Identification of common variants in BRCA2 and MAP2K4 for susceptibility to sporadic pancreatic cancer

Liming Huang1, Chen Wu1, Dianke Yu1, Chengfeng Wang1, Xu Che2, Xiaoping Miao2, Kan Zhai3, Jiang Chang4, Guoliang Jiang3, Xianghong Yang1, Guangwen Cao1, Zhibin Hu4, Yongjian Zhou5, Chaohui Zuo6, Chunyou Wang7, Xianghong Zhang11, Yifeng Zhou7, Xianjun Yu11, Wanjin Dai11, Zhaoshen Li11, Hongbing Shen1, Luming Liu15, Yanling Chen16, Sheng Zhang17, Xiaoji Wang18, Yu Liu19, Menghong Sun19, Wei Cao5, Jun Gao14, Ying Ma5, Xiongwei Zheng20, Siu Ming Cheung13, Yongfeng Jia21, Wen Tan1, Tongchun Wu6 and Dongxin Lin14,18

1State Key Laboratory of Molecular Oncology, Cancer Institute and Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China, 2Department of Abdominal Surgery, Cancer Institute and Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China, 3Key Laboratory for Environment and Health, School of Public Health, Tongji Medical College, Huazhong University of Sciences and Technology, Wuhan, China, 4Department of Radiation Oncology, Cancer Hospital, Fudan University, Shanghai, China, 5Department of Pathology, Shengjing Hospital, China Medical University, Shenyang, China, 6Department of Epidemiology, Second Military Medical University, Shanghai, China, 7Department of Epidemiology and Biostatistics, Cancer Center, Nanjing Medical University, Nanjing, China, 8Department of Gastrointestinal Surgery, Union Hospital of Fujian Medical University, Fuzhou, China, 9Department of Gastroenterology and Biostatistics, Huazhong University of Sciences and Technology, Wuhan, China, 10Department of Experimental Pathology, Hebei Medical University, Shijiazhuang, China, 11Laboratory of Cancer Molecular Genetics, Medical College of Soochow University, Suzhou, China, 12Department of Pancreas and Hepatobiliary Surgery, Cancer Hospital, Fudan University, Shanghai, China, 13Department of Gastroenterology, First Affiliated Hospital, Second Military Medical University, Shanghai, China, 14Department of Integrative Oncology, Cancer Hospital, Fudan University, Shanghai, China, 15Department of Hepatobiliary and Pancreatic Surgery, Union Hospital of Fujian Medical University, Fuzhou, China, 16Department of Pathology, Cancer Hospital, Chinese Academy of Medical Sciences (Beijing) and Cancer Hospital, Fudan University (Shanghai) between 2000 and 2011; controls were cancer-free individuals selected from community cancer screening programs for early detection of cancer (Beijing) or nutritional survey (Shanghai) conducted during the same time period as the patients were collected. The study subjects for the second-stage analysis consisted of 2603 cases and 2877 controls, which were recruited from 16 provinces or cities in China during the GWAS period. The diagnosis of pancreatic ductal adenocarcinoma was confirmed by histopathological or cytological analysis according to the World Health Organization classification. Demographic characteristics including sex and age were obtained from patient’s medical records. All controls were selected on the basis of physical examinations and frequency matched for sex and age to cases at recruitment, informed consent was obtained from each subject. This study was approved by the Institutional Review Board of Chinese Academy of Medical Sciences Cancer Institute.

Abbreviations: CI, confidence interval; cDNA, complementary DNA; FPRP, false-positive report probability; GWAS, genome-wide association studies; LD, linkage disequilibrium; OR, odds ratio; qPCR, quantitative PCR; SNP, single nucleotide polymorphism; UTR, untranslated region

# Materials and methods

## Study subjects

This study was an extension of our previous GWAS, in which the genome-wide scan sample comprised 981 pancreatic cancer cases and 1991 controls, and replication samples comprised 2603 cases and 2877 controls. The sources and characteristics of these study subjects were described previously (3). Briefly, for the genome-wide scan, cases were recruited in Cancer Hospital, Chinese Academy of Medical Sciences (Beijing) and Cancer Hospital, Fudan University (Shanghai) between 2000 and 2011; controls were cancer-free individuals selected from community cancer screening programs for early detection of cancer (Beijing) or nutritional survey (Shanghai) conducted during the same time period as the patients were collected. The study subjects for the second-stage analysis consisted of 2603 cases and 2877 controls, which were recruited from 16 provinces or cities in China during the GWAS period. The diagnosis of pancreatic ductal adenocarcinoma was confirmed by histopathological or cytological analysis according to the World Health Organization classification. Demographic characteristics including sex and age were obtained from patient’s medical records. All controls were selected on the basis of physical examinations and frequency matched for sex and age to cases at recruitment, informed consent was obtained from each subject. This study was approved by the Institutional Review Board of Chinese Academy of Medical Sciences Cancer Institute.

## Selection of candidate genes and SNPs and statistical analysis

We used NextBio Public database (http://www.nextbio.com/b/nextbio.nb) (7) to choose genes that have been well established to play critical roles in the pancreas that included 981 cases and 1991 controls in the first stage followed by a second stage (2603 cases and 2877 controls). Using an approach based on candidate genes whose roles in pancreatic cancer have been well known, we identified two new susceptibility loci. rs12939944 located in the MAP2K4 intron was associated with decreased risk (odds ratio = 0.82, 95% confidence interval = 0.74–0.91, P = 0.0001) in a dominant manner. Our results demonstrate for the first time that common variants in BRCA2 and MAP2K4 are susceptibility to sporadic pancreatic cancer.
development of pancreatic cancer as candidates. By using keyword 'pancreatic cancer', numerous genes appeared in the database and we selected the top 10 genes for genotyping in this study. They were TP53, KRAS, CDKN2A, SMAD4, MAP2K4, BRCA2, MEN1, STK11, PALLD and APC in order. Because MEN1, a gene mainly associated with pancreatic endocrine tumors, is often mutated and methylated in pancreatic ductal adenocarcinoma (8), it was also chosen as a candidate. All selection of markers was based on our previous GWAS scan results with standard quality-control process including testing Hardy-Weinberg equilibrium, principal component analysis and so on (3). We identified a total of 672 directly genotyped or well-imputed SNPs in these 10 genes and their 10kb upstream regions with minor allele frequency (MAF) of more than or equal to 0.05 for genetic association analysis (Supplementary Table I, available at Carcinogenesis Online). Associations were then analyzed by multivariate logistic-regression model using sex and age as covariates. In the first-stage analysis, we used dominant and recessive models in logistic regression to calculate P values and odds ratios (ORs) and their 95% confidence intervals (CIs) of each SNP. The lower P values from the two models were used to evaluate the statistical significance of the associations. We then calculated false-positive report probability (FPPR) for statistically significant associations as described by Wacholder et al. (9). In this study, we set a stringent significance value of FPPR at 0.20, which indicates that any finding with a FPPR value of less than 0.20 is significant. The prior probability was set at 0.10 for all SNPs and the statistical power to detect an OR of 1.50 (or its reciprocal 0.67) was used for calculating FPPR value. We also computed correlation coefficient (r) of each pair of adjacent SNPs on the same gene to assess the linkage disequilibrium (LD) status. SNPs with r^2 > 0.8 were considered to be in one LD block and we thus selected the most significant SNP (with lowest P value) in the block for replication. With this criterion, five SNPs were then selected for replication analysis. To assess the significance in the replication stage, we performed Bonferroni correction for multiple testing. LD structures of interest loci were generated using Haploview v4.2 software.

Genotype analysis
The preparation of genomic DNA, genotyping analysis of the GWAS samples, imputation and LD analysis have been described previously (3). Briefly, besides a part of genomic DNA samples of cases for the second-stage analysis (n = 814) was isolated from surgically removed and paraffin-embedded normal tissue adjacent to pancreatic cancer using KAPA Express Extract Kits (KAPA Biosystems, Woburn, MA), the remains were isolated from peripheral blood tissues of patients carrying the rs11571836 AG genotype (n = 16) and converted to complementary DNA (cDNA) using oligo(dT) primer and M-MLV Reverse Transcriptase (Promega). Allele-specific expression of BRCA2 RNA was analyzed by real-time quantitative PCR (qPCR) on an ABI 7900HT system as described by Chen et al. (3). The p-A and p-G plasmids as described above were used as standard samples to generate standard curves for rs11571836A or rs11571836G allele. The qPCR results were analyzed using the software SDS 2.3 (Applied Biosystems) and the BRCA2 transcript (expressed as cDNA) copy numbers for different rs11571836 alleles were estimated based on the respective standard curve. The ratio of rs11571836A transcript copy numbers to rs11571836A transcript copy numbers was calculated as an indicator of allele-specific expression levels. Because allele imbalance of genomic BRCA2 DNA might exist in individuals and would influence allele-specific BRCA2 transcript levels, we also examined BRCA2 DNA copy numbers in the same individual using allele-specific qPCR and took into account of DNA copy numbers when comparing allele-specific BRCA2 transcript copy numbers. If there is no allele expression imbalance, the ratio of rs11571836A transcript copy numbers to rs11571836A transcript copy numbers would be equal to the ratio of rs11571836A DNA copy numbers to rs11571836A DNA copy numbers. Therefore, if the SNPs were selected for qPCR determination of different allelic BRCA2 genomic DNA and transcript are shown in Supplementary Table II, available at Carcinogenesis Online.

Results
We first used NextBio Public database (7) to choose candidate genes that have been well recognized to play critical roles in the development of pancreatic cancer and 10 top genes, that is TP53, KRAS, CDKN2A, SMAD4, MAP2K4, BRCA2, MEN1, STK11, PALLD and APC, were identified. Then we analyzed the associations between pancreatic cancer risk and directly genotyped and well-imputed SNPs within each of these gene loci and its 10kb upstream region in our GWAS scan results (3) using both dominant and recessive genetic models in the multivariate logistic regression with sex and age as covariates. As a result, we found that among the identified 672 SNPs in these gene loci, 3 in KRAS, 1 in CDKN2A, 4 in SMAD4, 30 in MAP2K4, 5 in BRCA2, and 23 in PALLD were associated with the risk at P < 0.05 (Supplementary Table I, available at Carcinogenesis Online). However, no SNPs with P < 0.05 were found in TP53, MEN1, STK11 and APC. We then estimated each of these SNPs with P < 0.05 by FPPR and found that only two in MAP2K4 (rs12939944 and rs12942507), two in BRCA2 (rs9567639 and rs11571836) and two in PALLD (rs7666151 and rs13120709) passed the FPRP test (FPRP < 0.20; Table I and Supplementary Table I, available at Carcinogenesis Online), indicating that the associations of these SNPs with pancreatic cancer risk were promising for replication. Because the markers rs9567639 and rs11571836 in the BRCA2 gene are in perfect linkage disequilibrium (r^2 = 1.00), we only selected one tagging marker, rs11571836, for further analysis.

Replication of the five markers in the second-stage consisting of 2603 pancreatic cancer cases and 2877 controls verified rs12939944 in MAP2K4 and rs11571836 in BRCA2 significantly associated with pancreatic cancer risk in the same direction as observed in the GWAS scan (P = 0.0019 and P = 0.0041, respectively). However, the other three signals, rs12942507 in MAP2K4 and rs7666151 and rs13120709 in PALLD failed to be replicated (all P > 0.05). rs12939944 SNP in MAP2K4 was associated with a decreased pancreatic cancer risk in a dominant manner; in combined sample, the adjusted OR for individuals with at least one C allele (TC and CC genotype) was 0.82 (95% CI = 0.74–0.91, P = 0.0001) compared with those with the TT genotype. rs11571836 SNP in BRCA2 was significantly associated with an increased pancreatic cancer risk in a recessive manner; in combined sample, the adjusted OR for individuals with the GG genotype was 1.30 (95% CI = 1.14–1.47, P = 7.64 x 10^-5) compared
Polymorphisms in BRCA2 and MAP2K4 and pancreatic cancer

with those with at least one A allele (AA and AG genotype; Table I). Either in the second-stage analysis or in combined analysis, the associations for rs12939944 and rs11571836 were significant after Bonferroni correction for multiple testing.

Because rs11571836 is located in BRCA2 3′UTR, it might affect regulation of the gene. To investigate its potential function, we constructed two luciferase reporter plasmids consisting of the BRCA2 3′UTR with either the rs11571836A or rs11571836G allele and transiently transfected them into two pancreatic cancer cell lines, AsPC-1 and PANC-1. We found that reporter gene with the rs11571836A allele was nearly 1.00 (mean ± standard error, 4.46 ± 0.09 versus 2.24 ± 0.04 in AsPC-1 cells and 2.89 ± 0.09 versus 1.94 ± 0.13 in PANC-1 cells (both P < 0.0001, Figure 1A). These results suggest that the rs11571836 might have functional relevance.

To test whether rs11571836 SNP affects BRCA2 expression in vivo, we performed allele-specific quantitative PCR to measure copy number of the BRCA2 transcript and the corresponding genomic DNA with different rs11571836A or rs11571836G allele in normal pancreatic tissues adjacent to tumors obtained from different individuals with the heterozygous rs11571836AG genotype (n = 16). We found that the ratios of BRCA2 DNA copy number for the rs11571836G allele to that for the rs11571836A allele were nearly 1.00 (mean ± standard error, 1.00 ± 0.04) in all samples, indicating that no genomic copy number imbalance of different BRCA2 alleles existed. However, the BRCA2 cDNA copy number was significantly lower for the rs11571836G allele than that for the rs11571836A allele (mean ± standard error, 0.80 ± 0.05) (Figure 1B and Supplementary Table III, available at Carcinogenesis Online), suggesting a potential allele expression deficiency of BRCA2 for the rs11571836G allele compared with the rs11571836A allele.

Discussion

In this study, we employed a new candidate gene-based strategy to further discover susceptibility loci for pancreatic cancer using our previous GWAS data (3). By focusing on the top 10 genes that have been well established to play pivotal roles in the development of pancreatic cancer, we identified rs11571836 in BRCA2 and rs12939944 in MAP2K4 as new susceptibility loci for sporadic pancreatic cancer. Biochemical assays linked the rs11571836A>G change to decreased constitutional levels of BRCA2 transcript, which might be the underlying mechanism for the association of this SNP with susceptibility to pancreatic cancer. To the best of our knowledge, this is the first study showing that genetic polymorphisms in BRCA2 and MAP2K4 as well are associated with susceptibility to sporadic pancreatic cancer.

The main function of BRCA2 is to maintain genome stability by repairing DNA damage and deficiency of BRCA2 leads to the accumulation of chromosomal aberrations in cells (14,15). More importantly, germline mutations of the BRCA2 gene have been shown to cause hereditary breast and ovarian cancer syndrome and linked to 5- to 10-fold increased risk of familial pancreatic cancer (2,16,17). These findings strongly support our results that genetic variation in BRCA2 is associated with susceptibility to sporadic pancreatic cancer. The SNP rs11571836 is located in the BRCA2 3′UTR and might influence BRCA2 expression via potential mechanisms such as creating a micro-RNA binding site and affecting 3′ end formation of pre-messenger RNA or messenger RNA secondary structure as reported in several studies (16,18,19). Indeed, our reporter gene assays and allele-specific expression analysis of BRCA2 in normal pancreatic tissues demonstrated that the rs11571836 SNP is functional relevant and the G allele had a significantly lower BRCA2 messenger RNA expression compared with the rs11571836A allele. These results are consistent with our results, indicating the G allele as risk allele. Because BRCA2 is a tumor suppressor, subtle deficiency of it due to functional genetic polymorphism would be anticipated to confer susceptibility to the cancer. We observed that the rs11571836 SNP confers susceptibility to pancreatic cancer in a recessive manner, which is consistent with the nature of action of a tumor suppressor.

MAP2K4 is a member of the mitogen-activated protein kinase family and acts as a direct activator of mitogen-activated protein kinases

<table>
<thead>
<tr>
<th>Genotype</th>
<th>GWAS stage (981/1991)*</th>
<th>Replication stage (2603/2877)*</th>
<th>Combined sample (3584/4866)*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N*</td>
<td>OR (95% CI)</td>
<td>P</td>
</tr>
<tr>
<td>MAP2K4, rs12939944</td>
<td>TT</td>
<td>360/652</td>
<td>1.00 (Reference)</td>
</tr>
<tr>
<td></td>
<td>TC</td>
<td>454/1001</td>
<td>0.82 (0.69–0.97)</td>
</tr>
<tr>
<td></td>
<td>CC</td>
<td>167/338</td>
<td>0.89 (0.71–1.12)</td>
</tr>
<tr>
<td></td>
<td>Dominant model</td>
<td>454/338</td>
<td>0.89 (0.71–1.12)</td>
</tr>
<tr>
<td>MAP2K4, rs12942507</td>
<td>TT</td>
<td>292/512</td>
<td>1.00 (Reference)</td>
</tr>
<tr>
<td></td>
<td>TC</td>
<td>443/797</td>
<td>0.80 (0.67–0.96)</td>
</tr>
<tr>
<td></td>
<td>CC</td>
<td>237/492</td>
<td>0.83 (0.67–1.03)</td>
</tr>
<tr>
<td></td>
<td>Dominant model</td>
<td>443/492</td>
<td>0.83 (0.67–1.03)</td>
</tr>
<tr>
<td>BRCA2, rs11571836</td>
<td>AA</td>
<td>379/822</td>
<td>1.00 (Reference)</td>
</tr>
<tr>
<td></td>
<td>AG</td>
<td>458/945</td>
<td>1.03 (0.88–1.22)</td>
</tr>
<tr>
<td></td>
<td>GG</td>
<td>144/224</td>
<td>1.39 (1.09–1.77)</td>
</tr>
<tr>
<td></td>
<td>Recessive model</td>
<td>144/224</td>
<td>1.39 (1.09–1.77)</td>
</tr>
<tr>
<td>PALLD, rs7666151</td>
<td>GG</td>
<td>811/1721</td>
<td>1.00 (Reference)</td>
</tr>
<tr>
<td></td>
<td>GT</td>
<td>175/252</td>
<td>1.29 (1.04–1.61)</td>
</tr>
<tr>
<td></td>
<td>TT</td>
<td>131/147</td>
<td>1.72 (1.43–2.07)</td>
</tr>
<tr>
<td></td>
<td>Dominant model</td>
<td>175/147</td>
<td>1.72 (1.43–2.07)</td>
</tr>
<tr>
<td>PALLD, rs13120709</td>
<td>TT</td>
<td>302/643</td>
<td>1.00 (Reference)</td>
</tr>
<tr>
<td></td>
<td>TC</td>
<td>453/964</td>
<td>1.02 (0.85–1.22)</td>
</tr>
<tr>
<td></td>
<td>CC</td>
<td>226/384</td>
<td>1.26 (1.02–1.56)</td>
</tr>
<tr>
<td></td>
<td>Dominant model</td>
<td>453/384</td>
<td>1.26 (1.02–1.56)</td>
</tr>
</tbody>
</table>

P values are two sided and were calculated by logistic-regression analysis with age and sex as covariates.

*Number of cases/number of controls.

#Significant after Bonferroni correction for multiple test.
in response to various environmental stresses or mitogenic stimuli, such as growth factors, hormones, and pro-inflammatory cytokines (20). Although important, the roles that MAP2K4 plays in cancer development appear complicated and contradictory results have been reported over the past decade. It was found that about 5% of several types of cancer including pancreatic cancer had genetic inactivation of MAP2K4 (21–23) and dominant-negative mutant of MAP2K4 could promote cell transformation and tumorigenicity (24). Furthermore, it has been shown that loss of MAP2K4 protein was associated with distant metastasis and bad survival of pancreatic cancer (25). All of these studies imply MAP2K4 a tumor suppressor. However, other studies provided evidence supporting a pro-oncogenic role of MAP2K4. It was shown that ectopic expression of MAP2K4 in MAP2K4-negative breast and pancreatic cancer cell lines stimulates cell proliferation and invasion, and knockdown of MAP2K4 expression led to suppressed tumor growth in mouse xenograft model and increased cell susceptibility to apoptosis (26). Another study reported that MAP2K4-null pancreatic cancer cells produced fewer metastases in vivo compared with wild-type cells (27). In the present study, we observed a significant association between risk of pancreatic cancer and rs11571836A (mean ratio ± standard error = 0.80 ± 0.05), whereas the copy numbers of rs11571836G cDNA were significantly lower than that of rs11571836A = 16). The copy numbers of the heterozygous rs11571836AG genotype (corresponding genomic DNA (gDNA) samples from pancreatic tissues with transcript (cDNA) and BRCA2 (20)). Allele-specific expression of BRCA2 rs11571836G and rs11571836A.

A previous study in European ancestry reported null significant association between risk of pancreatic cancer and SNPs in hereditary pancreatic cancer genes including BRCA2 (29). The discrepancy between their results and ours might reflect the ethnic-specific nature of cancer susceptibility. It is important to note that genotype frequencies of rs11571836 are 70.8% (AA), 27.4% (AG) and 1.8% (GG) in population of European ancestry (HapMap phase III databases) compared with 42.6% (AA), 45.7% (AG) and 11.7% (GG) in our control population of Chinese ancestry. Given extremely rare frequency of the 11571836 GG genotype in populations of European ancestry, it would not be surprising to see null association in the study where statistical power was limited (29). Germline mutations in CDKN2A, STK11 and APC have been correlated to unusual high lifetime risk of developing familial pancreatic cancer (30–33). However, in the current study, we did found any significant association of SNPs in these genes with risk of pancreatic cancer, which is consistent with the previous study (29).

The evaluation of susceptibility loci in GWAS necessitates the use of stringent P values to prevent false-positive results; however, this strategy might yield false-negative results (6). Thus, other empirical complementary approaches are required to discover additional loci. Indeed, employing an empirical approach focusing on genes that their roles in pancreatic cancer are well established, we have identified rs11571836 in BRCA2 and rs12939944 in MAP2K4, which had not reached genome-wide significance in our previous GWAS (3), as new susceptibility loci. These results demonstrated that our approach in the present study is useful and can be used as a complementary approach in the analysis of GWAS data.

In conclusion, in our two-stage association study with candidate gene approach, we have identified for the first time two new susceptibility loci for pancreatic cancer in Chinese populations. Strikingly, one of these susceptibility loci is BRCA2, whose germline mutations have been well known to cause hereditary breast and ovarian cancer syndrome and familial pancreatic cancer. These results have substantially extended our previous findings and advanced our understanding of the etiology of pancreatic cancer.

Supplementary material
Supplementary Tables I–III can be found at http://carcin.oxfordjournals.org/

Funding
National High-Tech Research and Development Program of China (2009AA022706); National Natural Science Foundation of China (81021061).

Acknowledgements
We thank the subjects and the surgeons who recruited them: M. Yang, L. Wang, X. Wang, C. Yang, L. Du, J. Li, F. Qi, X. Song and L. Zeng.

Conflict of Interest Statement: None declared.

References

Received November 22, 2012; accepted December 30, 2012