and Amundadottir LT. Epidemiology and Inherited Predisposition for Sporadic Pancreatic Adenocarcinoma. Hematol Oncol Clin North Am. 2015; 29:619-640.
9. Klein AP. Genetic susceptibility to pancreatic cancer. Molecular carcinogenesis. 2012; 51:14-24.
10. Amundadottir L, Kraft P, Stolzenberg-Solomon RZ, Fuchs CS, Petersen GM, Arslan AA, Bueno-de-Mesquita HB, Gross M, Helzlouer K, Jacobs EJ, LaCroix A, Zheng W, Albanes D, Bamlet W, Berg CD, Berrino F, et al. Genome-wide association study identifies variants in the ABO locus
associated with susceptibility to pancreatic cancer. Nat Genet. 2009; 41:986-990.

11. Petersen GM, Amundadottir L, Fuchs CS, Kraft P, Stolzenberg-Solomon RZ, Jacobs KB, Arslan AA, Bueno-de-Mesquita HB, Gallinger S, Gross M, Helzlsouer K, Holly EA, Jacobs EJ, Klein AP, LaCroix A, Li D, et al. A genome-wide association study identifies pancreatic cancer susceptibility loci on chromosomes 13q22.1, 1q32.1 and 5p15.33. Nat Genet. 2010; 42:224-228.

12. Wolpin BM, Rizzato C, Kraft P, Kooperberg C, Petersen GM, Wang Z, Arslan AA, Beane-Freeman L, Bracci PM, Buring J, Canzian F, Duell EJ, Gallinger S, Giles GG, Goodman GE, Goodman PJ, et al. Genome-wide association study identifies multiple susceptibility loci for pancreatic cancer. Nat Genet. 2014; 46:994-1000.

13. Wu C, Miao X, Huang L, Che X, Jiang G, Yu D, Yang X, Cao G, Hu Z, Zhou Y, Zuo C, Wang C, Zhang X, Yu X, Dai W, Li Z, et al. Genome-wide association study identifies five loci associated with susceptibility to pancreatic cancer in Chinese populations. Nat Genet. 2012; 44:62-66.

14. Low SK, Kuchiba A, Zembutsu H, Saito A, Takahashi A, Kubo M, Daigo Y, Kamatani N, Chiku S, Totsuka H, Ohnami S, Hirose H, Shimada K, Okusaka T, Yoshida T, Nakamura Y, et al. Genome-wide association study of pancreatic cancer in Japanese population. PloS one. 2010; 5:e1824.

15. Childs EJ, Mocci E, Campa D, Bracci PM, Gallinger S, Goggins M, Li D, Neale RE, Olson SH, Scelo G, Amundadottir LT, Bamlet WR, Bijlisma MF, Blackford A, Borges M, Brennan P, et al. Common variation at 2p13.3, 3q29, 7p13 and 17q25.1 associated with susceptibility to pancreatic cancer. Nat Genet. 2015; 47:911-916.

16. Marchini J and Howie B. Genotype imputation for genome-wide association studies. Nat Rev Genet. 2010; 11:499-511.

17. 1000 Genomes Consortium. A map of human genome variation from population-scale sequencing. Nature. 2010; 467:1061-1073.

18. Campa D, Rizzato C, Capurso G, Giese N, Petursdottir K, Petersen GM, Wang Z, Arslan AA, Beane-Freeman L, Bracci PM, Buring J, Canzian F, Duell EJ, Gallinger S, Giles GG, Goodman GE, Goodman PJ, et al. Genome-wide association study identifies multiple susceptibility loci for pancreatic cancer. Hum Mol Genet. 2014; 23:6616-6633.

19. Fayard E, Auwerx J and Schoonjans K. LRH-1: an orphan nuclear receptor involved in development, metabolism and digestive and liver disease. 2013; 45:95-99.

20. Kelly VR, Xu B, Kuick R, Koenig RJ and Hammer GD. Dax1 up-regulates Oct4 expression in mouse embryonic stem cells via LRH-1 and SRA. Mol Endocrinol. 2010; 24:2281-2291.

21. Rafnar T, Sulem P, Thorleifsson G, Vermeulen SH, Helgason H, Saemundsdottir J, Gudjonsson SA, Sigurdsson A, Stacey SN, Gudmundsson J, Johannsdottir H, Alexiusdottir K, Petursdottir V, Nikulasson S, Geirsson G, Jonsson T, et al. Genome-wide association study yields variants at 20p12.2 that associate with urinary bladder cancer. Hum Mol Genet. 2014; 23:5545-5557.

22. Kiemeneij LA, Thorlacius S, Sulem P, Geller F, Aben KK, Stacey SN, Gudmundsson J, Jakobsdottir M, Berghorssson JT, Sigurdsson A, Blondal T, Witjes JA, Vermeulen SH, Hulsbergen-van de Kaa CA, Swinkels DW, Ploeg M, et al. Sequence variant on 8q24 confers susceptibility to urinary bladder cancer. Nat Genet. 2008; 40:1307-1312.

23. Rothman N, Garcia-Closas M, Chatterjee N, Malats N, Wu X, Figueroa JD, Real FX, Van Den Berg D, Matullo G, Baris D, Thun M, Kiemeneij LA, Vineis P, De Vivo I, Albanes D, Purdue MP, et al. A multi-stage genome-wide association study of bladder cancer identifies multiple susceptibility loci. Nat Genet. 2010; 42:978-984.

24. Wang Z, Zhu B, Zhang M, Parikh H, Jia J, Chung CC, Sampson JN, Hoskins JW, Hutchinson A, Burdette L, Ibrahim A, Hautman C, Raj PS, Abnet CC, Adjei AA, Ahlbom A, et al. Imputation and subset-based association analysis across different cancer types identifies multiple independent risk loci in the TERT-CLPTM1L region on chromosome 5p15.33. Hum Mol Genet. 2014; 23:6616-6633.

25. Ward LD and Kellis M. HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. Nucleic Acids Res. 2012; 40:D930-934.

26. Boyle AP, Hong EL, Haririyan M, Cheng Y, Schaumb MA, Kasowski M, Karcewski KJ, Park J, Hitz BC, Weng S, Cherry JM and Snyder M. Annotation of functional variation in personal genomes using RegulomeDB. Genome Res. 2012; 22:1790-1797.

27. GTEx Consortium. The Genotype-Tissue Expression (GTEx) pilot analysis: Multitissue gene regulation in humans. Science. 2015; 348:648-660.

28. Hoskins JW, Jia J, Flandez M, Parikh H, Xiao W, Collins I, Emmanuel MA, Ibrahim A, Powell J, Zhang L, Malats N, Bamlet WR, Petersen GM, Real FX and Amundadottir LT. Transcriptome analysis of pancreatic cancer reveals a tumor suppressor function for HNF1A. Carcinogenesis. 2014; 35:2670-2678.

29. Holmstrom SR, Deering T, Swift GH, Poelwijk FJ, Mangelsdorf DJ, Kliwer SA and MacDonald RJ. LRH-1 and PTF1-L coregulate an exocrine pancreas-specific transcriptional network for digestive function. Genes & development. 2011; 25:1674-1679.

30. Flandez M, Cendrowski J, Canamero M, Salas A, Del Pozo N, Schoonjans K and Real FX. Nr5a2 heterozygosity sensitises to, and cooperates with, inflammation in KRasG12V-driven pancreatic tumourigenesis. Gut. 2014; 63:467-55.

31. von Figura G, Morris JPt, Wright CV and Hebrok M. Nr5a2 maintains acinar cell differentiation and constrains oncogenic Kras-mediated pancreatic neoplastic initiation.
32. Benod C, Vinogradova MV, Jouravel N, Kim GE, Fletterick RJ and Sablin EP. Nuclear receptor liver receptor homologue 1 (LRH-1) regulates pancreatic cancer cell growth and proliferation. Proc Natl Acad Sci U S A. 2011; 108:16927-16931.

33. Lin Q, Aihara A, Chung W, Li Y, Huang Z, Chen X, Weng S, Carlson RI, Wands JR and Dong X. LRH1 as a driving factor in pancreatic cancer growth. Cancer letters. 2014; 345:85-90.

34. Sur I, Tuupanen S, Whittington T, Aaltonen LA and Taipale J. Lessons from functional analysis of genome-wide association studies. Cancer Res. 2013; 73:4180-4184.

35. Gudmundsson J, Sulem P, Gudbjartsson DF, Masson G, Petursdottir V, Hardarson S, Gudjonsson SA, Johannsdottir H, Helgadottir HT, Stacey SN, Magnusson OT, Helgason H, Panadero A, van der Zanden LF, Aben KK, Vermeulen SH, et al. A common variant at 8q24.21 is associated with renal cell cancer. Nature communications. 2013; 4:2776.

36. Park SL, Chang SC, Cai L, Cordon-Cardo C, Ding BG, Greenland S, Hussain SK, Jiang Q, Liu S, Lu ML, Mao JT, Morgenstern H, Mu LN, Ng LJ, Pantuck A, Rao J, et al. Associations between variants of the 8q24 chromosome and nine smoking-related cancer sites. Cancer Epidemiol Biomarkers Prev. 2008; 17:3193-3202.

37. Shete S, Hosking FJ, Robertson LB, Dobbins SE, Sanson M, Malbrer B, Simon M, Marie Y, Boiselier B, Delattre JY, Hoang-Xuan K, El Hallani S, Idbaih A, Zelenika D, Andersson U, Henriksen R, et al. Genome-wide association study identifies five susceptibility loci for glioma. Nat Genet. 2009; 41:899-904.

38. Goode EL, Chenevix-Trench G, Song H, Ramus SJ, Notaridou M, Lawrenson K, Widenschwender M, Vierkant R, Larson MC, Kjaer SK, Birrer MJ, Berchuck A, Schildkraut J, Tomlinson I, Kiemeney LA, Cook LS, et al. A genome-wide association study identifies susceptibility loci for ovarian cancer at 2q31 and 8q24. Nat Genet. 2010; 42:874-879.

39. Dang CV, Ostnoll KA, Zeller KI, Nguyen T, O'sus RC and Li F. The c-Myc target gene network. Semin Cancer Biol. 2006; 16:253-264.

40. Pomerantz MM, Ahmadiyeh N, Jia L, Herman P, Verzi MP, Doddapaneni H, Beckwith CA, Chan JA, Hilds A, Davis M, Yao K, Keoh SM, Lenz HJ, Haiman CA, Yan C, Henderson BE, et al. The 8q24 cancer risk variant rs6983267 shows long-range interaction with MYC in colorectal cancer. Nat Genet. 2009; 41:882-884.

41. Dryden NH, Broome LR, Dudbridge F, Johnson N, Orr N, Schoenfelder S, Nagano T, Andrews S, Wingett S, Kozarewa I, Assiotis I, Fenwick K, Maguire SL, Campbell J, Natarajan R, Lambros M, et al. Unbiased analysis of potential targets of breast cancer susceptibility loci by Capture Hi-C. Genome Res. 2014; 24:1854-1868.

42. Ahmadiyeh N, Pomerantz MM, Grisanzio C, Herman P, Jia L, Almendro V, He HH, Brown M, Liu XS, Davis M, Caswell JL, Beckwith CA, Hills A, Macennaol L, Coetzee GA, Regan MM, et al. 8q24 prostate, breast, and colon cancer risk loci show tissue-specific long-range interaction with MYC. Proc Natl Acad Sci U S A. 2010; 107:9742-9746.

43. Breyer JP, Dorset DC, Clark TA, Bradley KM, Wahlfors TA, McReynolds KM, Maynard WH, Chang SS, Cookson MS, Smith JA, Schleutker J, Dupont WD and Smith JR. An expressed retrogene of the master embryonic stem cell gene POU5F1 is associated with prostate cancer susceptibility. Am J Hum Genet. 2014; 94:395-404.

44. Ling H, Spizzo R, Atalay Y, Nicoloso M, Shimizu M, Redis RS, Nishida N, Gafa R, Song J, Guo Z, Ivan C, Barbarotto E, De Vries I, Zhang X, Ferracin M, Churchman M, et al. CCAT2, a novel noncoding RNA mapping to 8q24, underlies metastatic progression and chromosomal instability in colon cancer. Genome Res. 2013; 23:1446-1461.

45. Meyer KB, Maia AT, O’Reilly M, Ghousaini M, Prathalingam R, Porter-Gill P, Amb S, Prokunina-olsson L, Carroll J and Ponder BA. A functional variant at a prostate cancer predisposition locus at 8q24 is associated with PVT1 expression. PLoS Genet. 2011; 7:e1002165.

46. Wang M, Zhang W, Yuan L, Fu G, Wei Q and Zhang Z. Common genetic variants on 8q24 contribute to susceptibility to bladder cancer in a Chinese population. Carcinogenesis. 2009; 30:991-996.

47. Tseng YY, Moriarity BS, Gong W, Akiyama R, Tiwari A, Kawakami H, Ronning P, Reuland B, Guenther K, Beadnell TC, Essig J, Otto GM, O’Sullivan MG, Largaespada DA, Schwertfeger KL, Marahrens Y, et al. PVT1 dependence in cancer with MYC copy-number increase. Nature. 2014; 512:82-86.

48. Colombo T, Farina L, Macino G and Paci P. PVT1: a rising star among oncogenic long noncoding RNAs. Biomed Res Int. 2015; 2015:304208.

49. Huang C, Yu W, Wang Q, Cui H, Wang Y, Zhang L, Han F and Huang T. Increased expression of the lncRNA PVT1 is associated with poor prognosis in pancreatic cancer patients. Minerva Med. 2015; 106:143-149.

50. Tumbles C, Rapley EA, Seal S, Pernet D, Ruck J, Hughes D, Ricketts M, Linger R, Nsenjima J, Deloukas P, Huddart RA, Bishop DT, Easton DF, Stratton MR and Rahman N. Variants near DMRT1, TERT and ATF7IP1 with MYC. Proc Natl Acad Sci U S A. 2010; 107:604-607.

51. Haiman CA, Chen GK, Vachon CM, Canzian F, Dunning A, Millikan RC, Wang X, Ademuyiwa F, Ahmed S, Ambrosone CB, Baglietto L, Balleine R, Bandera EV, Beckmann MW, Berg CD, Bernstein L, et al. A common variant at the TERT-CLPTM1L locus is associated with estrogen receptor-negative breast cancer. Nat Genet. 2011; 43:1210-1214.
52. Berndt SI, Skibola CF, Joseph V, Camp NJ, Nieters A, Wang Z, Cozen W, Monnereau A, Wang SS, Kelly RS, Lan Q, Teras LR, Chatterjee N, Chung CC, Yeager M, Brooks-Wilson AR, et al. Genome-wide association study identifies multiple risk loci for chronic lymphocytic leukemia. Nat Genet. 2013; 45:868-876.

53. Wang Y, Broderick P, Webb E, Wu X, Vijayakrishnan J, Matakodou A, Qureshi M, Dong Q, Gu X, Chen WV, Spitz MR, Eisen T, Amos CI and Houlston RS. Common 5p15.33 and 6p21.33 variants influence lung cancer risk. Nat Genet. 2008; 40:1407-1409.

54. McKay JD, Hung RJ, Gaborieau V, Boffetta P, Chabrier A, Byrnes G, Zardizde D, Mukeria A, Szeszenia-Dabrowska N, Lissowska J, Rudnai P, Fabianova E, Mates D, Bencko V, Foretova L, Janout V, et al. Lung cancer susceptibility locus at 5p15.33. Nat Genet. 2008; 40:1404-1406.

55. Landi MT, Chatterjee N, Yu K, Goldin LR, Goldstein AM, McKay JD, Hung RJ, Gaborieau V, Boffetta P, Chabrier A, Byrnes G, Zardizde D, Mukeria A, Szeszenia-Dabrowska N, Lissowska J, Rudnai P, Fabianova E, Mates D, Bencko V, Foretova L, Janout V, et al. A genome-wide association study of lung cancer identifies a region of chromosome 5p15 associated with risk for adenocarcinoma. Am J Hum Genet. 2009; 85:679-691.

56. Rafnar T, Sulem P, Stacey SN, Geller F, Gundmundsson J, Sigurdsson A, Jakobsdottir M, Helgadottir H, Thorlacius S, Aben KK, Blondal T, Hemminki K, Thorleifsson G, Kristjansson K, Thorisdottir K, Ragnarsson R, et al. Sequence variants at the TERT-CLPTM1L locus associate with many cancer types. Nat Genet. 2009; 41:221-227.

57. Stacey SN, Sulem P, Masson G, Gurzau E, et al. New common variants affecting susceptibility to basal cell carcinoma. Nat Genet. 2009; 41:909-914.

58. Yang X, Yang B, Li B and Liu Y. Association between TERT-CLPTM1L rs401681[CC] allele and NMBC cancer risk: a meta-analysis including 45,184 subjects. Arch Dermatol. Res. 2013; 305:49-52.

59. Beesley J, Pooley KA, Johnatty SE, Beesley J, Michailidou K, Tyrer JP, Edwards SL, Pickett HA, Shen HC, Smart CE, Hillman KM, Mai PL, Lawrenson K, Stutz MD, Lu Y, Karevan R, et al. Multiple independent variants at the TERT locus are associated with telomere length and risks of breast and ovarian cancer. Nat Genet. 2013; 45:371-384.

60. Campa D, Rizzato C, Stolzenberg-Solomon R, Pacetti P, Vodicka P, Cleary SP, Capuro G, Bueno-de-Mesquita HB, Werner J, Gaozuli M, Butterbach K, Ivanauskas A, Giese N, Petersen GM, Fogar P, Wang Z, et al. TERT gene harbors multiple variants associated with pancreatic cancer susceptibility. Int J Cancer. 2015; 137:2175-83.

61. Zhang Y, Calado R, Rao M, Hong JA, Meeker AK, Dumitriu B, Atay S, McCormick PJ, Garfield SH, Wangsa D, Padilla-Nash HM, Burkett S, Zhang M, Kunst TF,
Peterson NR, Xi S, et al. Telomerase variant A279T induces telomere dysfunction and inhibits non-canonical telomerase activity in esophageal carcinomas. PloS one. 2014; 9:e101010.

72. Howie BN, Donnelly P and Marchini J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. PLoS Genet. 2009; 5:e1000529.

73. Marchini J, Howie B, Myers S, McVean G and Donnelly P. A new multipoint method for genome-wide association studies by imputation of genotypes. Nat Genet. 2007; 39:906-913.

74. de Bakker PI, Ferreira MA, Jia X, Neale BM, Raychaudhuri S and Voight BF. Practical aspects of imputation-driven meta-analysis of genome-wide association studies. Hum Mol Genet. 2008; 17:R122-128.

75. Howie B, Marchini J and Stephens M. Genotype Imputation with Thousands of Genomes. G3: Genes, Genomes, Genetics. 2011; 1:457-470.

76. Luna A and Nicodemus KK. snp.plotter: an R-based SNP/haplotype association and linkage disequilibrium plotting package. Bioinformatics. 2007; 23:774-776.

77. Berndt S YM, Wang Z, Jacobs K, Sampson J, Margaret Tucker M, Chanock C. Pegasus GWAS. Nature Genetics, in preparation. 2013.

78. Browning BL and Browning SR. A unified approach to genotype imputation and haplotype-phase inference for large data sets of trios and unrelated individuals. Am J Hum Genet. 2009; 84:210-223.