Published in final edited form as:
Clin Cancer Res. 2012 July 15; 18(14): 3942–3951. doi:10.1158/1078-0432.CCR-11-2856.

A Replication Study and Genome-wide Scan of Single Nucleotide Polymorphisms Associated with Pancreatic Cancer Risk and Overall Survival

Jason A. Willis[1], **Sara H. Olson**[2], **Irene Orlow**[2], **Semanti Mukherjee**[1,3], **Robert R. McWilliams**[5], **Robert C. Kurtz**[4], and **Robert J. Klein**[1]

[1]Program in Cancer Biology and Genetics, Memorial Sloan-Kettering Cancer Center, New York, NY 10065 USA

[2]Department of Epidemiology and Biostatistics, Memorial Sloan-Kettering Cancer Center, New York, NY 10065 USA

[3]Gerstner Sloan-Kettering Graduate School of Biomedical Sciences, New York, NY 10065, USA

[4]Department of Medicine, Memorial Sloan-Kettering Cancer Center, New York, NY 10065 USA

[5]Division of Medical Oncology, Mayo Clinic, Rochester, Minnesota

Abstract

Purpose—To explore the effects of single nucleotide polymorphisms (SNPs) on pancreatic cancer risk and overall survival.

Experimental Design—The germline DNA of 531 pancreatic cancer cases and 305 healthy controls from a hospital-based study was genotyped at SNPs previously reported to be associated with pancreatic cancer risk or clinical outcome. We analyzed putative risk SNPs for replication of their reported effects on risk and tested for novel effects on overall survival (OS). Similarly, we analyzed putative survival-associated SNPs for replication of their reported effects on OS and tested for novel effects on risk. Lastly, we performed a genome-wide association study of OS using a subset of 252 cases, with two subsequent validation sets of 261 and 572 patients, respectively.

Results—Among seven risk SNPs analyzed, two (rs505922, rs9543325) were associated with risk ($p<0.05$). Among 24 survival-associated SNPs analyzed, one (rs9350) was associated with OS ($p<0.05$). No putative risk SNPs or putative survival-associated SNPs were found to be associated with OS or risk, respectively. Further, our GWAS identified a novel SNP (rs1482426, combined stage 1 and 2 $p = 1.7 \times 10^{-6}$, per-allele HR = 1.74, 95% CI 1.38–2.18) to be putatively associated with OS.

Conclusions—The effects of SNPs on pancreatic cancer risk and overall survival were replicated in our study, though further work is necessary to understand the functional mechanisms underlying these effects. More importantly, the putative association with OS identified by GWAS suggests that GWAS may be useful in identifying SNPs associated with clinical outcome in pancreatic cancer.

Corresponding Author: Robert J. Klein, Program in Cancer Biology and Genetics, Memorial Sloan-Kettering Cancer Center, 417 East 68th Street, New York, NY 10065. Phone: 646-888-2525; Fax: 646-422-0197; kleinr@mskcc.org.

Conflicts of Interest: none

Keywords

pancreatic cancer; susceptibility; overall survival; genome-wide association study (GWAS)

INTRODUCTION

Pancreatic adenocarcinoma is a rapidly fatal disease with poor long-term prognosis; the 5-year survival rate is estimated to be <6%. Despite its relatively low incidence (3% of new cancer cases in the United States) and recent advances in surgical treatment, pancreatic cancer remains the fourth-leading cause of cancer mortality in the U.S (1, 2). Thus, intense efforts are underway to develop novel strategies for both screening and treatment.

One approach to this challenge has been the identification of inherited genetic variations that are involved with susceptibility to pancreatic cancer. Recently, genome-wide association studies (GWAS) performed by the PanScan consortium identified four loci on chromosomes 9q34 (*ABO*), 13q22, 1q32 (*NR5A2*) and 5p15 (*CLPTM1L-TERT*) associated with increased risk (3, 4). A similar study performed in the Japanese population identified three additional risk loci on chromosomes 6p26, 12p11 (*BICD1*), and 7q36 (*DPP6*) (5). Independent replication of these findings is a critical step towards understanding their functional significance. However, while three of the associations reported by PanScan were moderately replicated in the Japan-based study, there have been no reported attempts to independently replicate the risk loci identified in the Japan-based study.

There is also evidence that inherited genetic variations may influence the clinical outcome of pancreatic cancer. In several recent studies, a candidate-gene approach was used to identify SNPs within different cellular pathways (*e.g.* DNA damage repair and mitotic regulation) that are associated with overall survival (6–17). These findings could potentially identify novel therapeutic targets and/or strategies. Yet, because the majority of these associations have not been replicated in independent patient groups, their generalizability is still unknown. Furthermore, the effects on overall survival contributed by genetic variations in unexplored regions of the genome are not well understood.

Here, we addressed these issues by performing a combined replication and discovery study of genetic variants associated with pancreatic cancer risk and overall survival within a multi-ethnic, hospital-based case-control group. First, we attempted to replicate the risk associations of SNPs identified by the PanScan or Japan-based pancreatic cancer risk GWAS, and we assessed these SNPs for association with overall survival. Next, we attempted to replicate the overall survival associations of SNPs reported by previous candidate-gene studies of pancreatic cancer clinical outcome, and evaluated these SNPs for association with risk. Finally, we performed a GWAS to identify novel variants associated with overall survival.

MATERIALS AND METHODS

Study Population and Sample Collection

Memorial Sloan-Kettering Cancer Center Study—Participants were part of an ongoing hospital-based case-control project conducted in conjunction with the Familial Pancreatic Tumor Registry (FPTR) at Memorial Sloan-Kettering Cancer Center (MSKCC). Patients were eligible if they were age 21 or over, spoke English, and had pathologically or cytologically confirmed adenocarcinoma of the pancreas. Patients were recruited between June 2003 and July 2009 from the surgical and medical oncology clinics at MSKCC at the time of their initial diagnosis or during follow-up. Controls were spouses of patients or

visitors accompanying patients with other diseases, had the same age and language eligibility requirements as the cases, had no personal history of cancer (except for non-melanoma skin cancer), and were not blood relatives of the cases. A total of 531 cases and 305 controls from MSKCC participated in our analyses. The participation rate among approached and eligible individuals was 76% among cases and 56% among controls. The study was approved by the MSKCC Institutional Review Board and all enrolled participants signed informed consent.

Participants provided a blood or buccal (mouthwash or saliva) sample to the MSKCC FPTR research study assistant and completed risk factor and family history questionnaires administered by the research study assistant in person or via telephone. Biospecimens were subsequently delivered for genomic DNA extraction and banking to the Molecular Epidemiology Laboratory. DNA was isolated from mouthwash specimens using the Puregene DNA purification kit (Qiagen, Inc; Valencia CA, from saliva samples with the Oragene saliva kits (DNA Genotek; Kanata, Ontario, Canada), and from whole blood using the Gentra Puregene blood kit (Qiagen Inc; Valencia CA). DNA samples were hydrated in 1× TE buffer.

Mayo Clinic Study—Participants from the Mayo Clinic were identified and recruited as described previously (18). A subset of patients was selected for genotyping based on the following criteria: 1) consented to enroll in the Mayo Clinic SPORE registry; 2) diagnosed with histopathologically-confirmed pancreatic adenocarcinoma (not invasive intraductal papillary mucinous neoplasm or other histology); and 3) a risk factor questionnaire had been completed. In total, 572 genotyped patients from Mayo Clinic with non-Hispanic Caucasian ethnicity were included in our analyses.

Whole-genome SNP Array Genotyping and Quality-Control

MSKCC Study—Genomic DNA samples from 263 cases were each genotyped on the Illumina CNV370 SNP bead array (either the Illumina CNV370-Duo or Illumina CNV370-Quad) at the Genomics Core Laboratory of MSKCC according to the manufacturer's protocol. The array contains probes for 351,496 SNPs. Genotype calls were made in the Illumina BeadStudio software package and exported to PLINK (version 1.07) (19) for processing.

After genotyping, quality control was first applied by removing SNPs with call rates < 99%, or minor allele frequency (MAF) < 1% in the genotyped cases. In addition, we excluded SNPs that showed extreme deviation from Hardy-Weinberg equilibrium (HWE) in the genotyped cases. Because we intended to perform a case-only survival analysis and deviation from HWE in cases may be a sign of a true risk SNP, we chose a less stringent approach by excluding SNPs with HWE P-value < 1×10^{-20}. in the genotyped cases. Identify-by-descent (IBD) analysis was performed in PLINK to confirm that none of the genotyped cases were blood relatives. Furthermore, none of the genotyped cases were excluded from analysis on the basis of sample-level quality control metrics (overall genotyping rate ≥ 90%; the minimum genotyping rate was found to be 91.5%). Thus, after quality control, 301,250 SNPs (MAF ≥ 1%) were available for downstream statistical analysis in 263 cases.

Mayo Clinic Study—Whole-genome SNP genotyping of pancreatic cancer cases from the Mayo Clinic study was performed as previously described (3, 4). In this study, we extracted SNP rs1482426 genotypes from the Mayo Clinic data set to perform stage 3 analysis in the GWAS.

Mass Spectrometry-Based SNP Genotyping and Quality-Control

We performed matrix-assisted laser desorption/ionization-time of flight (MALDI-TOF) mass spectrometry-based SNP genotyping using Sequenom iPLEX Gold assays (Sequenom, Inc.; San Diego, CA). iPLEX assays were designed and multiplexed using the MassARRAY Assay Designer software package (Sequenom Inc., version 4.0). In total, we designed 3 multiplexes comprised of 64 individual assays targeting risk- and survival-associated SNPs selected for replication study as well as novel survival-associated SNPs identified by GWAS.

Polymerase chain reaction (PCR) amplification primers and single-base extension (SBE) oligonucleotides were synthesized for each SNP assay according to the designed specifications (Integrated DNA Technologies, Inc.; Coralville, IA). PCR primers within the same multiplex were mixed to a final concentration of 500nM each. SBE oligonucleotide concentrations were optimized according Sequenom's recommendations. Prior to use, all DNA samples were whole-genome amplified (WGA) using the Illustra GenomiPhi v2 DNA Amplification Kit (GE Healthcare), following manufacturer's recommendations. The WGA reactions were then diluted by adding 120 μL of reduced TE buffer; an additional 2-fold dilution of the DNA was made as it empirically improved genotyping performance in our hands.

Genotyping reactions were performed in batches of 384-well plates based on Sequenom's recommendations. Briefly, for PCR amplification, template DNA was added to each well along with a master mix containing amplification primers, dNTPs, and HotStar Taq DNA polymerase PCR reagents (Qiagen, Inc.; Valencia CA). After PCR cycling, unincorporated dNTPs were neutralized by treatment with shrimp alkaline phosphatase (SAP) supplied by the iPLEX Gold reagent kit (Sequenom Inc.). SAP treatment was then followed by single base extension (SBE) reactions using the iPLEX Gold enzyme and reagent kit (Sequenom Inc.). Subsequent steps, including reaction desalting, spotting, and SpectroCHIP analysis, were performed by the MSKCC Sequenom Facility according to protocol. Mass spectra were analyzed with the Typer software package (Sequenom Inc., version 4.0) to make genotype calls. Each individual batch of 384-well plates was reviewed for quality control and the data were exported to PLINK. Batch results were merged together in PLINK. Further quality control was applied by removing six SNPs with call rates < 90% and individuals with overall genotyping rates < 90%.

A subset of individuals (n = 263 cases and 203 controls) who were genotyped by iPLEX assays was also genotyped by array-based methods (described above). Thus, to assess the overall quality our iPLEX genotyping experiment, we analyzed the results for concordance with array-based calls. In the analysis of 39 SNPs that were present on the array and in the iPLEX experiment, we found the overall concordance rate to be >99.7%.

Statistical Methods

Risk Association—SNPs were analyzed for association with pancreatic cancer risk by use of a logistic regression model whereby genotypes were coded as either 0,1 or 2 depending on the number of minor alleles carried by an individual; race, age (years) and gender (male or female) were included as covariates. This model estimated a log-additive effect of each SNP on disease risk per additional minor allele. Statistical significance was determined using a 1-degree-of-freedom (d.f) Wald test. For the replication study of previously-reported risk SNPs, power analysis was performed using the CaTS tool (20) based on a disease prevalence of 0.01%, the reported effect size of each SNP, the minor allele frequency of each SNP observed in controls, our sample size, and a significance level of 0.05.

Overall Survival Association—Survival analyses were performed separately in the MSKCC and Mayo Clinic studies:

MSKCC Study: SNPs and clinical variables were analyzed for association with overall survival by use of a Cox proportional hazards model as implemented in the *survival* package for R. We measured overall survival from the date of diagnosis until the date of death (any cause) or last follow-up. Notably, the cases in our study were recruited at variable times following diagnosis — a setting which potentially may lead to biased hazard ratio estimates (21). Therefore, we modified the analysis to allow for left truncation, whereby cases were considered to be at risk of death only after the time of study recruitment. Of the initial 531 cases included in our study, 18 cases were either missing the date-of-diagnosis (n = 1) or were lost to follow-up (n = 17) after their initial recruitment. Therefore, follow-up data was available for a total of 513 cases.

To analyze a given SNP for association with overall survival, we considered one or more of the following Cox regression models. *Model 1*: a univariate Cox model with SNP genotypes coded as either 0,1, or 2 depending on the number of minor alleles carried by an individual; *Model 2*: a multivariate Cox model including the SNP (coded as in *Model 1*) and race/ethnicity as a covariate; *Model 3*: a multivariate Cox model including the SNP (coded as in Model 1), race/ethnicity, Eastern Cooperative Oncology Group (ECOG) performance status (coded as integer values between 0 and 5); clinical stage at diagnosis (which was abstracted from the medical record and categorized as either localized, locally advanced or metastatic), and radiation treatment (yes/no) as covariates. The clinical variables included in *Model 3* were each chosen on the basis of having significant ($p \leq 0.05$) hazard ratio (HR) estimates when individually tested in a univariate Cox model (Supplemental Table 1). Using the "strata()" function in the R *survival* package, Model 3 was further stratified by surgical resection status (yes/no), as this variable was found to violate the Cox proportional hazards assumption (data not shown).

In all of the above models, a log-additive effect (hazard ratio per additional minor allele) was estimated for each SNP; statistical significance was determined using a 1-degree of freedom (d.f.) Wald test. Cumulative survival probabilities within different SNP genotype strata were estimated by Kaplan-Meier analysis methods.

The first stage of the survival GWAS utilized whole-genome SNP array data for 252 of the original subset of 263 cases genotyped on the Illumina CNV370 chip that had follow-up information. SNPs from the array with MAF ≥ 1% were each tested for association with overall survival using *Model 1* (described above) in the full (n = 252) discovery case set. SNPs were carried forward to a validation stage based on the following criteria: 1) the proportional hazards assumption was not violated, as determined by evaluation of the Schoenfeld residuals; and 2) association test $p < 1\times10^{-4}$. In the validation stage, an independent set of 261 cases was used for analysis. Each SNP carried forward from the first stage of the GWAS was analyzed using the validation case set under *Model 2* (to adjust for the multi-ethnic composition of the validation case set).

For the replication study of SNPs previously reported for association with overall survival, power analysis was performed using methods described by Schoenfeld (22) based on the minor allele frequency of each SNP in the CEU HapMap sample and its reported effect size

Mayo Clinic Study: Overall survival was measured from the date of diagnosis until the date of death (from all causes) or last follow-up for those not known to be deceased at the time of analysis. SNPs were coded as the number of minor alleles carried (0,1, or 2) by an individual. Under an additive genetic model, associations between SNPs and overall survival

were analyzed using Cox proportional hazards regression as implemented in the R *survival* package. Significance was determined using a chi-square test comparing the association model with and without the SNP using the *anova.coxph* function. In addition to univariate SNP analyses, multivariable models were also considered which adjusted for age at diagnosis, sex, Karnofsky Performance Score, BMI and pancreatic cancer stage.

RESULTS

Study Design and Characteristics of the Study Participants

The overall design and principal findings of our combined replication/discovery study are summarized in Table 1. A total of 836 individuals (531 pancreatic cancer cases and 305 healthy, unrelated controls) from a case-control study group based at the Memorial Sloan-Kettering Cancer Center (MSKCC) were included in the analyses of previously-reported risk SNPs, previously-reported survival SNPs, or stages 1 and 2 of the GWAS for novel survival SNPs described below. Table 2 describes the overall demographics of MSKCC cases and controls used in our study. Table 3 describes the clinical characteristics of MSKCC cases pertaining to survival analysis. An additional 572 pancreatic cancer cases from the Mayo Clinic (described in ref. 26) were included as part of the stage 3 analysis of a putative survival-associated SNP identified by GWAS.

Replication Analysis of SNPs Associated with Pancreatic Cancer Risk

We first determined whether pancreatic cancer risk associations identified by GWAS could be replicated in the MSKCC case-control group. From the PanScan studies, we selected SNPs rs3790844, rs401681, rs505922, and rs9543325, which were reported to have the strongest association signal (lowest P-value) at risk loci on chromosomes 1q32, 5p15, 9q34, and 13q22, respectively. From the Japan-based study, we selected SNPs rs9502893 (6p26), rs6464375 (7q36), and rs708224 (12q11) on the basis of having the three smallest reported P-values in that study ($p = 3.3\times10^{-7}$, 4.4×10^{-7}, and 3.3×10^{-7}, respectively).

After genotyping cases and controls, each SNP was analyzed for association with risk by use of a logistic regression model adjusted for race, age and gender (Table 4). Notably, a subset (n = 283) of individuals in the MSKCC case-control group also participated in the PanScan study. Thus, for the purpose of independent replication, we excluded these individuals from the analysis of risk SNPs identified by PanScan. Overall, at a nominal P-value of 0.05, we replicated the risk associations of SNPs rs505922 ($p = 0.002$, per-allele OR = 1.65, 95% CI 1.20–2.26) located in the *ABO* gene, and rs9543325 (p = 0.02, per-allele OR = 1.42, 95% CI 1.06–1.90) located in a non-genic region of chromosome 13q22.

Survival Analysis of SNPs Associated with Pancreatic Cancer Risk

We next determined whether each of the seven SNPs selected for risk replication was also associated with overall survival of pancreatic cancer in the MSKCC case group. The SNPs were analyzed by use of Cox proportional hazard models adjusted for race (Supplemental Table 2). However, no statistically significant ($p < 0.05$) associations were observed.

Replication Analysis of SNPs Associated with Overall Survival of Pancreatic Cancer

Utilizing a candidate-gene approach, several recent studies have reported SNPs that are associated with overall survival of pancreatic cancer. For replication analysis, we selected a set of 24 previously reported SNPs based on three criteria: i) we had adequate (>80%) study power to detect the reported hazard ratio (HRs) under a dominant model, assuming $\alpha = 0.05$; ii) a genotyping assay could be designed; and iii) genotyping met quality-control criteria (see Materials and Methods). After genotyping, we tested each SNP for association with

overall survival in the MSKCC case group using a Cox proportional Hazards model adjusted for race (Table 5).

We observed one nominally significant association ($p < 0.05$) at SNP rs9350 ($p = 0.007$, per-allele HR = 1.26, 95% CI 1.07–1.50), located in the *EXO1* gene, with an estimated HR that was consistent with previous reports. Patients who were heterozygous or homozygous minor allele carriers of rs9350 had an estimated median survival of 0.99 years (95% CI 0.80–1.21), versus 1.28 years (95% CI 1.12–1.49) for homozygous major allele carriers (Supplemental Figure 1a). In multivariate Cox analysis — stratified by surgical resection (yes/no) and adjusted for race, ECOG status, clinical stage and radiation treatment— rs9350 ($p = 0.002$, adjusted per-allele HR = 1.33, 95% CI 1.11–1.60) remained statistically significantly associated with overall survival (Supplemental Table 4).

Notably, we observed one additional SNP — rs8191754 ($p = 0.02$, per-allele HR 0.77, 95% CI 0.62–0.95), located in the *IGF2R* gene — to be associated with overall survival at $p < 0.05$ in our study. However, the estimated HR for this SNP was in the opposite direction of previous reports (9).

Risk Analysis of SNPs Associated with Overall Survival of Pancreatic Cancer

We next tested whether each of the 24 SNPs chosen for overall survival analysis was also associated with pancreatic cancer risk in the MSKCC case-control group. The SNPs were analyzed using a log-additive logistic regression model adjusted for race, age, and gender (Supplemental Table 3). However, no statistically significant ($p < 0.05$) associations were observed.

Genome-wide scan for novel SNPs associated with Overall Survival

To identify novel SNPs associated with overall survival, we performed a three-stage GWAS. The first stage of analysis was performed within a set of 252 MSKCC cases for which we had follow-up information and whole-genome SNP array data. We tested each of 301,250 SNPs (with MAF ≥ 1% in cases) for association with overall survival under a univariate Cox proportional hazards model. None of SNPs tested in this analysis reached genome-wide significance ($p < 5\times10^{-8}$) (Supplemental Figures 2 and 3). However, 22 candidate SNPs with $p < 1\times10^{-4}$ were carried forward to a second (validation) stage. Genotyping was performed in the validation case set (n = 261 MSKCC cases) and each SNP was tested for association with overall survival using a Cox proportional hazards model adjusted for race (Table 6).

Simultaneously, under the hypothesis that different genetic variants play a role in the survival of pancreatic cancer cases that undergo surgical tumor resection compared to cases that do not, we performed similar analyses (up to stage 1 and 2) in those respective case sub-groups (data not shown). However, as no SNPs tested in stage 2 were significant in those analyses, we focused the remainder of our study on candidate SNPs identified in full-group analysis.

Notably, of the 22 candidate SNPs tested in stage 2, we successfully validated the association of SNP rs1482426 ($p = 0.001$, per-allele HR = 1.73, 95% CI 1.25–2.40), which is located in a non-genic region of chromosome 12q21 (Supplemental Figure 4). In a combined analysis using both the GWAS (discovery) and validation case sets, rs1482426 was found to improve in overall rank and P-value ($p = 1.7\times10^{-6}$, per-allele HR = 1.74, 95% CI 1.38–2.18). Patients who were heterozygous or homozygous minor allele carriers of rs1482426 had an estimated median survival of 0.78 years (95% CI 0.53–0.95), versus 1.28 years (95% CI 1.19–1.49) for homozygous major allele carriers (Supplemental Figure 1b). Furthermore, multivariate Cox analysis of the combined case sets — stratified by surgical

resection (yes/no) and adjusted for race, ECOG status, clinical stage, and radiation treatment – showed that rs1482426 was independently associated with overall survival ($p = 9.0\times10^{-6}$, per-allele HR = 1.70, 95% CI 1.34–2.14) (Supplemental Table 4).

To gain further statistical evidence of its association with overall survival, we selected SNP rs1482426 for a third (replication) stage of analysis involving an independent case group (n = 572) based at the Mayo Clinic. In univariate Cox regression analysis of rs1482426, we observed suggestive evidence of association ($p = 0.0755$, per-allele HR = 1.19) in the Mayo case group that, although weaker, was consistent with the direction of effect estimated in stages 1 and 2 of the GWAS.

DISCUSSION

In this study, we have replicated previously reported pancreatic cancer risk and survival SNPs, and we have identified a novel putative pancreatic cancer survival SNP. We first focused on a replication analysis of SNPs reported by recent pancreatic cancer risk GWAS. Both the *ABO* locus (marked by SNP rs505922) and the 13q22 locus (marked by SNP rs9543325) were found to be associated with risk in our case-control group. Notably, our estimate of the per-allele odds ratio for SNP rs505922 (per-allele OR = 1.65, 95% CI 1.20–2.26) is larger than those reported by PanScan or the Japan-based GWAS. Indeed, it is also larger than most odds ratios estimated for other common cancer risk SNPs. However, since the 95% confidence interval of our estimate overlaps with previous reports, we cannot draw specific conclusions about this observation.

In addition, we observed that three other SNPs in our risk replication analysis — rs708224 (*BICD1*), rs9502893 (*FOXQ1*) and rs401681 (*CLPTM1L*) — had OR estimates that were trending towards significance ($p < 0.2$) and were consistent with previous GWAS reports. Indeed, based on our power calculations, we had only weak-to-moderate study power to replicate the associations of previously-reported risk SNPs chosen for analysis (Table 4). Thus, we emphasize that a major limitation of our risk replication analysis was its relatively small sample size. Nonetheless, for several of the SNPs, our results contribute suggestive evidence of their roles in pancreatic cancer.

We next turned towards the analysis of 24 SNPs that have been previously implicated in modulating overall survival from pancreatic cancer. We focused this analysis on SNPs for which we had >80% to replicate their effects on overall survival. Using a Cox regression model adjusted for race, we observed only one SNP association having a P-value ≤ 0.05 and an estimated HR that was consistent with the direction and magnitude of previous reports: rs9350, a non-synonymous coding SNP in the DNA damage repair gene exonuclease 1 (*EXO1*). Notably, whereas we primarily considered the log-additive effects of minor alleles on overall survival, previous studies of rs9350 considered a dominant model of effect. Specifically, rs9350 was first analyzed by Dong et al. (10) by comparing the survival of CT/TT versus CC genotype groups (HR = 1.89, 95% CI 1.25–2.87). To compare our results more directly with previous studies, we examined differences in estimated survival probability according to rs9350 genotypes after applying a dominant-model grouping scheme (Supplemental Figure 1a). In this setting, we once again observed a significant and consistent effect of SNP rs9350 (CT/TT vs. CC HR = 1.41, 95% CI 1.14–1.74).

Further, in multivariate analysis adjusting for other clinical factors related to overall survival, SNP rs9350 (*EXO1*) remained independently associated. *EXO1* encodes a 5'–3' exonuclease that interacts with several components of the DNA mismatch repair pathway (23). It is currently unknown whether rs9350 might itself be a functional SNP modulating *EXO1* activity, or whether it lies in linkage-disequilibrium (LD) with the true functional

SNP. However, in conjunction with recent follow-up studies by Li and colleagues that found additional SNPs in *EXO1* associated with OS (11), our results provide a third line of evidence suggesting the importance of *EXO1* in pancreatic tumor biology. Additional studies are needed to understand the functional mechanisms by which this gene (and the associated SNPs) are modulating clinical outcome in patients with pancreatic cancer.

By contrast, the direction of our results for SNP rs8191754 (per-allele HR = 0.77, 95% CI 0.62–0.95) was nominally significant ($p < 0.05$) but not consistent with previous reports. Overall, we were unable to replicate the survival associations of most previously-reported SNPs analyzed in this study. Importantly, however, we emphasize that such discordant findings do not directly imply that previous reports were false-positives. Rather, given the complex spectrum of factors that influence survival from pancreatic cancer, we would first speculate that clinical and/or genetic ancestral differences between our patient population and previously-studied populations might explain the discordant results.

In this study, we also assessed whether SNPs involved in pancreatic cancer risk might also be involved in overall survival (and vice versa). This hypothesis was suggested by previous studies that found the same factors could influence both pancreatic cancer risk and survival (*e.g.* self-reported allergies or high BMI) (24–30). We tested SNPs identified by pancreatic cancer risk GWAS for association with overall survival. Conversely, we tested SNPs identified by candidate-gene studies of pancreatic cancer survival for risk association. Our results, however, do not provide strong evidence that either phenomenon is occurring for these sets of SNPs.

Finally, using an unbiased genome-wide approach, we sought to identify preliminary evidence of novel SNPs associated with the clinical outcome of pancreatic cancer. Putative associations were identified by GWAS of overall survival and then validated using independent case groups. This approach yielded at least one novel putative association at SNP rs1482426, which lies in a non-genic region of chromosome 12q21. Interestingly, the gene *SLC6A15* (solute carrier family 6, neutral amino acid transporter member 15) is located ~730kb downstream of rs1482426, adjacent to the linkage disequilibrium region containing rs1482426, and was recently found to be somatically mutated in two sequenced pancreatic cancer exomes (31).

However, we emphasize that our GWAS results should be interpreted as preliminary. Additional studies are needed to confirm this association in larger patient samples and to understand the functional significance of this region in pancreatic cancer survival. Importantly, although the direction of its effect on overall survival was consistent throughout all three stages of our GWAS, rs1482426 was estimated to have a much weaker effect (HR ≈ 1.2) in stage 3 compared to stages 1 and 2 (per-allele HR ≈ 1.7). Several factors could have led to this disparity. For example, it is possible that the effect observed in stages 1 and 2 was an overestimate of rs1482426's true effect size in the population — i.e., the winner's curse phenomenon (32). In addition, the patients analyzed in stages 1 and 2 were based at MSKCC, while patients analyzed in stage 3 were based at the Mayo Clinic. Thus, we also speculate that differences in clinical and/or demographic characteristics between these two patient samples may partially explain the discordant effect sizes. Lastly, we note that the overall results of our GWAS were not strictly significant at the genome-wide level. Therefore, we cannot rule out the possibility that the observed effect of rs1482426 on survival was a false-positive finding.

Nonetheless, towards the goal of identifying novel loci that play a role in pancreatic cancer outcome, our study demonstrates both the potential utility of GWAS and some of the challenges faced in its design. Ideally, a large multi-center consortium study like PanScan

would be well powered to identify robust and novel associations of SNPs with overall survival. However, that approach would be faced with the complex challenge of merging clinical data across the various centers. Here, at the cost of overall study power, we chose to focus the discovery and validation stages of our GWAS on patients treated at a single center (MSKCC), followed by replication in patients from an independent center (Mayo Clinic).

In conclusion, we have provided evidence of independent replication for several previously-reported SNPs associated with pancreatic cancer risk and overall survival. Further, we used an unbiased, genome-wide approach to identify a novel locus putatively associated with overall survival. Our study adds further supporting evidence that inherited genetic variations may play important biological roles in pancreatic cancer.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We are very grateful to the patients and their families who participated in this study. We sincerely thank Dr. Mithat Gonën for helpful advice regarding study design and manuscript preparation; Dr. Laetitia Borsu, Angela Marchetti (MSKCC Sequenom Facility),and Xiaoni Gao for their assistance with Sequenom assay processing; Drs. Agnes Viale and Jeffrey Zhao (MSKCC Genomics Core Facility) for assistance with array-based genotyping; Sarah Yoo and Pampa Roy (Molecular Epidemiology Laboratory) for their work with DNA extraction and sample processing; and William Bamlet (Mayo Clinic) for biostatistical analysis. We are very grateful to Dr. Gloria Petersen (Mayo Clinic) for allowing us to use the Mayo Clinic Pancreas SPORE Registry.

GRANT SUPPORT

This work was supported by the Geoffrey Beene Cancer Research Center at MSKCC and the Emerald Foundation. J.A.W. was partially supported by National Institutes of Health (NIH) grant GM07739. The MSKCC Sequenom Facility is supported by the Anbinder Fund. The Mayo Clinic Pancreas registry is supported by NIH grant SPORE P50 CA 102701 (PI G. Petersen).

REFERENCES

1. Jemal A, Siegel R, Ward E, Hao Y, Xu J, Thun MJ. Cancer Statistics, 2009. CA Cancer J Clin. 2009; 59(4):225–249. [PubMed: 19474385]
2. Jemal A, Siegel R, Xu J, Ward E. Cancer statistics, 2010. CA Cancer J Clin. 2010; 60(5):277–300. [PubMed: 20610543]
3. Amundadottir L, Kraft P, Stolzenberg-Solomon RZ, Fuchs CS, Petersen GM, Arslan AA, et al. Genome-wide association study identifies variants in the ABO locus associated with susceptibility to pancreatic cancer. Nat Genet. 2009
4. Petersen GM, Amundadottir L, Fuchs C, Kraft P, Stolzenberg-Solomon RZ, Jacobs KB, et al. A genome-wide association study identifies pancreatic cancer susceptibility loci on chromosomes 13q22.1, 1q32.1 and 5p15.33. Nat Genet. 2010; 42(3):224–228. [PubMed: 20101243]
5. Low S-K, Kuchiba A, Zembutsu H, Saito A, Takahashi A, Kubo M, et al. Genome-Wide Association Study of Pancreatic Cancer in Japanese Population. PLoS ONE. 2010; 5(7) e11824 EP -.
6. Chen J, Li D, Killary AM, Sen S, Amos CI, Evans DB, et al. Polymorphisms of p16, p27, p73, and MDM2 modulate response and survival of pancreatic cancer patients treated with preoperative chemoradiation. Ann Surg Oncol. 2009; 16(2):431–439. [PubMed: 19020940]
7. Cotton RT, Li D, Scherer SE, Muzny DM, Hodges SE, Catania RL, et al. Single nucleotide polymorphism in RECQL and survival in resectable pancreatic adenocarcinoma. HPB (Oxford). 2009; 11(5):435–444. [PubMed: 19768149]
8. Couch FJ, Wang X, Bamlet WR, de Andrade M, Petersen GM, McWilliams RR. Association of mitotic regulation pathway polymorphisms with pancreatic cancer risk and outcome. Cancer Epidemiol Biomarkers Prev. 2010; 19(1):251–257. [PubMed: 20056645]

9. Dong X, Javle M, Hess KR, Shroff R, Abbruzzese JL, Li D. Insulin-like Growth Factor Axis Gene Polymorphisms and Clinical Outcomes in Pancreatic Cancer. Gastroenterology. 2010
10. Dong X, Jiao L, Li Y, Evans DB, Wang H, Hess KR, et al. Significant associations of mismatch repair gene polymorphisms with clinical outcome of pancreatic cancer. J Clin Oncol. 2009; 27(10):1592–1599. [PubMed: 19237629]
11. Dong X, Li Y, Hess KR, Abbruzzese JL, Li D. DNA mismatch repair gene polymorphisms affect survival in pancreatic cancer. Oncologist. 2011; 16(1):61–70. [PubMed: 21212431]
12. Halfdanarson TR, Wang L, Bamlet WR, de Andrade M, McWilliams RR, Cunningham JM, et al. Mitochondrial genetic polymorphisms do not predict survival in patients with pancreatic cancer. Cancer Epidemiol Biomarkers Prev. 2008; 17(9):2512–2513. [PubMed: 18768524]
13. Li D, Frazier M, Evans DB, Hess KR, Crane CH, Jiao L, et al. Single nucleotide polymorphisms of RecQ1, RAD54L, and ATM genes are associated with reduced survival of pancreatic cancer. J Clin Oncol. 2006; 24(11):1720–1728. [PubMed: 16520463]
14. Li D, Li Y, Jiao L, Chang DZ, Beinart G, Wolff RA, et al. Effects of base excision repair gene polymorphisms on pancreatic cancer survival. Int J Cancer. 2007; 120(8):1748–1754. [PubMed: 17230526]
15. Li D, Liu H, Jiao L, Chang DZ, Beinart G, Wolff RA, et al. Significant effect of homologous recombination DNA repair gene polymorphisms on pancreatic cancer survival. Cancer Res. 2006; 66(6):3323–3330. [PubMed: 16540687]
16. Okazaki T, Javle M, Tanaka M, Abbruzzese JL, Li D. Single nucleotide polymorphisms of gemcitabine metabolic genes and pancreatic cancer survival and drug toxicity. Clin Cancer Res. 2010; 16(1):320–329. [PubMed: 20028759]
17. Okazaki T, Jiao L, Chang P, Evans DB, Abbruzzese JL, Li D. Single-nucleotide polymorphisms of DNA damage response genes are associated with overall survival in patients with pancreatic cancer. Clin Cancer Res. 2008; 14(7):2042–2048. [PubMed: 18381943]
18. McWilliams RR, Matsumoto ME, Burch PA, Kim GP, Halfdanarson TR, de Andrade M, et al. Obesity adversely affects survival in pancreatic cancer patients. Cancer. 2010; 116(21):5054–5062. [PubMed: 20665496]
19. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al. PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. The American Journal of Human Genetics. 2007; 81(3):559–575.
20. Skol AD, Scott LJ, Abecasis GR, Boehnke M. Joint analysis is more efficient than replication-based analysis for two-stage genome-wide association studies. Nat Genet. 2006; 38(2):209–213. doi: http://www.nature.com/ng/journal/v38/n2/suppinfo/ng1706_S1.html. [PubMed: 16415888]
21. Azzato EM, Greenberg D, Shah M, Blows F, Driver KE, Caporaso NE, et al. Prevalent cases in observational studies of cancer survival: do they bias hazard ratio estimates? Br J Cancer. 2009; 100(11):1806–1811. [PubMed: 19401693]
22. Schoenfeld DA. Sample-size formula for the proportional-hazards regression model. Biometrics. 1983; 39(2):499–503. [PubMed: 6354290]
23. Zhang Y, Yuan F, Presnell SR, Tian K, Gao Y, Tomkinson AE, et al. Reconstitution of 5'-directed human mismatch repair in a purified system. Cell. 2005; 122(5):693–705. [PubMed: 16143102]
24. Olson SH, Chou JF, Ludwig E, O'Reilly E, Allen PJ, Jarnagin WR, et al. Allergies, obesity, other risk factors and survival from pancreatic cancer. International journal of cancer Journal international du cancer. 2010
25. Olson SH, Orlow I, Simon J, Tommasi D, Roy P, Bayuga S, et al. Allergies, variants in IL-4 and IL-4R alpha genes, and risk of pancreatic cancer. Cancer Detect Prev. 2007; 31(5):345–351. [PubMed: 18031948]
26. Turner MC, Chen Y, Krewski D, Ghadirian P. An overview of the association between allergy and cancer. International journal of cancer Journal international du cancer. 2006; 118(12):3124–3132. [PubMed: 16395696]
27. Gandini S, Lowenfels AB, Jaffee EM, Armstrong TD, Maisonneuve P. Allergies and the risk of pancreatic cancer: a meta-analysis with review of epidemiology and biological mechanisms. Cancer Epidemiol Biomarkers Prev. 2005; 14(8):1908–1916. [PubMed: 16103436]

28. Li D, Morris JS, Liu J, Hassan MM, Day RS, Bondy ML, et al. Body mass index and risk, age of onset, and survival in patients with pancreatic cancer. JAMA. 2009; 301(24):2553–2562. [PubMed: 19549972]

29. Fleming JB, Gonzalez RJ, Petzel MQB, Lin E, Morris JS, Gomez H, et al. Influence of obesity on cancer-related outcomes after pancreatectomy to treat pancreatic adenocarcinoma. Arch Surg. 2009; 144(3):216–221. [PubMed: 19289659]

30. McWilliams, RR.; Bamlet, WR.; Matsumoto, ME.; Petersen, GM.; Halfdanarson, TR. Correlation of high usual adult body mass index with survival in pancreatic adenocarcinoma; 2009 Gastrointestinal Cancers Symposium; 2009.

31. Jones S, Zhang X, Parsons DW, Lin JC, Leary RJ, Angenendt P, et al. Core signaling pathways in human pancreatic cancers revealed by global genomic analyses. Science. 2008; 321(5897):1801–1806. [PubMed: 18772397]

32. Lohmueller KE, Pearce CL, Pike M, Lander ES, Hirschhorn JN. Meta-analysis of genetic association studies supports a contribution of common variants to susceptibility to common disease. Nat Genet. 2003; 33(2):177–182. [PubMed: 12524541]

STATEMENT OF TRANSLATIONAL RELEVANCE

The identification of inherited genetic variations that influence susceptibility to and the clinical outcome of pancreatic cancer may yield novel insights about the development and progression of this highly fatal disease. In this study, we provide supporting evidence that genetic variants located on chromosomes 9q34 (*ABO*) and 13q22 are involved in pancreatic cancer susceptibility, and that a variant within the *EXO1* gene may be involved in overall survival from pancreatic cancer. More importantly, using genome-wide methods, we putatively identify a novel locus on chromosome 12q21 associated with overall survival. Functional analysis of these regions may lead to the development of new strategies for risk prediction and treatment of pancreatic cancer.

Table 1

Summary of the overall study design and results.

SNPs Analyzed	Study Samples	Findings: Risk Analysis	Findings: Survival Analysis
Seven (7) SNPs identified by previous genome-wide association study of pancreatic cancer **risk** (refs. 3–5)	531 cases, 305 controls (MSKCC)	Independent replication of the risk associations of SNPs: rs505922 (chr 9q34, *ABO*) rs9543325 (chr 13q22)	no significant findings
Twenty-four (24) SNPs identified by **previous candidate-gene/pathway studies** of pancreatic cancer clinical **outcome** (refs. 6–17)	513 cases (MSKCC)	no significant findings	Independent replication of the overall survival association of SNP: rs9350 (chr 1q43, *EXO1*)
Genome-wide association study of overall survival from pancreatic cancer (this study)		not performed	Putative identification of a novel SNP associated with overall survival: rs1482426 (chr 12q21)
Stage 1 (Discovery): 301,250 SNPs	252 cases (MSKCC)		
Stage 2 (Validation): 22 SNPs	261 cases (MSKCC)		
Stage 3 (Replication): 1 SNP	572 cases (Mayo Clinic)		

Table 2

Demographics of the MSKCC case-control study population.

	Cases	Controls
	n = 531	**n = 305**
	n (%)	**n (%)**
Gender		
Male	290 (55)	127 (42)
Female	241 (45)	178 (58)
Age		
≤50	83 (16)	80 (26)
51–60	129 (24)	86 (28)
61–70	184 (35)	100 (33)
>70	135 (25)	39 (13)
Race/Ethnicity		
White/Caucasian	495 (93)	283 (93)
Black/African-American	21 (4)	10 (3)
Asian (East)	10 (2)	11 (4)
Asian (Indian)	5 (1)	1 (0)

Table 3

Clinical characteristics of MSKCC cases utilized in survival analysis.

	Patient Subset	
	Stage 1: GWAS	**Stage 2: Validation**
	n = 252	**n = 261**
	n (%)	**n (%)**
Gender		
Male	141 (56)	141 (54)
Female	111 (44)	120 (46)
Race/Ethnicity		
White/Caucasian	252 (100)	226 (87)
Black/African-American	0 (0)	21 (8)
Asian (East)	0 (0)	10 (4)
Asian (Indian)	0 (0)	4 (2)
Surgical Resection		
Yes	102 (40)	109 (42)
No	150 (60)	152 (58)
ECOG Performance Status		
Less than or equal to 1	203 (81)	208 (80)
Greater than 1	11 (4)	15 (6)
Missing	38 (15)	38 (15)
Clinical Stage		
Localized	89 (35)	100 (38)
Locally Advanced	65 (26)	56 (21)
Metastatic	87 (35)	82 (31)
Missing	11 (4)	23 (9)
Chemotherapy		
Yes	232 (92)	234 (90)
No	20 (8)	27 (10)
Radiation		
Yes	90 (36)	89 (34)
No	162 (64)	172 (66)

Table 4

Replication of pancreatic cancer risk associations for SNPs identified by previous GWAS.

SNP, Minor Allele, Chr, Genes,	Ref.	Subjects		MAF		Reported OR (95% CI)	Estimated OR[a] (95% CI)	P-value[b]	Power[c]
		Controls	Cases	Controls	Cases				
rs505922 * C 9q34 *ABO*	(3)	149	385	0.34	0.43	1.20 (1.12–1.28)	1.65 (1.20–2.26)	0.002	26%
rs9543325 * G 13q22.1 none	(4)	148	386	0.40	0.49	1.26 (1.18–1.35)	1.42 (1.06–1.90)	0.02	39%
rs708224 A 12p11.21 *BICD1*	(5)	303	525	0.38	0.42	1.32 (1.19–1.47)	1.21 (0.98–1.50)	0.08	77%
rs9502893 G 6p25.3 *FOXQ1*	(5)	303	528	0.42	0.46	1.29 (1.17–1.43)	1.20 (0.97–1.48)	0.09	70%
rs401681 * A 5p15.33 *CLPTM1L*	(4)	149	390	0.43	0.48	1.19 (1.11–1.27)	1.22 (0.91–1.62)	0.18	25%
rs6464375 A 7q36.2 *DPP6*	(5)	305	531	0.05	0.05	3.73 (2.24–6.21) +	0.44 (0.04–5.39) +	0.52	9%
rs3790844 * G 1q32.1 *NR5A2*	(4)	152	392	0.24	0.21	0.77 (0.71–0.84)	0.93 (0.66–1.30)	0.66	33%

[a] multivariable logistic regression model adjusted for race, gender, and age;

[b] 1-d.f. Wald test for significance;

[c] Power calculation based on a disease prevalence of 0.01% in the population, the reported effect size, the MAF observed in controls, the indicated sample size, and significance level of 0.05

* SNP was analyzed after excluding cases and controls who participated in the PanScan study

+ SNP was analyzed under a recessive model consistent with its original report in Ref. 5.

Abbreviations: OR = odds ratio; CI = confidence interval; Ref. = reference; Chr. = chromosome; MAF = minor allele frequency

Table 5

Replication of overall survival associations for SNPs identified by previous candidate-gene studies of pancreatic cancer clinical outcome.

SNP	Ref.	Gene	Minor Allele	Estimated HR[a] (95% CI)	P-value[b]
rs9350	(10)	*EXO1*	T	1.26 (1.07–1.50)	0.007
rs8191754	(9)	*IGF2R*	G	0.77 (0.62–0.95)	0.02
rs8041224	(9)	*IGF1R*	T	1.15 (0.99–1.33)	0.06
rs12090453	(8)	*GPSM2*	C	1.15 (0.99–1.33)	0.07
rs2854744	(9)	*IGFBP3*	A	1.14 (0.99–1.31)	0.07
rs735943	(10)	*EXO1*	A	0.88 (0.76–1.02)	0.08
rs2272615	(14)	*POLB*	G	1.19 (0.94–1.51)	0.14
rs2946834	(9)	*IGF1*	T	0.91 (0.77–1.08)	0.28
rs3218536	(15)	*XRCC2*	T	0.88 (0.69–1.13)	0.31
rs2953993	(14)	*POLB*	A	1.16 (0.86–1.55)	0.33
rs2134808	(8)	*TUBG1*	C	0.93 (0.80–1.08)	0.34
rs302864	(8)	*TEX14*	T	1.14 (0.86–1.50)	0.36
rs2431238	(8)	*APC*	T	1.05 (0.91–1.22)	0.49
rs293794	(14)	*hOGG1*	C	1.07 (0.87–1.31)	0.52
rs664143	(17)	*ATM*	T	0.95 (0.81–1.11)	0.53
rs1805355	(10)	*MSH3*	A	0.94 (0.73–1.21)	0.61
rs3743262	(9)	*IGF1R*	T	0.92 (0.67–1.28)	0.63
rs12437963	(9)	*IGF1R*	G	0.96 (0.78–1.17)	0.67
rs664677	(13)	*ATM*	C	0.97 (0.83–1.14)	0.74
rs2066827	(6)	*p27*	C	0.98 (0.82–1.16)	0.78
rs7928320	(8)	*KIAA0999*	T	0.98 (0.81–1.18)	0.85
rs11079571	(8)	*AXIN2*	A	1.02 (0.84–1.23)	0.88
rs521102	(17)	*CHEK1*	T	1.00 (0.86–1.17)	0.95
rs5742933	(10)	*PMS1*	C	1.00 (0.84–1.18)	0.96

[a] Cox proportional hazards model adjusted for race;

[b] 1-d.f. Wald test for significance

Abbreviations: HR = per-allele hazard ratio; Ref = reference

Table 6

Summarized results from stages 1 and 2 of the GWAS of overall survival from pancreatic cancer.

	Discovery		Validation		Combined	
SNP, Minor Allele, Chromosome, Gene[a]	**HR[b] (95% CI)**	**P-value[c]**	**HR[d] (95% CI)**	**P-value[c]**	**HR[d] (95% CI)**	**P-value[c]**
rs1482426 G 12q21 -	1.67 (1.16–2.41)	3.6×10^{-5}	1.73 (1.25–2.40)	0.001	1.74 (1.38–2.18)	1.7×10^{-6}
rs4285214 G 5q23 ZNF608	1.65 (1.35–2.02)	7.9×10^{-7}	1.09 (0.90–1.31)	0.39	1.33 (1.16–1.53)	4.1×10^{-5}
rs7849571 G 9p21 -	1.63 (1.28–2.07)	7.2×10^{-5}	1.14 (0.90–1.43)	0.27	1.35 (1.14–1.59)	3.7×10^{-4}
rs11151040 T 18q23 -	2.57 (1.61–4.10)	7.6×10^{-5}	1.27 (0.76–2.15)	0.36	1.84 (1.30–2.61)	5.6×10^{-4}
rs3747572 C 16p13 GLIS2	1.83 (1.39–2.40)	1.3×10^{-5}	1.10 (0.86–1.42)	0.44	1.38 (1.15–1.66)	6.2×10^{-4}
rs10189511 A 2q34 -	2.11 (1.51–2.95)	1.2×10^{-5}	1.07 (0.77–1.48)	0.70	1.50 (1.19–1.89)	6.7×10^{-4}
rs1344963 T 12q21 -	2.59 (1.69–3.97)	1.4×10^{-5}	1.18 (0.85–1.66)	0.32	1.56 (1.20–2.03)	8.5×10^{-4}
rs4903736 A 14q24 -	1.59 (1.27–1.98)	4.5×10^{-5}	1.02 (0.78–1.32)	0.90	1.33 (1.12–1.58)	9.8×10^{-4}

SNP, Minor Allele, Chromosome, Gene[a]	Discovery		Validation		Combined	
	HR[b] (95% CI)	P-value[c]	HR[d] (95% CI)	P-value[c]	HR[d] (95% CI)	P-value[c]
rs11024097 T 11p15 PLEKHA7	1.54 (1.25–1.90)	4.2×10^{-5}	1.01 (0.81–1.26)	0.91	1.28 (1.10–1.49)	0.001
rs4903741 C 14q24 -	1.57 (1.25–1.96)	8.7×10^{-5}	1.04 (0.83–1.31)	0.72	1.29 (1.10–1.52)	0.002
rs12835268 C Xp21 -	1.50 (1.24–1.82)	3.4×10^{-5}	1.05 (0.86–1.28)	0.64	1.23 (1.08–1.41)	0.003
rs16867625 T 8q22 -	2.33 (1.55–3.51)	5.0×10^{-5}	1.15 (0.75–1.75)	0.52	1.55 (1.16–2.08)	0.003
rs7016046 A 8q24 -	1.83 (1.38–2.43)	3.1×10^{-5}	1.02 (0.76–1.38)	0.89	1.36 (1.11–1.67)	0.004
rs2056096 T 18q12 -	1.56 (1.20–2.03)	7.3×10^{-5}	0.97 (0.78–1.20)	0.76	1.24 (1.07–1.44)	0.005
rs8034546 T 15q14 -	1.62 (1.29–2.04)	4.0×10^{-5}	0.97 (0.75–1.26)	0.85	1.27 (1.07–1.51)	0.005
rs1867348 T 6q25 IGF2R	2.02 (1.43–2.85)	6.3×10^{-5}	0.98 (0.70–1.37)	0.90	1.38 (1.08–1.75)	0.009
rs956518 G 4p15 -	0.62 (0.50–0.79)	5.7×10^{-5}	1.13 (0.92–1.40)	0.24	0.82 (0.70–0.95)	0.01

SNP, Minor Allele, Chromosome, Gene[a]	Discovery HR[b] (95% CI)	Discovery P-value[c]	Validation HR[d] (95% CI)	Validation P-value[c]	Combined HR[d] (95% CI)	Combined P-value[c]
rs10167103 C 2q14 -	1.53 (1.25–1.88)	5.0×10^{-5}	0.92 (0.76–1.13)	0.43	1.20 (1.04–1.38)	0.01
rs38402 G 7p15 GGCT	4.75 (2.19–10.31)	8.2×10^{-5}	1.25 (0.65–2.41)	0.50	1.86 (1.13–3.05)	0.01
rs3795244 T 17q11 ZNF207	2.46 (1.69–3.58)	3.0×10^{-6}	0.89 (0.61–1.29)	0.53	1.34 (1.03–1.74)	0.03
rs1944395 C 18q12 -	1.50 (1.23–1.84)	8.3×10^{-5}	0.86 (0.70–1.06)	0.16	1.16 (1.01–1.35)	0.04
rs1388193 T 13q31 -	1.77 (1.34–2.34)	6.5×10^{-5}	0.72 (0.53–0.99)	0.04	1.13 (0.92–1.38)	0.25

[a] closest RefSeq annotated gene within +/− 20kb;

[b] univariate Cox proportional hazards model;

[c] 1-d.f. Wald test for significance;

[d] Cox proportional hazards model adjusted for race;

Abbreviations: HR = per-allele hazard ratio; CI = confidence interval