




# Validation of genome-wide association study-identified single nucleotide polymorphisms in a case-control study of pancreatic cancer from Taiwan

Yan-Shen Shan<sup>1,2†</sup>, Li-Tzong Chen<sup>3,4,5,6†</sup>, Jin-Shang Wu<sup>7†</sup>, Yin-Fan Chang<sup>7</sup>, Chih-Ting Lee<sup>7</sup>, Chih-Hsing Wu<sup>7</sup>, Nai-Jung Chiang<sup>3,4</sup>, Hsin-En Huang<sup>7</sup>, Chia-Jui Yen<sup>4</sup>, Ying-Jui Chao<sup>1</sup>, Hui-Jen Tsai<sup>3,4</sup>, Chiung-Yu Chen<sup>4</sup>, Jui-Wen Kang<sup>4</sup>, Chin-Fu Kuo<sup>8</sup>, Chia-Rung Tsai<sup>3</sup>, Ya-Ling Weng<sup>3</sup>, Han-Chien Yang<sup>3</sup>, Hui-Chin Liu<sup>9</sup> and Jeffrey S. Chang<sup>3\*</sup> 

## Abstract

**Background:** Due to differences in genetic background, it is unclear whether the genetic loci identified by the previous genome-wide association studies (GWAS) of pancreatic cancer also play significant roles in the development of pancreatic cancer among the Taiwanese population.

**Methods:** This study aimed to validate the 25 pancreatic cancer GWAS-identified single nucleotide polymorphisms (SNPs) in a case-control study (278 cases and 658 controls) of pancreatic cancer conducted in Taiwan. Statistical analyses were conducted to determine the associations between the GWAS-identified SNPs and pancreatic cancer risk. Gene-environment interaction analysis was conducted to evaluate the interactions between SNPs and environmental factors on pancreatic cancer risk.

**Results:** Among the 25 GWAS-identified SNPs, 7 (rs2816938 (~ 11 kb upstream of *NR5A2*), rs10094872 (~ 28 kb upstream of *MYC*), rs9581943 (200 bp upstream of *PDX1*) and 4 chromosome 13q22.1 SNPs: rs4885093, rs9573163, rs9543325, rs9573166) showed a statistically significant association with pancreatic cancer risk in the current study. Additional analyses showed two significant gene-environment interactions (between poor oral hygiene and *NR5A2* rs2816938 and between obesity and *PDX1* rs9581943) on the risk of pancreatic cancer.

**Conclusions:** The current study confirmed the associations between 7 of the 25 GWAS-identified SNPs and pancreatic risk among the Taiwanese population. Furthermore, pancreatic cancer was jointly influenced by lifestyle and medical factors, genetic polymorphisms, and gene-environment interaction. Additional GWAS is needed to determine the genetic polymorphisms that are more relevant to the pancreatic cancer cases occurring in Taiwan.

(Continued on next page)

\* Correspondence: [jeffreychang@nhri.org.tw](mailto:jeffreychang@nhri.org.tw)

<sup>†</sup>Yan-Shen Shan, Li-Tzong Chen and Jin-Shang Wu contributed equally to this work.

<sup>3</sup>National Institute of Cancer Research, National Health Research Institutes, 1F No 367, Sheng-Li Road, Tainan 70456, Taiwan

Full list of author information is available at the end of the article



© The Author(s). 2020 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

(Continued from previous page)

**Keywords:** Genome-wide association, Epidemiology and prevention, Pancreatic cancer, Gene-environment interaction

## Introduction

According to the International Agency for Research on Cancer (IARC), approximately 458,918 cases (incidence rate = 4.8 per 100,000) of pancreatic cancer are diagnosed worldwide annually [1]. The survival of pancreatic cancer is extremely poor with a 5-year survival rate of only 9% [2], resulting in the high number of deaths (approximately 432,242 deaths per year) that almost equals the number of incident cases [1].

Despite the numerous epidemiologic investigations of pancreatic cancer, the only established environmental risk factors (defined as factors other than genetic factors such as genetic polymorphisms or mutations) with strong evidence are cigarette smoking and chronic diabetes [3, 4]. Environmental factors that have shown a strong association with an increased pancreatic cancer risk include obesity and chronic pancreatitis [5–7], while allergies have been consistently associated with a decreased pancreatic cancer risk [8]. Other environmental factors that may possibly influence the risk of pancreatic cancer but require more investigations are diet, physical exercise, oral hygiene/health, and alcohol use. Diet rich in fruits and vegetables has been associated with a reduced pancreatic cancer risk while diet with higher levels of meat and animal products has been linked to an increased pancreatic cancer risk [9–11]. A reduced risk of pancreatic cancer has been linked to higher levels of physical activities, although results have not been consistent across studies [12–15]. Huang et al. used the number of teeth, dental plaque, and oral mucosal lesions as indicators of oral hygiene status and reported that lower number of teeth, higher level of dental plaque, and presence of oral mucosal lesions were associated with a higher risk of pancreatic cancer [16]. For alcohol drinking, studies suggested that only high amount of alcohol consumption may increase the risk of pancreatic cancer [14, 17, 18].

Only five to 10 % of pancreatic cancers are familial with strong genetic predisposition due to rare genetic mutations [19, 20]. The majority of the pancreatic cancer cases are sporadic with an unclear level of genetic contribution. Numerous candidate gene studies that selected genetic polymorphisms based on a priori hypothesis have reported genetic polymorphisms associated with pancreatic cancer risk, including polymorphisms of DNA repair genes, carcinogen-metabolizing genes, folate-metabolizing genes, and alcohol-metabolizing gene. These studies were documented by a systematic review and meta-analysis by Dai et al. [21]. According to their study, many of the

polymorphisms were investigated by only 1 or 2 studies and therefore the reliability of the results required further validation. They performed meta-analysis on polymorphisms that had been investigated by three or more studies, which focused mainly on the DNA repair and folate-metabolizing genes. Their analysis found that polymorphisms of the DNA repairs genes, including *XRCC1* Arg399Gln and Arg194Trp, *ERCC1* rs11615 and rs3212986, and *ERCC2* rs13181, showed significant association with pancreatic cancer risk. No significant association was found between the C677T and A1298C polymorphisms of the folate-metabolizing gene, *MTHFR*, and pancreatic cancer risk [21]. Because most candidate gene studies were conducted with a relatively small sample size, the results across studies were often inconsistent and not replicable.

In contrast to candidate gene studies, genome-wide association studies (GWAS) are usually performed with a large sample size consisting of discovery and replication groups and therefore the results of GWAS are usually considered more valid than those generated by the candidate gene studies. GWAS have identified genetic polymorphisms that are associated with the risk of pancreatic cancer [22–33]. These studies, conducted among different racial/ethnic groups, have identified genetic loci that are common across and unique to different racial/ethnic groups. Due to differences in genetic background, it is unclear whether these genetic loci identified by the previous GWAS of pancreatic cancer also play significant roles in the development of pancreatic cancer among the Taiwanese population. The current study aimed to validate the pancreatic cancer GWAS-identified single nucleotide polymorphisms (SNPs) in a case-control study of pancreatic cancer conducted in Taiwan.

Previous studies have reported significant gene-environment interactions on the risk of pancreatic cancer. For example, the increased pancreatic cancer risk due to cigarette smoking was modified by polymorphisms of carcinogen-metabolizing, folate-metabolizing, and DNA repair genes [34–36]. An interaction between diabetes and the polymorphisms of hexokinase 2 (*HK2*), a glucose metabolism gene, on the risk of pancreatic cancer was reported, with *HK2* R844K GA/AA genotype showing an inverse association with pancreatic cancer among individuals without diabetes, but a positive association with pancreatic cancer among diabetic patients [37]. These studies indicated that pancreatic cancer could result from the interplay between environmental

and genetic factors; therefore, the current study also performed exploratory analysis to evaluate the interaction between the GWAS-identified SNPs and the environmental factors on the risk of pancreatic cancer.

### Materials and methods

The current study was approved by institutional review boards of the National Cheng Kung University Hospital (ethic approval number: B-BR-102-070) and the National Health Research Institutes (ethic approval number: EC1030109-E). A signed informed consent was collected from every study participant who agreed to participate in the study.

#### Subject recruitment

All subjects were from an ongoing case-control study of pancreatic cancer conducted at the National Cheng Kung University Hospital. Cases were recruited from the Department of General Surgery or the Division of Hemato-Oncology, Department of Internal Medicine. The eligible criteria for the cases were: 1) diagnosis of pancreatic ductal adenocarcinoma 2) no history of previous cancer diagnosis; 3) age = 20 years or more; and 4) the ability to understand the purpose of the study and provide informed consent. Control subjects were recruited from the Department of Family Medicine and the eligibility criteria were: 1) the purpose of the clinical visit was for regular physical examination, vaccination, or for minor illnesses (e.g. common cold, headache) not related to cigarette smoking or metabolic diseases (diabetes, hypertension, hyperlipidemia); 2) no history of previous cancer diagnosis; 3) age = 20 years or more; and 4) the ability to understand the purpose of the study and provide informed consent. The current analysis included subjects recruited from November 19, 2013 to February 7, 2018.

#### In-person interview

Each study subject was interviewed in-person by a trained interviewer using a standardized questionnaire to collect information on: 1) demographic information, including age, sex, and education; 2) lifestyle factors, including cigarette smoking, alcohol drinking, and diet; 3) medical history, including history of allergy and diabetes/glucose intolerance; 4) oral hygiene habits, including regular dental visits, frequency of tooth brushing, and use of dental floss; and 5) biometric measurements, including height and weight 2 years before the pancreatic cancer diagnosis for the cases or before the interview date for the controls. We asked about the weight 2 years before the diagnosis for the cases because we wanted to collect the weight information before the development of pancreatic cancer to avoid reverse causation.

#### DNA sample collection and processing

DNA samples were obtained from blood or buccal swab for those who could not provide blood samples. Blood samples were collected in a vacutainer tube containing EDTA (lavender-top). Buccal swab samples were obtained by gently brushing the buccal mucosa with FLOQSwabs (Copan Flock Technologies, Brescia, Italy). Blood samples were centrifuged to separate out the buffy coat. Genomic DNA was extracted from the buffy coat and the buccal swab samples using a commercially available DNA purification kit. DNA samples were stored in the  $-80^{\circ}\text{C}$  refrigerator until ready to use.

#### SNP selection

Literature search was conducted to identify pancreatic cancer GWAS articles and related articles published by December 31, 2017. A total of 12 articles were identified (Supplementary Table 1). Eight of the 12 articles included all or mostly subjects with European ancestry and 4 articles included only Asian study subjects (two were Chinese studies and two were Japanese studies). SNPs that showed statistically significant association with pancreatic cancer and were successfully validated in another independent group were identified from these articles. The minor allele frequencies of these SNPs in the East Asian population recorded by the NCBI dbSNP database (<https://www.ncbi.nlm.nih.gov/snp/>) were reviewed. SNPs with a minor allele frequency  $< 0.05$  were excluded. Twenty-five SNPs were selected for genotyping for the current study. Detailed information for these 25 SNPs can be found in Supplementary Table 1.

#### Genotyping

Genotyping of the 25 SNPs was performed using the MassArray System (Agena Bioscience, San Diego, CA, USA), which is a mass spectrometry-based detection system for medium-throughput genotyping. All of the 25 SNPs had a call rate of more than 98%. Samples with a call rate of  $< 90\%$  were excluded from the analysis.

#### Statistical analysis

The average age was compared between cases and controls using t-test. The distributions of sex and education levels were compared between cases and controls using chi-squared test. For analysis with the environmental (non-genetic) factors, we only analyzed factors with strong evidence for association with pancreatic cancer according to the current literature [38], including cigarette smoking, oral hygiene, consumption of vegetables and fruits, allergy, diabetes/glucose intolerance, and body mass index (BMI). First, unconditional logistic regression analysis was conducted to estimate the odds ratio (OR) and 95% confidence interval (CI) to determine the association between each environmental factor and

the risk of pancreatic cancer, adjusted for sex, age, and education. Subsequently, another model including all of the above-mentioned environmental factors plus adjustment for sex, age, and education was constructed to assess the independent contribution of each environmental factor while adjusting for the potential confounding effect of the other environmental factors. Cigarette smoking was analyzed as ever (current + former) vs. never. An oral hygiene score was created to evaluate oral hygiene. The oral hygiene score = regular dental visit + frequency of tooth brushing + use of dental floss: regular dental visit: yes = 0, no = 1; frequency of tooth brushing:  $\geq 2$  times per day = 0,  $< 2$  times per day = 1; and use of dental floss: yes = 0, no = 1. A higher oral hygiene score indicates poorer oral hygiene. We have used this oral hygiene score previously in several studies of head and neck cancer [39–41]. Furthermore, we have shown that a higher oral hygiene score was correlated with a higher percentage of periodontopathogenic bacteria in the saliva, thus validating the use of the oral hygiene score as an indicator of oral hygiene status [40]. Consumption of vegetables and fruits was evaluated according to the portion and the frequency of intake. Diabetes/glucose intolerance was evaluated according to the time of onset with  $< 2$  years being recent onset and  $\geq 2$  years being chronic. BMI was calculated with the formula: (weight in kilograms)/(height in meters)<sup>2</sup>. BMI was divided into four categories according to the guideline set by the Ministry of Health and Welfare, Taiwan:  $< 18.5$  = underweight; 18.5–23.9: normal; 24–26.9 = overweight; and 27 or more = obese.

Unconditional logistic regression analysis was conducted to estimate the OR and 95% CI for the association between each SNP and the risk of pancreatic cancer, adjusted for sex, age, and education. SNPs with a  $p < 0.05$  were considered statistically significant. Linkage disequilibrium (LD) analysis was performed to identify the statistically significant SNPs in high LD ( $r^2 > 0.8$ ) and the SNP with the lowest  $p$  among the highly-linked SNPs was chosen to perform further analyses.

Exploratory gene-environment interaction analysis was performed to determine whether the association between each of the environmental factors and pancreatic cancer risk might differ by the genotype of the SNPs significantly associated with pancreatic cancer. To evaluate the significance of the gene-environment interaction, an interaction term (environmental factor  $\times$  SNP) was included in the logistic regression model. Log-likelihood ratio test was used to compare the logistic regression model with the interaction term to the model without the interaction term. A  $p < 0.05$  for the interaction term was considered statistically significant.

Receiver Operating Characteristic (ROC) curve analysis was conducted to evaluate the capability of the

environmental factors and SNPs in differentiating between pancreatic cancer cases and controls. An ROC curve was generated from each of the four risk models constructed using logistic regression: Model 1 included age, sex, education, and environmental factors significantly associated with pancreatic cancer; Model 2 included age, sex, education, and the unlinked SNPs significantly associated with pancreatic cancer. Model 3 included age, sex, education, environmental factors, and SNPs; and Model 4 included age, sex, education, environmental factors, SNPs, and interactions between environmental factors and SNPs. The  $c$ -statistic was used to calculate the area under the curve (AUC) to evaluate how well each of the models could differentiate between cases and controls. The AUCs of the different models were compared using the nonparametric approach proposed by Delong et al. [42].

## Results

The current study initially included 285 pancreatic cancer cases and 664 controls, with a participation rate of 86% for the cases and 81% for the controls. Thirteen subjects (7 cases and 6 controls) were excluded from the analysis due to a low sample call rate ( $< 90\%$ ) for the 25 SNPs. The final analysis included 278 pancreatic cancer cases and 658 controls. The majority (86%) of the cases were interviewed within 6 months of the pancreatic cancer diagnosis (76% were interviewed within 3 months and 10% were interviewed between 3 and 6 months). Due to the ongoing study subject recruitment, the age and the sex distributions were imbalanced despite the implementation of frequency matching. Cases were older in mean age and had a lower percentage of women compared to the controls (Table 1). Compared to the cases, a higher percentage of the controls completed at least a college education.

In the logistic regression adjusted only for age, sex, and education (analysis 1), cigarette smoking, poor oral hygiene (i.e. higher oral hygiene score), diabetes/glucose intolerance, and BMI  $\geq 27$  were all associated with an increased pancreatic cancer risk (Table 2). History of allergy and more frequent consumption of vegetables ( $> 3$  portions per week) and fruits ( $> 1$  portion per day) were associated with a reduced pancreatic cancer risk. In analysis 2, a multivariate logistic regression model containing all of the above-mentioned environmental factors showed that all except for consumption of fruits remained significantly associated with pancreatic cancer risk.

Among the 25 pancreatic cancer GWAS SNPs, 7 (rs2816938 (~ 11 kb upstream of *NR5A2*), rs10094872 (~ 28 kb upstream of *MYC*), rs9581943 (200 bp upstream of *PDX1*), and 4 chromosome 13q22.1 SNPs: rs4885093, rs9573163, rs9543325, rs9573166) showed statistically significant association with pancreatic cancer risk (Table 3 and the results for the entire 25

**Table 1** Demographic characteristics of the pancreatic cancer patients and control subjects

Characteristics	Cases N = 278 n (%)	Controls N = 658 n (%)	P-value <sup>a</sup>
<b>Age (years)</b>			
Mean (SE)	62.1 (0.4)	58.4 (0.6)	< 0.0001
<b>Sex</b>			
Men	164 (59.0)	282 (42.9)	< 0.0001
Women	114 (41.0)	376 (57.1)	
<b>Education</b>			
≤ Junior high	134 (48.2)	191 (29.0)	< 0.0001
High school/technical school	84 (30.2)	178 (27.1)	
College	50 (18.0)	229 (34.8)	
Graduate school	10 (3.6)	60 (9.1)	

Abbreviations: N number, SE standard error

<sup>a</sup>P-values were generated using T-tests (for continuous variables) or chi-squared tests (for categorical variables)

SNPs can be found in Supplementary Table 2). For these 7 SNPs, having at least one copy of the minor allele (dominant inheritance model) was associated with an increased pancreatic cancer risk ranging from 1.4 to 1.7 times. The 4 chromosome 13q22.1 SNPs, rs4885093, rs9573163, rs9543325, and rs9573166, are in high LD ( $r^2 > 0.93$ ); therefore, out of these 4 linked SNPs, we chose the SNP with the lowest p (rs4885093) for further analysis.

We conducted a multivariate analysis that included the four unlinked SNPs (*NR5A2* rs2816938, *MYC* rs10094872, *PDX1* rs9581943 and a chromosome 13q22 SNP, rs4885093). The results of the multivariate analysis showed that each of the four SNPs was independently associated with the risk of pancreatic cancer (Table 4).

We conducted gene-environment analyses with four SNPs (*NR5A2* rs2816938, *MYC* rs10094872, *PDX1* rs9581943 and a chromosome 13q22 SNP, rs4885093) significantly associated with pancreatic cancer risk. Two significant (p-interaction < 0.05) gene-environment interactions were observed (Table 5). Poor oral hygiene was associated with a more significantly elevated risk of pancreatic cancer among carriers of the *NR5A2* rs2816938 AT or AA genotype (OR for 1 point increment of the oral hygiene score = 2.83, 95% CI: 1.65–4.84) compared to those with the TT genotype (OR for 1 point increment of the oral hygiene score = 1.57, 95% CI: 1.30–1.90). BMI  $\geq$  27 was associated with an increased pancreatic cancer risk only among individuals with the *PDX1* rs9581943 GA or AA genotype (OR = 2.99, 95% CI: 1.89–4.72) but not among those with the GG genotype (OR = 1.13, 95% CI: 0.63–2.00).

Fig. 1 presents the results of the ROC curve analysis. ROC curve 1, constructed with age, sex, education, and

the four SNPs (*NR5A2* rs2816938, *MYC* rs10094872, *PDX1* rs9581943 and a chromosome 13q22 SNPs, rs4885093), had an AUC = 0.7115. ROC curve 2, constructed with age, sex, education, and the seven environmental factors (cigarette smoking, oral hygiene score, vegetable consumption, fruit consumption, allergy, diabetes/glucose intolerance, and BMI) had a better AUC = 0.7818 than the AUC of the ROC curve 1 ( $p < 0.0001$ ). ROC curve 3, constructed with the 7 environmental factors and the four SNPs plus age, sex, and education, marginally improved the AUC to 0.7924 compared to the AUC of ROC curve 2 ( $p = 0.05$ ). ROC curve 4, which included age, sex, education, the 7 environmental factors, the four SNPs and the two significant gene-environment interactions (poor oral hygiene x *NR5A2* rs2816938 and BMI x *PDX1* rs9581943) had the best AUC = 0.7982. ( $p = 0.006$  when compared to the AUC of ROC curve 2).

## Discussion

Consistent with results of the previous literature, our study showed that cigarette smoking, poor dental hygiene, diabetes, and obesity were associated with an increased pancreatic cancer risk, while having allergy and more frequent consumption of vegetables and fruits were associated with a lower pancreatic cancer risk. Among the 25 pancreatic cancer GWAS-identified SNPs, 7 (rs2816938 (~ 11 kb upstream of *NR5A2*), rs10094872 (~ 28 kb upstream of *MYC*), rs9581943 (200 bp upstream of *PDX1*) and 4 chromosome 13q22.1 SNPs: rs4885093, rs9573163, rs9543325, rs9573166) showed a statistically significant association with pancreatic cancer risk. In addition, our results showed two significant gene-environment interactions (between poor oral hygiene and *NR5A2* rs2816938 and between obesity and *PDX1* rs9581943) on the risk of pancreatic cancer. The best model in differentiating pancreatic cancer cases from controls was the one that included age, sex, education, environmental factors (cigarette smoking, oral hygiene, consumption of vegetables and fruits, allergy, diabetes/glucose intolerance, and BMI), four unlinked SNPs, and the two gene-environment interactions (between poor oral hygiene and *NR5A2* rs2816938 and between obesity and *PDX1* rs9581943).

Of the 25 GWAS-identified SNPs, we were able to validate the association between 7 SNPs on four chromosomal loci (1q32.1 near *NR5A2*, 8q24.21 near *MYC*, 13q12.2 near *PDX1*, and 13q22.1) and pancreatic risk. One of the 7 SNPs, rs2816938, was previously discovered as an independent pancreatic cancer risk SNP located on a known pancreatic cancer locus, 1q32.1, by a study of subjects with European ancestry [22]. This SNP is located ~ 11 kb upstream of *NR5A2*, which encodes a transcription factor involved in the organogenesis of

**Table 2** The association between lifestyle and clinical factors and pancreatic cancer risk

Characteristics	Cases N = 278 n (%)	Controls N = 658 n (%)	Analysis 1 OR (95% CI) <sup>a</sup>	Analysis 2 OR (95% CI) <sup>b</sup>
<b>Cigarette smoking</b>				
Never	162 (58.3)	496 (75.4)	Reference	Reference
Ever	116 (41.7)	162 (24.6)	1.63 (1.10–2.40)	1.61 (1.05–2.46)
<b>Oral hygiene score<sup>c</sup></b>				
0 (Good)	27 (9.7)	180 (27.4)	Reference	Reference
1	65 (23.4)	219 (33.3)	1.63 (0.98–2.70)	1.44 (0.85–2.43)
2	118 (42.4)	198 (30.1)	2.88 (1.77–4.69)	2.15 (1.28–3.62)
3 (Poor)	60 (21.6)	52 (7.9)	4.60 (2.56–8.29)	2.58 (1.37–4.86)
Unknown	8 (2.9)	9 (1.4)	–	–
<b>Vegetable consumption</b>				
3 portions or less per week	106 (38.1)	100 (15.2)	Reference	Reference
> 3 portions per week	172 (61.9)	558 (84.8)	0.29 (0.21–0.41)	0.34 (0.24–0.49)
<b>Fruit consumption</b>				
1 portion or less per day	85 (30.6)	143 (21.7)	Reference	Reference
> 1 portion per day	193 (69.4)	513 (78.0)	0.71 (0.51–1.00)	0.95 (0.80–1.13)
Unknown	0 (0.0)	2 (0.3)	–	–
<b>Allergy</b>				
No	227 (81.6)	405 (61.6)	Reference	Reference
Yes	51 (18.4)	251 (38.1)	0.46 (0.32–0.65)	0.47 (0.32–0.69)
Unknown	0 (0.0)	2 (0.3)	–	–
<b>Diabetes mellitus/glucose intolerance (DM/GI)</b>				
No DM/GI	196 (70.5)	602 (91.5)	Reference	Reference
< 2 years	22 (7.9)	16 (2.4)	4.08 (2.01–8.25)	3.04 (1.42–6.50)
> 2 years	59 (21.2)	39 (5.9)	3.71 (2.36–5.87)	2.84 (1.74–4.61)
Unknown	1 (0.4)	1 (0.2)	–	–
<b>BMI 2 years ago<sup>d</sup></b>				
< 18.5	4 (1.4)	36 (5.5)	0.37 (0.12–1.09)	0.38 (0.12–1.17)
18.5–23.9	102 (36.7)	319 (48.5)	Reference	Reference
24–26.9	83 (29.9)	191 (29.0)	1.07 (0.75–1.53)	0.99 (0.67–1.47)
27 or more	87 (31.3)	111 (16.9)	2.09 (1.43–3.06)	1.68 (1.11–2.54)
Unknown	2 (0.7)	1 (0.1)	–	–

Abbreviations: BMI body mass index, CI confidence interval, DM/GI diabetes mellitus/glucose intolerance, OR odds ratio

<sup>a</sup>OR and 95% CI of each independent variable in Model 1 was calculated using unconditional logistic regression, adjusted for sex, age, and education

<sup>b</sup>Model 2 included all of the independent variables in the same model with additional adjustment for sex, age, and education

<sup>c</sup>Oral hygiene score = tooth brushing + use of dental floss + regular dental visit, with tooth brushing:  $\geq 2$  times per day = 0, < 2 times per day = 1; use of dental floss: yes = 0, no = 1; and regular dental visit: yes = 0, no = 1

<sup>d</sup>BMI at two years before the pancreatic cancer diagnosis for the cases or before the interview date for the controls

pancreas, including acinar differentiation [43]. Flandez et al. showed that mice with *NR5A2* heterozygosity developed a more severe acute pancreatitis with increased acino-ductal metaplasia and defective recovery from pancreatitis-induced damage [44]. In addition, *NR5A2* heterozygosity cooperated with *Kras* mutation in the oncogenesis of pancreatic cancer [44]. These suggested that *NR5A2* could be a pancreatic tumor suppressor [44]. A second SNP (rs9581943) validated by our study

was previously reported by a pancreatic cancer GWAS study with study subjects of European ancestry [28]. It is located 200 bp upstream of *PDX1*, which has been shown to regulate the initiation and the maintenance of pancreatic cancer, playing both the tumor-suppressive and oncogenic roles throughout the different stages of the pancreatic cancer development and progression [45]. Another pancreatic cancer GWAS-identified SNP (rs10094872) validated by our study was previously

**Table 3** Results for the significant associations between 7 genome-wide association study-identified single nucleotide polymorphisms and the risk of pancreatic cancer

SNPs (chromosome locus, nearest gene (s))	Cases n (%)	Controls n (%)	OR (95% CI) <sup>a</sup>	P
<b>rs2816938</b> (1q32.1, <i>NR5A2</i> )				
TT	232 (83.5)	583 (88.7)	Reference	
AT	44 (15.8)	72 (11.0)	1.68 (1.09–2.58)	0.02
AA	2 (0.7)	2 (0.3)	1.58 (0.22–11.39)	0.65
AT+AA	46 (16.6)	74 (11.3)	1.68 (1.10–2.55)	0.02
Every 1 copy of A			1.61 (1.08–2.40)	0.02
<b>rs10094872</b> (8q24.21, <i>MYC</i> )				
AA	159 (57.2)	429 (65.3)	Reference	
TA	111 (39.9)	196 (29.8)	1.61 (1.19–2.22)	0.002
TT	8 (2.9)	32 (4.9)	0.62 (0.27–1.44)	0.27
TA + TT	119 (42.8)	228 (34.7)	1.48 (1.09–2.00)	0.01
Every 1 copy of T			1.23 (0.95–1.59)	0.11
<b>rs9581943</b> (13q12.2, <i>PDX1</i> )				
GG	111 (40.1)	323 (49.1)	Reference	
GA	126 (45.5)	250 (38.0)	1.41 (1.03–1.94)	0.03
AA	40 (14.4)	85 (12.9)	1.38 (0.87–2.17)	0.17
GA + AA	166 (59.9)	335 (50.9)	1.40 (1.04–1.89)	0.03
Every 1 copy of A			1.23 (1.00–1.52)	0.05
<b>rs4885093</b> (13q22.1, intergenic region)				
AA	59 (21.2)	196 (29.8)	Reference	
AG	150 (54.0)	335 (50.9)	1.43 (0.99–2.06)	0.06
GG	69 (24.8)	127 (19.3)	1.65 (1.07–2.54)	0.02
AG + GG	219 (78.8)	462 (70.2)	1.49 (1.05–2.11)	0.02
Every 1 copy of G			1.29 (1.04–1.59)	0.02
<b>rs9573163</b> (13q22.1, intergenic region)				
GG	59 (21.2)	195 (29.6)	Reference	
GC	150 (54.0)	337 (51.2)	1.40 (0.97–2.02)	0.07
CC	69 (24.8)	126 (19.2)	1.64 (1.06–2.53)	0.03
GC + CC	219 (78.8)	463 (70.4)	1.47 (1.04–2.08)	0.03
Every 1 copy of C			1.28 (1.03–1.59)	0.02
<b>rs9543325</b> (13q22.1, intergenic region)				
TT	62 (22.3)	198 (30.1)	Reference	
TC	146 (52.5)	333 (50.7)	1.40 (0.99–1.98)	0.06
CC	70 (25.2)	126 (19.2)	1.77 (1.18–2.67)	0.006
TC + CC	216 (77.7)	459 (69.9)	1.44 (1.02–2.02)	0.04
Every 1 copy of C			1.28 (1.03–1.58)	0.02
<b>rs9573166</b> (13q22.1, intergenic region)				
AA	62 (22.4)	193 (29.4)	Reference	
GA	145 (52.3)	335 (51.0)	1.29 (0.90–1.86)	0.17
GG	70 (25.3)	129 (19.6)	1.56 (1.02–2.39)	0.04
GA + GG	215 (77.6)	464 (70.6)	1.37 (0.97–1.93)	0.08
Every 1 copy of G			1.25 (1.01–1.55)	0.04

**Abbreviations:** CI confidence interval, OR odds ratio

<sup>a</sup>OR and 95% CI were calculated using unconditional logistic regression, adjusted for age, sex, and education

**Table 4** Multivariate analysis to assess the independent association between the four unlinked single nucleotide polymorphisms and pancreatic cancer risk

SNPs (chromosome locus, nearest gene (s))	Cases n (%)	Controls n (%)	OR (95% CI) <sup>a</sup>	P
<b>rs2816938</b> (1q32.1, <i>NR5A2</i> )				
TT	232 (83.5)	583 (88.7)	Reference	
AT+AA	46 (16.6)	74 (11.3)	1.71 (1.12–2.63)	0.01
<b>rs10094872</b> (8q24.21, <i>MYC</i> )				
AA	159 (57.2)	429 (65.3)	Reference	
TA + TT	119 (42.8)	228 (34.7)	1.50 (1.10–2.04)	0.01
<b>rs9581943</b> (13q12.2, <i>PDX1</i> )				
GG	111 (40.1)	323 (49.1)	Reference	
GA + AA	166 (59.9)	335 (50.9)	1.39 (1.03–1.88)	0.03
<b>rs4885093</b> (13q22.1, intergenic region)				
AA	59 (21.2)	196 (29.8)	Reference	
AG + GG	219 (78.8)	462 (70.2)	1.48 (1.04–2.10)	0.03

**Abbreviations** CI confidence interval, OR odds ratio

<sup>a</sup>OR and 95% CI were calculated using multivariate unconditional logistic regression model that included, age, sex, education, and the four unlinked SNPs

discovered as a pancreatic cancer risk SNP according to a study with subjects of European ancestry [22]. This SNP is located ~ 28 kb upstream of *MYC*. Previous studies have demonstrated the major contribution of *MYC* in the oncogenesis of pancreatic cancer and showed that *MYC* could be activated by most of the genetic and epigenetic events involved in the initiation and the progression of pancreatic cancer [46]. Four (rs4885093, rs9573163, rs9543325, and rs9573166) of the 7 SNPs validated by our study are in high LD and located on 13q22.1, which is a non-genic region previously identified as a pancreatic cancer risk locus by three studies with subjects of mostly European ancestry and a Chinese study. Besides pancreatic cancer, polymorphisms on 13q22.1 have also been associated with the risk of esophageal squamous cell carcinoma, gastric cardia cancer, and endometrial cancer [47, 48]. Functional study of this region discovered an rs386772267 (an insertion/deletion polymorphism)-containing sub-region that may regulate the expression of *DIS3* through long-range (570 kb) physical interaction [33]. More investigations are needed to decipher how this may affect the development of pancreatic cancer.

Our analysis showed a significant interaction between poor oral hygiene and *NR5A2* rs2816938 on pancreatic cancer risk. Cobo et al. showed that *NR5A2* is involved in the suppression of inflammation in pancreas and a pre-inflammatory state could be observed in the pancreas of human subjects with low *NR5A2* mRNA levels [49]. Poor oral hygiene may lead to the overgrowth of oral pathogenic bacteria, which have been associated with an increased pancreatic cancer risk. Michaud et al. reported that higher levels of serum antibody for *P*

*gingivalis* ATTC53978, a periodontopathogenic bacterium, were associated with an elevated risk of pancreatic cancer [50]. By sequencing the 16S rRNA gene of the bacteria in the oral wash samples of 361 pancreatic cancer cases and 371 controls, Fan et al. found an increased risk of pancreatic cancer among individuals with carriage of oral pathogens, including *P. gingivalis* and *A. actinomycetemcomitans* [51]. Several possible biological mechanisms have been proposed to explain the association between infection with oral pathogens and pancreatic cancer. Infection with oral pathogens may generate systemic inflammation, including inflammation at a distant site, such as the pancreas, and chronic inflammation may promote carcinogenesis [52, 53]. Alternatively, oral pathogenic bacteria may travel through circulation to induce local inflammation in the pancreas to promote the oncogenesis of pancreatic cancer. The interaction between poor oral hygiene and *NR5A2* rs2816938 pointed toward the role of inflammation pathways in the oncogenesis of pancreatic cancer, although the biological mechanism underlying this interaction requires further investigation.

Other studies, although involving factors other than oral hygiene and the *NR5A2* gene, have also indicated the role of inflammation in the occurrence of pancreatic cancer. Duell et al. showed that pro-inflammatory conditions, such as pancreatitis and smoking, and pro-inflammatory genes, may have a combined effect to affect the development of pancreatic cancer [54]. Antwi et al. reported that pro-inflammatory diet, which is diet associated with higher levels of circulating inflammatory biomarkers (IL-1 $\beta$ , IL-4, IL-6, IL-10, TNF- $\alpha$ , and C-reactive protein), was associated with an increased pancreatic cancer risk [55].



**Table 5** The interaction between 4 genome-wide association study-identified single nucleotide polymorphisms and lifestyle and clinical factors on the risk of pancreatic cancer risk

	rs2816938 (1q32.1, NR5A2)		rs10094872 (8q24.21, MYC)		rs9581943 (13q12.2, PDX1)		rs4885093 (13q22.1, intergenic region)	
	TT	AT+AA	AA	TA + TT	GG	GA + AA	AA	AG + GG
<b>Cigarette smoking</b>	OR (95% CI) <sup>a</sup>	OR (95% CI) <sup>a</sup>	OR (95% CI) <sup>a</sup>	OR (95% CI) <sup>a</sup>	OR (95% CI) <sup>a</sup>	OR (95% CI) <sup>a</sup>	OR (95% CI) <sup>a</sup>	OR (95% CI) <sup>a</sup>
Never	Referent	Referent	Referent	Referent	Referent	Referent	Referent	Referent
Ever	1.32 (0.87–1.99)	7.41 (2.10–26.09)	1.51 (0.91–2.51)	1.92 (1.04–3.54)	2.04 (1.11–3.74)	1.39 (0.83–2.34)	0.85 (0.35–2.08)	1.91 (1.22–2.98)
	P-interaction = 0.07		P-interaction = 0.65		P-interaction = 0.73		P-interaction = 0.35	
<b>Oral hygiene score<sup>b</sup></b>	Referent	Referent	Referent	Referent	Referent	Referent	Referent	Referent
0 (Good)	Referent	Referent	Referent	Referent	Referent	Referent	Referent	Referent
1	1.51 (0.88–2.60)	2.97 (0.61–14.45)	1.65 (0.84–3.26)	1.63 (0.75–3.56)	2.07 (0.99–4.34)	1.30 (0.65–2.63)	2.06 (0.60–7.12)	1.50 (0.85–2.63)
2	2.35 (1.39–3.98)	13.26 (2.79–63.06)	2.83 (1.47–5.46)	2.93 (1.37–6.24)	2.72 (1.31–5.65)	2.71 (1.39–5.31)	3.75 (1.14–12.34)	2.55 (1.48–4.42)
3 (Poor)	3.78 (2.01–7.10)	17.64 (2.37–131.31)	4.17 (1.91–9.09)	4.74 (1.87–11.98)	4.12 (1.67–10.21)	4.20 (1.89–9.37)	4.35 (1.11–17.09)	4.37 (2.23–8.57)
	P-interaction = 0.21		P-interaction = 0.73		P-interaction = 0.66		P-interaction = 0.68	
Every 1 point increment	1.57 (1.30–1.90)	2.83 (1.65–4.84)	1.64 (1.31–2.07)	1.75 (1.33–2.31)	1.58 (1.22–2.06)	1.72 (1.35–2.19)	1.63 (1.11–2.40)	1.68 (1.38–2.05)
	P-interaction = 0.047		P-interaction = 0.71		P-interaction = 0.62		P-interaction = 0.38	
<b>Allergy</b>	Referent	Referent	Referent	Referent	Referent	Referent	Referent	Referent
No	Referent	Referent	Referent	Referent	Referent	Referent	Referent	Referent
Yes	0.44 (0.29–0.65)	0.51 (0.21–1.23)	0.36 (0.22–0.59)	0.61 (0.36–1.05)	0.42 (0.25–0.73)	0.49 (0.30–0.79)	0.26 (0.12–0.61)	0.54 (0.36–0.81)
	P-interaction = 0.83		P-interaction = 0.16		P-interaction = 0.77		P-interaction = 0.07	
<b>Diabetes mellitus/glucose intolerance (DM/GI)</b>	Referent	Referent	Referent	Referent	Referent	Referent	Referent	Referent
No DM/GI	Referent	Referent	Referent	Referent	Referent	Referent	Referent	Referent
< 2 years	5.61 (2.43–12.98)	1.29 (0.30–5.52)	2.67 (1.08–6.61)	8.94 (2.34–34.23)	6.20 (2.03–18.93)	2.83 (1.12–7.11)	9.11 (2.04–40.72)	3.18 (1.42–7.12)
> 2 years	3.37 (2.09–5.44)	18.50 (2.09–163.62)	3.84 (2.13–6.92)	3.48 (1.68–7.19)	4.05 (2.02–8.10)	3.27 (1.76–6.07)	4.54 (1.67–12.33)	3.70 (2.19–6.26)
	P-interaction = 0.06		P-interaction = 0.32		P-interaction = 0.46		P-interaction = 0.47	
<b>Vegetable consumption</b>	Referent	Referent	Referent	Referent	Referent	Referent	Referent	Referent
3 portions or less per week	Referent	Referent	Referent	Referent	Referent	Referent	Referent	Referent
> 3 portions per week	0.28 (0.19–0.40)	0.41 (0.16–1.06)	0.27 (0.17–0.42)	0.34 (0.20–0.59)	0.28 (0.17–0.48)	0.31 (0.20–0.50)	0.22 (0.10–0.48)	0.31 (0.21–0.46)
	P-interaction = 0.43		P-interaction = 0.49		P-interaction = 0.71		P-interaction = 0.55	
<b>Fruit consumption</b>	Referent	Referent	Referent	Referent	Referent	Referent	Referent	Referent
1 portion or less per day	Referent	Referent	Referent	Referent	Referent	Referent	Referent	Referent
> 1 portion per day	0.74 (0.51–1.07)	0.44 (0.16–1.21)	0.77 (0.50–1.20)	0.63 (0.36–1.08)	0.80 (0.47–1.37)	0.71 (0.45–1.11)	0.76 (0.38–1.51)	0.68 (0.46–1.01)
	P-interaction = 0.55		P-interaction = 0.56		P-interaction = 0.88		P-interaction = 0.76	

**Table 5** The interaction between 4 genome-wide association study-identified single nucleotide polymorphisms and lifestyle and clinical factors on the risk of pancreatic cancer risk (Continued)

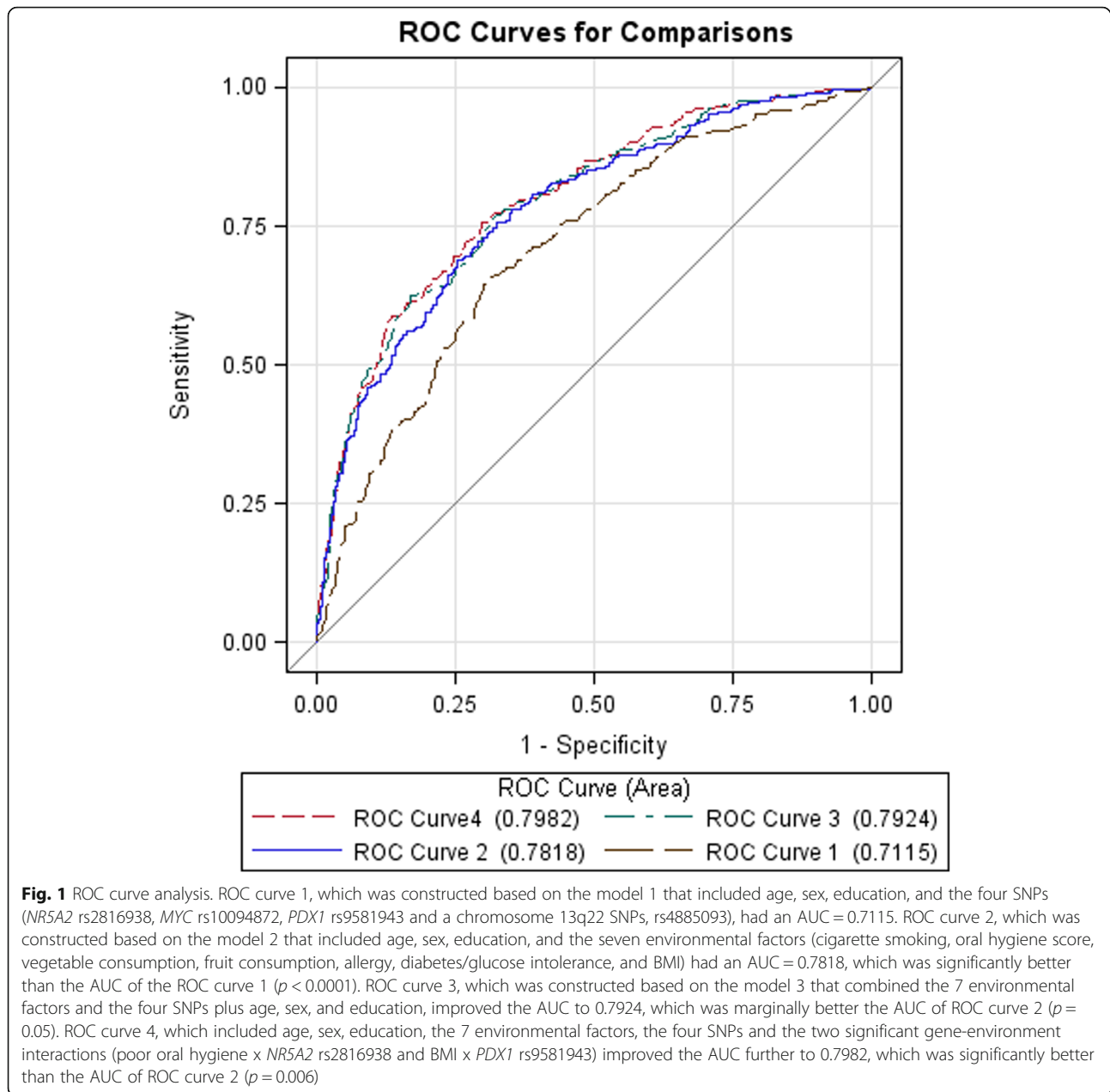
	rs2816938 (1q32.1, NR5A2)		rs10094872 (8q24.21, MYC)		rs9581943 (13q12.2, PDX1)		rs4885093 (13q22.1, intergenic region)	
	TT	AT+AA	AA	TA + TT	GG	GA + AA	AA	AG + GG
<b>BMI 2 years ago<sup>c</sup></b>								
< 27	Referent	Referent	Referent	Referent	Referent	Referent	Referent	Referent
27 or more	2.21 (1.53–3.20)	1.45 (0.54–3.90)	2.00 (1.29–3.09)	2.37 (1.34–4.19)	1.13 (0.63–2.00)	2.99 (1.89–4.72)	1.48 (0.64–3.42)	2.37 (1.61–3.49)
	P-interaction = 0.46		P-interaction = 0.67		P-interaction = 0.009		P-interaction = 0.18	

Abbreviations: BMI body mass index, CI confidence interval, DM/GI diabetes mellitus/glucose intolerance, OR odds ratio

<sup>a</sup>OR and 95% CI were calculated using unconditional logistic regression, adjusted for sex, age, and education

<sup>b</sup>Oral hygiene score = tooth brushing + use of dental floss + regular dental visit, with tooth brushing:  $\geq 2$  times per day = 0, < 2 times per day = 1; use of dental floss: yes = 0, no = 1; and regular dental visit: yes = 0, no = 1

<sup>c</sup>BMI at two years before the pancreatic cancer diagnosis for the cases or before the interview date for the controls



Another significant gene-environment interaction in our study was between obesity and *PDX1* rs9581943. *PDX1* plays an important role in maintaining the function and the survival of pancreatic beta-cells, which produce and secrete insulin [56]. A reduced expression of *PDX1* may result in the pathogenesis of diabetes [56], which is a known risk factor of pancreatic cancer. Obesity is a risk factor for both pancreatic cancer and diabetes [5, 6, 57]. The interaction between obesity and *PDX1* rs9581943 in our analysis suggested the involvement of metabolic disease-related pathways in the development of pancreatic cancer.

Although to our knowledge, our study is the first to observe an interaction between obesity and *PDX1* rs9581943, polymorphisms of other genes have been reported to interact with obesity to influence pancreatic cancer risk. Tang et al. studied the polymorphisms of obesity and diabetes-associated genes and found that the associations between pancreatic cancer and the variants of *FTO* and *ADIPOQ*, which encodes the protein adiponectin, differed according to the overweight status [58]. Nakao et al. reported that the associations between the polymorphisms of *IGF1* and pancreatic cancer were significant only among the overweight individuals [59]. Studies have indicated an interplay between IGF-1 and

adiponectin to influence the development of obesity, insulin resistance, diabetes, and cancer [60]. These studies again suggested the roles of metabolic disease-related pathways in the occurrence of pancreatic cancer.

This study has several limitations. First, for a hospital-based case-control study, it is difficult to determine whether the cases and the controls are from the same source population, and often the controls may not be representative of the general population. The potential selection bias of the hospital controls may have less impact for genetic association studies, unless the genetic polymorphisms were somehow associated with the hospital visits of the control subjects. Hospital controls could be unrepresentative of the general population with regards to environmental factors since they might be less healthy compared to the general population. In our control subjects, the percentage of current smokers was 21.3% for men and 3.1% for women, which were similar to the national prevalence of current smokers (26% for men and 2.3% for women) according to the Annual Report of the Health Promotion Administration of Taiwan [61]. Similarly, in our control subjects, the percentage of BMI  $\geq 24$  (overweight + obese) 2 years before the interview date was 45.9%, which was similar to the 47.1% at the national level [61]. The comparison between our control subjects and the national data indicated that our control subjects could be representative of the general population. Second, due to the case-control study design, subjects were asked to recall exposure information in the past, such that the random recall errors might have biased the results towards the null. In addition, the case subjects might be more likely to ruminate about the possible causes of developing pancreatic cancer than the controls and this recall bias could have biased the results away from the null. Third, the sample size of the current study might not have a sufficient statistical power to validate some of the GWAS-identified SNPs, especially if the minor allele frequency is low. The minor allele frequencies of the 25 SNPs examined by the current study ranged from 0.05 to 0.50. If the minor allele frequency = 0.05 and assuming a two-sided  $\alpha = 0.05$ , the current study would have a sufficient statistical power to detect a dominant OR = 1.8 and a log-additive OR = 1.75. If the minor allele frequency = 0.50 and assuming a two-sided  $\alpha = 0.05$ , the current study would have a sufficient statistical power to detect a dominant OR = 1.65 and a log-additive OR = 1.33. Fourth, we did not consider adjusting for multiple testing, because all of the pancreatic cancer GWAS-identified SNPs had been well validated in the previous studies, and due to the relatively small sample size of the current study, we decided to use  $p < 0.05$  as a cut-off for statistical significance. Fifth, because of the differences in genetic background and lifestyle factors, the results of the current study may not be

generalizable to other ethnic/racial groups. Finally, there might be undiscovered environmental or genetic factors that could have confounded the associations observed by the current study. Further studies are needed to determine these factors.

This study has several strengths. This is the first study to validate the pancreatic cancer GWAS-identified SNPs using a case-control pancreatic cancer study conducted in Taiwan. The results indicated that some of the GWAS-identified pancreatic cancer SNPs may also contribute to the pancreatic cancer occurring in Taiwan but the non-replications of the majority of the 25 SNPs also indicated that there may be pancreatic cancer-associated genetic polymorphisms that are specific to the Taiwanese population and further investigations are required. Another strength of this study is the investigation of gene-environment interaction. We were able to show the significant interaction between poor oral hygiene and *NR5A2* rs2816938 and between obesity and *PDX1* rs9581943, which suggested the involvement of inflammation and metabolic disease-related pathways in the development of pancreatic cancer.

## Conclusions

Our study confirmed the association between 7 GWAS-identified SNPs and pancreatic risk among the Taiwanese population. Our results also showed that pancreatic cancer was jointly influenced by lifestyle and medical factors, genetic polymorphisms, and gene-environment interaction. Some of the GWAS-identified SNPs were not validated by our study, which might be explained partly by racial/ethnic differences in genetic background. In the future, we plan to conduct a GWAS of pancreatic cancer in Taiwan to determine additional genetic polymorphisms that are more relevant to the pancreatic cancer cases occurring in Taiwan.

## Supplementary information

**Supplementary information** accompanies this paper at <https://doi.org/10.1186/s12929-020-00664-9>.

**Additional file 1: Table S1.** Descriptions of the 25 genome-wide association study-identified single nucleotide polymorphisms of pancreatic cancer

**Additional file 2: Table S2.** The association between 25 genome-wide association study-identified single nucleotide polymorphisms and the risk of pancreatic cancer

## Abbreviations

AUC: Area under the curve; BMI: Body mass index; CI: Confidence interval; GWAS: Genome-wide association studies; IARC: International Agency for Research on Cancer; LD: Linkage disequilibrium; OR: Odds ratio; ROC: Receiver operating characteristic; SNP: Single nucleotide polymorphism

## Acknowledgments

Not applicable.

**Authors' contributions**

Conceptualization, YSS, LTC, JSC; Study design: YSS, LTC, JSC; Data acquisition: all authors; Project Administration: all authors; Data analysis: JSC; Data interpretation: YSS, LTC, JSC; Writing – Original Draft Preparation, JSC; Writing – Review & Editing: all authors; Funding Acquisition: YSS, LTC, JSC; All authors approved the final manuscript.

**Funding**

This work was supported by a grant from the Ministry of Science and Technology, Taiwan (105–2628-B-400 -006 -MY3).

**Availability of data and materials**

All data generated during this study are included in this published article.

**Ethics approval and consent to participate**

The current study was approved by institutional review boards of the National Cheng Kung University Hospital (ethic approval number: B-BR-102-070) and the National Health Research Institutes (ethic approval number: EC1030109-E). A signed informed consent was collected from every study participant who agreed to participate in the study.

**Consent for publication**

Not Applicable.

**Competing interests**

The authors declare that they have no competing interests.

**Author details**

<sup>1</sup>Department of Surgery, National Cheng Kung University Hospital, National Cheng Kung University, 138 Sheng Li Road, Tainan 70456, Taiwan. <sup>2</sup>Institute of Clinical Medicine, College of Medicine, National Cheng Kung University, Tainan, 138 Sheng Li Road, Tainan 70456, Taiwan. <sup>3</sup>National Institute of Cancer Research, National Health Research Institutes, 1F No 367, Sheng-Li Road, Tainan 70456, Taiwan. <sup>4</sup>Department of Internal Medicine, National Cheng Kung University Hospital, National Cheng Kung University, 138 Sheng Li Road, Tainan 70456, Taiwan. <sup>5</sup>Department of Internal Medicine, Kaohsiung Medical University Hospital, Kaohsiung Medical University, Ziyou 1st Road, Sanmin District, Kaohsiung 80756, Taiwan. <sup>6</sup>Institute of Molecular Medicine, College of Medicine, National Cheng Kung University, 138 Sheng Li Road, Tainan 70456, Taiwan. <sup>7</sup>Department of Family Medicine, National Cheng Kung University Hospital, National Cheng Kung University, 138 Sheng Li Road, Tainan 70456, Taiwan. <sup>8</sup>Preventive Medicine Center, Taichung Tzu Chi Hospital, 88 Section 1, Fengxing Road, Tanzi District, Taichung 427, Taiwan. <sup>9</sup>Department of Nursing, National Cheng Kung University Hospital, National Cheng Kung University, 138 Sheng Li Road, Tainan 70456, Taiwan.

Received: 6 August 2019 Accepted: 22 May 2020

Published online: 26 May 2020

**References**

- Cancer Today. Data visualization tool for exploring the global cancer burden in 2018. Lyon: International Agency for Research on Cancer; 2018. Available from: <http://gco.iarc.fr/today/home>. Accessed November 7 2018.
- Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. *CA Cancer J Clin*. 2019; 69:7–34.
- Bosetti C, Lucenteforte E, Silverman DT, Petersen G, Bracci PM, Ji BT, et al. Cigarette smoking and pancreatic cancer: an analysis from the international pancreatic Cancer case-control consortium (Panc4). *Ann Oncol*. 2012;23: 1880–8.
- Bosetti C, Rosato V, Li D, Silverman D, Petersen GM, Bracci PM, et al. Diabetes, antidiabetic medications, and pancreatic cancer risk: an analysis from the international pancreatic Cancer case-control consortium. *Ann Oncol*. 2015;25:2065–72.
- Aune D, Greenwood DC, Chan DS, Vieira R, Vieira AR, Navarro Rosenblatt DA, et al. Body mass index, abdominal fatness and pancreatic cancer risk: a systematic review and non-linear dose-response meta-analysis of prospective studies. *Ann Oncol*. 2012;23:843–52.
- Arslan AA, Helzlsouer KJ, Kooperberg C, Shu XO, Stepkowski E, Bueno-de-Mesquita HB, et al. Anthropometric measures, body mass index, and pancreatic cancer: a pooled analysis from the pancreatic Cancer cohort consortium (PanScan). *Arch Intern Med*. 2010;170:791–802.
- Raimondi S, Lowenfels AB, Morselli-Labate AM, Maisonneuve P, Pezzilli R. Pancreatic cancer in chronic pancreatitis: aetiology, incidence, and early detection. *Best Pract Res Clin Gastroenterol*. 2010;24:349–58.
- Gandini S, Lowenfels AB, Jaffee EM, Armstrong TD, Maisonneuve P. Allergies and the risk of pancreatic cancer: a meta-analysis with review of epidemiology and biological mechanisms. *Cancer Epidemiol Biomark Prevent*. 2005;14:1908–16.
- Chan JM, Gong Z, Holly EA, Bracci PM. Dietary patterns and risk of pancreatic cancer in a large population-based case-control study in the San Francisco Bay Area. *Nutr Cancer*. 2013;65:157–64.
- Bosetti C, Bravi F, Turati F, Edefonti V, Polesel J, Decarli A, et al. Nutrient-based dietary patterns and pancreatic cancer risk. *Ann Epidemiol*. 2013;23:124–8.
- Jansen RJ, Robinson DP, Stolzenberg-Solomon RZ, Bamlet WR, de Andrade M, Oberg AL, et al. Nutrients from fruit and vegetable consumption reduce the risk of pancreatic cancer. *J Gastrointest Cancer*. 2013;44:152–61.
- O'Rourke MA, Cantwell MM, Cardwell CR, Mulholland HG, Murray LJ. Can physical activity modulate pancreatic cancer risk? A systematic review and meta-analysis. *Int J Cancer*. 2010;126:2957–68.
- Keum N, Bao Y, Smith-Warner SA, Orav J, Wu K, Fuchs CS, et al. Association of Physical Activity by type and intensity with digestive system Cancer risk. *JAMA Oncol*. 2016;2:1146–53.
- Wu L, Zheng W, Xiang YB, Gao YT, Li HL, Cai H, et al. Physical activity and pancreatic Cancer risk among urban Chinese: results from two prospective cohort studies. *Cancer Epidemiol Biomark Prevent*. 2018;27:479–87.
- Farris MS, Mosli MH, McFadden AA, Friedenreich CM, Brenner DR. The association between leisure time physical activity and pancreatic Cancer risk in adults: a systematic review and meta-analysis. *Cancer Epidemiol Biomark Prevent*. 2015;24:1462–73.
- Huang J, Roosaar A, Axell T, Ye W. A prospective cohort study on poor oral hygiene and pancreatic cancer risk. *Int J Cancer*. 2016;138:340–7.
- Lucenteforte E, La Vecchia C, Silverman D, Petersen GM, Bracci PM, Ji BT, et al. Alcohol consumption and pancreatic cancer: a pooled analysis in the international pancreatic Cancer case-control consortium (Panc4). *Ann Oncol*. 2012;23:374–82.
- Naudin S, Li K, Jaouen T, Assi N, Kyro C, Tjonneland A, et al. Lifetime and baseline alcohol intakes and risk of pancreatic cancer in the European prospective investigation into Cancer and nutrition study. *Int J Cancer* 2018; 143:801–12.
- Klein AP. Genetic susceptibility to pancreatic cancer. *Mol Carcinog*. 2012;51: 14–24.
- Kleeff J, Korc M, Apte M, La Vecchia C, Johnson CD, Biankin AV, et al. Pancreatic cancer. *Nat Rev Dis Primers*. 2016;2:16022.
- Dai P, Li J, Li W, Qin X, Wu X, Di W, et al. Genetic polymorphisms and pancreatic cancer risk: a PRISMA-compliant systematic review and meta-analysis. *Medicine*. 2019;98:e16541.
- Zhang M, Wang Z, Obazee O, Jia J, Childs EJ, Hoskins J, et al. Three new pancreatic cancer susceptibility signals identified on chromosomes 1q32.1, 5p15.33 and 8q24.21. *Oncotarget*. 2016;7:66328–43.
- Petersen GM, Amundadottir L, Fuchs CS, Kraft P, Stolzenberg-Solomon RZ, Jacobs KB, et al. A genome-wide association study identifies pancreatic cancer susceptibility loci on chromosomes 13q22.1, 1q32.1 and 5p15.33. *Nat Genet*. 2010;42:224–8.
- Ueno M, Ohkawa S, Morimoto M, Ishii H, Matsuyama M, Kuruma S, et al. Genome-wide association study-identified SNPs (rs3790844, rs3790843) in the NR5A2 gene and risk of pancreatic cancer in Japanese. *Sci Rep*. 2015;5: 17018.
- Childs EJ, Mocchi E, Campa D, Bracci PM, Gallinger S, Goggins M, et al. Common variation at 2p13.3, 3q29, 7p13 and 17q25.1 associated with susceptibility to pancreatic cancer. *Nat Genet*. 2015;47:911–6.
- Wu C, Miao X, Huang L, Che X, Jiang G, Yu D, et al. Genome-wide association study identifies five loci associated with susceptibility to pancreatic cancer in Chinese populations. *Nat Genet*. 2011;44:62–6.
- Fang J, Jia J, Makowski M, Xu M, Wang Z, Zhang T, et al. Functional characterization of a multi-cancer risk locus on chr5p15.33 reveals regulation of TERT by ZNF148. *Nat Commun*. 2017;8:15034.
- Wolpin BM, Rizzato C, Kraft P, Kooperberg C, Petersen GM, Wang Z, et al. Genome-wide association study identifies multiple susceptibility loci for pancreatic cancer. *Nat Genet*. 2014;46:994–1000.
- Liu C, Wang Y, Huang H, Wang C, Zhang H, Kong Y, et al. Association between CLPTM1L-TERT rs401681 polymorphism and pancreatic cancer risk among Chinese Han population. *Tumour Biol*. 2014;35:5453–7.

30. Low SK, 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:e11824.
31. Willis JA, Olson SH, Orlow I, Mukherjee S, McWilliams RR, Kurtz RC, et al. A replication study and genome-wide scan of single-nucleotide polymorphisms associated with pancreatic cancer risk and overall survival. *Clin Cancer Res*. 2012;18:3942–51.
32. 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;41:986–90.
33. Hoskins JW, Ibrahim A, Emmanuel MA, Manmiller SM, Wu Y, O'Neill M, et al. Functional characterization of a chr13q22.1 pancreatic cancer risk locus reveals long-range interaction and allele-specific effects on DIS3 expression. *Hum Mol Genet*. 2016;25:4726–38.
34. Duell EJ, Holly EA, Bracci PM, Liu M, Wiencke JK, Kelsey KT. A population-based, case-control study of polymorphisms in carcinogen-metabolizing genes, smoking, and pancreatic adenocarcinoma risk. *J Natl Cancer Inst*. 2002;94:297–306.
35. Li D, Ahmed M, Li Y, Jiao L, Chou TH, Wolff RA, et al. 5,10-Methylenetetrahydrofolate reductase polymorphisms and the risk of pancreatic cancer. *Cancer Epidemiol Biomark Prevent*. 2005;14:1470–6.
36. Jiao L, Bondy ML, Hassan MM, Wolff RA, Evans DB, Abbruzzese JL, et al. Selected polymorphisms of DNA repair genes and risk of pancreatic cancer. *Cancer Detect Prev*. 2006;30:284–91.
37. Dong X, Li Y, Chang P, Tang H, Hess KR, Abbruzzese JL, et al. Glucose metabolism gene variants modulate the risk of pancreatic cancer. *Cancer Prevent Res*. 2011;4:758–66.
38. Tsai HJ, Chang JS. Environmental risk factors of pancreatic Cancer. *J Clin Med*. 2019;8:1427.
39. Tsai ST, Wong TY, Ou CY, Fang SY, Chen KC, Hsiao JR, et al. The interplay between alcohol consumption, oral hygiene, ALDH2 and ADH1B in the risk of head and neck cancer. *Int J Cancer*. 2014;135:2424–36.
40. Hsiao JR, Chang CC, Lee WT, Huang CC, Ou CY, Tsai ST, et al. The interplay between oral microbiome, lifestyle factors and genetic polymorphisms in the risk of oral squamous cell carcinoma. *Carcinogenesis*. 2018;39:778–87.
41. Chang CC, Lee WT, Hsiao JR, Ou CY, Huang CC, Tsai ST, et al. Oral hygiene and the overall survival of head and neck cancer patients. *Cancer Med*. 2019;8:1854–64.
42. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*. 1988;44:837–45.
43. Hale MA, Swift GH, Hoang CQ, Deering TG, Masui T, Lee YK, et al. The nuclear hormone receptor family member NR5A2 controls aspects of multipotent progenitor cell formation and acinar differentiation during pancreatic organogenesis. *Development*. 2014;141:3123–33.
44. Flandez M, Cendrowski J, Canamero M, Salas A, del Pozo N, Schoonjans K, et al. Nr5a2 heterozygosity sensitises to, and cooperates with, inflammation in KRas(G12V)-driven pancreatic tumourigenesis. *Gut*. 2014;63:647–55.
45. Roy N, Takeuchi KK, Ruggeri JM, Bailey P, Chang D, Li J, et al. PDX1 dynamically regulates pancreatic ductal adenocarcinoma initiation and maintenance. *Genes Dev*. 2016;30:2669–83.
46. Hessmann E, Schneider G, Ellenrieder V, Siveke JT. MYC in pancreatic cancer: novel mechanistic insights and their translation into therapeutic strategies. *Oncogene*. 2016;35:1609–18.
47. Chang J, Wei L, Miao X, Yu D, Tan W, Zhang X, et al. Two novel variants on 13q22.1 are associated with risk of esophageal squamous cell carcinoma. *Cancer Epidemiol Biomark Prevent*. 2015;24:1774–80.
48. Cheng TH, Thompson DJ, O'Mara TA, Painter JN, Glubb DM, Flach S, et al. Five endometrial cancer risk loci identified through genome-wide association analysis. *Nat Genet*. 2016;48:667–74.
49. Cobo I, Martinelli P, Flandez M, Bakiri L, Zhang M, Carrillo-de-Santa-Pau E, et al. Transcriptional regulation by NR5A2 links differentiation and inflammation in the pancreas. *Nature*. 2018;554:533–7.
50. Michaud DS, Izard J, Wilhelm-Benartzi CS, You DH, Grote VA, Tjonneland A, et al. Plasma antibodies to oral bacteria and risk of pancreatic cancer in a large European prospective cohort study. *Gut*. 2013;62:1764–70.
51. Fan X, Alekseyenko AV, Wu J, Peters BA, Jacobs EJ, Gapstur SM, et al. Human oral microbiome and prospective risk for pancreatic cancer: a population-based nested case-control study. *Gut*. 2016;67:120–7.
52. Hayashi C, Gudino CV, Gibson FC 3rd, Genco CA. Review: pathogen-induced inflammation at sites distant from oral infection: bacterial persistence and induction of cell-specific innate immune inflammatory pathways. *Mol Oral Microbiol*. 2010;25:305–16.
53. Coussens LM, Werb Z. Inflammation and cancer. *Nature*. 2002;420:860–7.
54. Duell EJ, Casella DP, Burk RD, Kelsey KT, Holly EA. Inflammation, genetic polymorphisms in proinflammatory genes TNF-A, RANTES, and CCR5, and risk of pancreatic adenocarcinoma. *Cancer Epidemiol Biomark Prevent*. 2006;15:726–31.
55. Antwi SO, Bamlet WR, Pedersen KS, Chaffee KG, Risch HA, Shivappa N, et al. Pancreatic cancer risk is modulated by inflammatory potential of diet and ABO genotype: a consortia-based evaluation and replication study. *Carcinogenesis*. 2018;39:1056–67.
56. Fujimoto K, Polonsky KS. Pdx1 and other factors that regulate pancreatic beta-cell survival. *Diabetes Obes Metab*. 2009;11(Suppl 4):30–7.
57. Blüher M. Obesity: global epidemiology and pathogenesis. *Nat Rev Endocrinol*. 2019;15:288.
58. Tang H, Dong X, Hassan M, Abbruzzese JL, Li D. Body mass index and obesity- and diabetes-associated genotypes and risk for pancreatic cancer. *Cancer Epidemiol Biomark Prevent*. 2011;20:779–92.
59. Nakao M, Hosono S, Ito H, Watanabe M, Mizuno N, Yatabe Y, et al. Interaction between IGF-1 polymorphisms and overweight for the risk of pancreatic cancer in Japanese. *Int J Mol Epidemiol Genet*. 2011;2:354–66.
60. Orru S, Nigro E, Mandola A, Alfieri A, Buono P, Daniele A, et al. A functional interplay between IGF-1 and Adiponectin. *Int J Mol Sci*. 2017;18:2145.
61. 2018 Health Promotion Administration Annual Report. Health Promotion Administration. Taiwan; 2018. Available from: <https://www.hpa.gov.tw/Pages/List.aspx?nodeid=1248>. Accessed 10 June 2019.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

