A pleiotropy scan to discover new susceptibility loci for pancreatic ductal adenocarcinoma

Giaccherini M.[1#], Rende M.[1#], Gentiluomo M.[1], Corradi C.[1], Archibugi L.[2,3], Ermini S.[4], Maiello E.[5], Morelli L.[6], van Eijck C.H.J.[7], Cavestro G.M.[8], Schneider M.[9], Mickevicius A.[10], Adamonis K.[11], Basso D.[12], Hlavac V.[13], Gioffreda D.[14], Talar-Wojnarowska R.[15], Schöttker B.[16,17], Lovecek M.[18], Vanella G.[2,3], Gazouli M.[19], Uno M.[20], Malecka-Wojciesko E.[15], Vodicka P.[21,22], Goetz M.[23], Bijlsma M.F.[24,25], Petrone M.C.[3], Bazzocchi F.[26], Kiudelis M.[10], Szentesi A.[27,28,29,30], Carrara S.[31], Nappo G.[32,33], Brenner H.[16,34,35], Milanetto A.C.[12], Soucek P.[13], Katzke V.[36], Peduzzi G.[1], Rizzato C.[1], Pasquali C.[12], Chen X.[16,37], Capurso G.[2,3], Hackert T.[9], Bueno-de-Mesquita B.[38], G. Uzunoglu F.G.[23], Hegyi P.[27,28,29,39], Greenhalf W.[40], E. Theodoropoulos G.E.[41], Sperti C.[12], Perri F.[14], Oliverius M.[42], Mambrini A.[43], Tavano F.[14], Farinella R.[1], Arcidiacono P.G.[3], Lucchesi M.[43], Bunduc S.[27,29,39,44,45], Kupcinskas J.[11], Di Franco G.[6], Stocker S.[16,17], Neoptolemos J.P.[9], Bambi F.[4], Jamroziak K.[46], Testoni S.G.G.[3], Aoki M.N.[47], Mohelnikova-Duchonova B.[48], Izbicki J.R.[23], Pezzilli R.[49], Lawlor R.T.[50], Kauffmann E.F.[51], López de Maturana E.[52], Malats N.[52], Canzian F.[53], Campa D.[1*]

1 Department of Biology, University of Pisa, Pisa, Italy

2 Digestive and Liver Disease Unit, Sant'Andrea Hospital, Rome, Italy

3 Pancreato-Biliary Endoscopy and Endosonography Division, Pancreas Translational and Clinical Research Center, IRSSC San Raffaele Scientific Institute, Milan, Italy

4 Blood Transfusion Service, Azienda Ospedaliero Universitaria Meyer, Florence, Italy

5 Department of Oncology, Fondazione IRCCS "Casa Sollievo della Sofferenza" Hospital, San Giovanni Rotondo, Italy

6 General Surgery Unit, Department of Translational Research and New Technologies in Medicine and Surgery, University of Pisa, Pisa, Italy

7 Department of Surgery, Erasmus MC University Medical Center, Rotterdam, The Netherlands

8 Gastrointestinal Endoscopy Unit, Vita-Salute San Raffaele University, IRCCS San Raffaele, Milan, Italy

9 Department of General, Visceral and Transplantation Surgery, Heidelberg University Hospital, Heidelberg, Germany

10 Surgery Department, Lithuanian University of Health Sciences, Kaunas, Lithuania

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

© The Author(s) 2024. Published by Oxford University Press on behalf of the UK Environmental Mutagen Society. All rights reserved. For commercial re-use, please contact reprints@oup.com for reprints and translation rights for reprints. All other permissions can be obtained through our RightsLink service via the Permissions link on the article page on our site—for further information please contact journals.permissions@oup.com.

Accepted Manuscript

Downloaded from https://academic.oup.com/mutage/advance-article/doi/10.1093/mutage/geae012/7644568 by Szegedi Tudomanyegyetem / University of Szeged user on 22 April 2024

12 Department of Surgery, Oncology and Gastroenterology-DiSCOG, University of Padova, Padua, Italy

13 Biomedical Center, Faculty of Medicine in Pilsen, Charles University, Pilsen, Czech Republic

14 Division of Gastroenterology and Research Laboratory, Fondazione IRCCS “Casa Sollievo della Sofferenza” Hospital, San Giovanni Rotondo, Italy

15 Department of Digestive Tract Diseases, Medical University of Lodz, Lodz, Poland

16 Division of Clinical Epidemiology and Aging Research, German Cancer Research Center (DKFZ), Heidelberg, Germany

17 Network Aging Research (NAR), Heidelberg University, Heidelberg, Germany

18 Department of Surgery I, University Hospital Olomouc, Olomouc, Czech Republic

19 Department of Basic Medical Sciences, Laboratory of Biology, Medical School, National and Kapodistrian University of Athens, Athens, Greece

20 Center for Translational Research in Oncology (LIM24), Instituto Do Câncer Do Estado de São Paulo, (ICESP), Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo (HCFMUSP), São Paulo, Brazil

21 Institute of Experimental Medicine, Czech Academy of Science, Prague, Czech Republic

22 Institute of Biology and Medical Genetics, 1st Medical Faculty, Charles University in Prague, Prague, Czech Republic

23 Department of General, Visceral and Thoracic Surgery, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

24 Laboratory for Experimental Oncology and Radiobiology, Center for Experimental and Molecular Medicine, Amsterdam UMC and Cancer Center Amsterdam, Amsterdam, The Netherlands

25 Oncode Institute, Amsterdam, The Netherlands

26 Department of Surgery, Fondazione IRCCS “Casa Sollievo della Sofferenza” Hospital, San Giovanni Rotondo, Italy

27 Institute for Translational Medicine, Medical School, University of Pécs, Pécs, Hungary

28 János Szentágothai Research Center, University of Pécs, Pécs, Hungary

29 Center for Translational Medicine, Semmelweis University, Budapest, Hungary

30 Centre for Translational Medicine, Department of Medicine, University of Szeged, Szeged, Hungary

31 Digestive Endoscopy Unit, Division of Gastroenterology, Humanitas Clinical and Research Center IRCCS, Milan, Italy

Accepted Manuscript

Downloaded from https://academic.oup.com/mutage/advance-article/doi/10.1093/mutage/geae012/7644568 by Szegedi Tudomanyegyetem / University of Szeged user on 22 April 2024

32 Pancreatic Unit, Humanitas Clinical and Research Center IRCCS, Milan, Italy

33 Department of Biomedical Sciences, Humanitas University, Milan, Italy

34 Division of Preventive Oncology, German Cancer Research Center (DKFZ) and National Center for Tumor Diseases (NCT), Heidelberg, Germany

35 German Cancer Consortium (DKTK), German Cancer Research Center (DKFZ), Heidelberg, Germany

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

37 Medical Faculty Heidelberg, Heidelberg University, Heidelberg, Germany

38 Centre for Nutrition, Prevention and Health Services, National Institute for Public Health and the Environment (RIVM), Bilthoven, The Netherlands

39 Division of Pancreatic Diseases, Heart and Vascular Center, Semmelweis University, Budapest, Hungary

40 Institute for Health Research Liverpool Pancreas Biomedical Research Unit, University of Liverpool, Liverpool, United Kingdom

41 First Department of Propaedeutic Surgery, Hippocration General Hospital, National and Kapodistrian University of Athens, Athens, Greece

42 Surgery Clinic Third Faculty of Medicine, Charles University, Prague, Czech Republic

43 Oncological Department Massa Carrara, Azienda USL Toscana Nord Ovest, Carrara, Italy

44 Carol Davila University of Medicine and Pharmacy, Bucharest, Romania

45 Fundeni Clinical Institute, Bucharest, Romania

46 Department of Hematology, Transplantology and Internal Medicine, Medical University of Warsaw, Warsaw, Poland

47 Laboratory for Applied Science and Technology in Health, Carlos Chagas Institute, Oswaldo Cruz Foundation (Fiocruz), Curitiba, Brazil

48 Department of Oncology, Palacky University Olomouc, Olomouc, Czech Republic

49 County Medical Association of Potenza, Potenza, Italy

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

51 Division of General and Transplant Surgery, Pisa University Hospital, Pisa, Italy

52 Genetic and Molecular Epidemiology Group, Spanish National Cancer Research Centre (CNIO), Madrid, Spain

Accepted Manuscript

Downloaded from https://academic.oup.com/mutage/advance-article/doi/10.1093/mutage/geae012/7644568 by Szegedi Tudomanyegyetem / University of Szeged user on 22 April 2024

53 Genomic Epidemiology Group, German Cancer Research Center (DKFZ), Heidelberg, Germany

[#]These authors share the first position

Corresponding author:

Daniele Campa
Department of Biology
University of Pisa
Via Derna 1
56126 Pisa
Italy
Email: daniele.campa@unipi.it
Phone +39-050-2211510

Accepted Manuscript

Downloaded from https://academic.oup.com/mutage/advance-article/doi/10.1093/mutage/geae012/7644568 by Szegedi Tudomanyegyetem / University of Szeged user on 22 April 2024

Abstract

Pleiotropic variants (*i.e.*, genetic polymorphisms influencing more than one phenotype) are often associated with cancer risk. A scan of pleiotropic variants was successfully conducted ten years ago in relation to pancreatic ductal adenocarcinoma susceptibility. However, in the last decade, genetic association studies performed on several human traits have greatly increased the number of known pleiotropic variants. Based on the hypothesis that variants already associated with a least one trait have a higher probability of association with other traits, 61,052 variants reported to be associated by at least one genome wide association study (GWAS) with at least one human trait were tested in the present study consisting of two phases (discovery and validation), comprising a total of 16,055 pancreatic ductal adenocarcinoma (PDAC) cases and 212,149 controls. The meta-analysis of the two phases showed two loci (10q21.1-rs4948550 ($P=6.52\times10^{-5}$) and 7q36.3-rs288762 ($P=3.03\times10^{-5}$) potentially associated with PDAC risk. 10q21.1-rs4948550 shows a high degree of pleiotropy and it is also associated with colorectal cancer risk while 7q36.3-rs288762 is situated 28,558 base pairs upstream of the Sonic Hedgehog (*SHH*) gene, which is involved in the cell differentiation process and PDAC etiopathogenesis. In conclusion, none of the single nucleotide polymorphisms (SNPs) showed a formally statistically significant association after correction for multiple testing. However, given their pleiotropic nature and association with various human traits including colorectal cancer, the two SNPs showing the best associations with PDAC risk merit further investigation through fine mapping and ad hoc functional studies.

Key words: pleiotropy, pancreatic cancer, single nucleotide polymorphism, genetic susceptibility

Accepted Manuscript

Downloaded from https://academic.oup.com/mutage/advance-article/doi/10.1093/mutage/geae012/7644568 by Szegedi Tudomanyegyetem / University of Szeged user on 22 April 2024

Introduction

Pancreatic ductal adenocarcinoma (PDAC) is a relatively rare disease with a crude incidence rate of 18.7 per 100,000 individuals per year in Europe[1]. Only a small number of potential PDAC risk factors have been identified, such as cigarette smoking, type 2 diabetes mellitus, chronic pancreatitis, overweight and non-O blood groups [2–4]. Recently a study carried out in the context of UK Biobiank (UKBB) has suggested stress as a major contributor for the development of the disease [5]. The genetic susceptibility to PDAC is due to rare high-penetrance mutations and common low-penetrance genetic variants, that alone or in combination are associated with increased risk of developing PDAC [6–24]. However, PDAC is a polygenic and complex multifactorial disease that shares a portion of the genetic background with several human traits [25,26]. For example, the *TERT-CLPTM1L* region is known to be associated with PDAC risk, but it is also associated with risk of melanoma, breast, and cervical cancers [26]. A genetic variant independently associated with more than one trait is defined as pleiotropic, a characteristic shared by many risk loci for a large number of human traits [27]. Pleiotropic polymorphisms could have a crucial role in the genetic architecture of complex diseases due to their influence on different pathways and biological mechanisms [28–30]. Pleiotropic single nucleotide polymorphisms (SNPs) are common in cancer, and there are regions in the genome called "nexus" that are associated with more than one type of cancer [31,32]. In addition, several PDAC risk loci are also associated with non-cancer phenotypes. For example, *ABO* and *TERT* SNPs are associated with a plethora of human traits, such as longevity [33], type II diabetes [34], male infertility [35], mitochondrial DNA copy number [36], and high-density lipoprotein cholesterol level [37]. Therefore, the study of pleiotropic SNPs could be instrumental in unravelling the genetic architecture of human diseases, as SNPs that are already associated with one trait have an increased chance of being associated with other phenotypes. The analysis of the possible association of pleiotropic variants with PDAC was completed in a study comprising 1,087 SNPs in 2,857 PDAC cases and 2,967 controls, in which the authors identified, for the first time, a new PDAC risk locus (rs7310409) in the HNF1 homeobox A *(HNF1A)* gene [38]. In the decade since that study our knowledge of SNP-phenotype associations has greatly increased. With these premises, we aimed to identify novel pleiotropic SNPs associated with PDAC risk in an extensive multi-ethnic study of 16,880 PDAC cases and 219,861 controls.

Material and methods

Study design

The present study was carried out in two phases. First, a discovery phase where pleiotropic SNPs reported to be associated with at least one human trait were tested for association in a case-control study consisting of 8,738 PDAC cases and 7,034 controls. The discovery phase consisted of the PanScan I, II, III [9,11,12], and PanC4 studies [14]that were imputed separately and then merged and analysed together. After the discovery phase two rounds of replication were carried out, the first (replication1) consisting of theEuropean Study into Digestive Illnesses and Genetics (PanGenEU) study [39] and the second (replication2) consisting of the Pancreatic Disease Research (PANDoRA) consortium [24], the

Accepted Manuscript

Downloaded from https://academic.oup.com/mutage/advance-article/doi/10.1093/mutage/geae012/7644568 by Szegedi Tudomanyegyetem / University of Szeged user on 22 April 2024

Japan Pancreatic Cancer Research Consortium (JaPAN) [7,10,40,41] and FinnGen [42]. Replication 2 was carried out in all SNPs that were significant in the discovery and replication 1 phases. **Table 1** shows the number of cases and controls analysed in each dataset, alongside age and sex distribution.

Discovery phase

A list of SNPs associated with at least one human trait at genome-wide significance level ($P<5x10^{-8}$) was obtained from the GWAS Catalog portal. The list contained 126,080 SNP-trait associations, consisting of 73,700 unique SNPs, among which 1,869 did not have reference SNP ID number (rs#) and therefore were not included in the analyses. The list included SNPs associated with any disease and/or any trait and was not restricted by ethnicity. All the selected SNPs were analysed using the genotypes of the PanScan I, II, III and PanC4 GWASs. The genotypes were downloaded from the database of Genotypes and Phenotypes (dbGaP; study accession nos. phs000206.v5.p3 and phs000648.v1.p1; project reference no. 12644). Genotyping and quality control details of these studies have been described in the original publications [9,11,12,14].

The four combined datasets included 9,563 PDAC cases and 8,073 controls. The genotypes were imputed using the Michigan Imputation Server (https://imputationserver.sph.umich.edu), and the Haplotype Reference Consortium (HRC, V.r1.1) as reference panel. The imputation for PanScan I, II, III and PanC4 GWASs was carried out separately for each dataset.

Before imputation, the datasets were filtered applying the following quality controls: removal of individuals with sex mismatches, missing genotypes >2%, relatedness issues (PI_HAT>0.2) and minimal or excessive heterozygosity (>3 standard deviations from the mean). Additionally, the SNPs with a minor allele frequency (MAF) <0.01, call-rate<98%, and evidence for violations of Hardy-Weinberg equilibrium (HWE, $P<1\times10^{-5}$) were discarded. Principal component analysis (PCA) was performed with PLINK 2.0, including the genotypes of phase 3 of the 1000 Genomes Project as reference panel [43]. Individuals not clustering in the PCA with the 1000 Genomes subjects of European descent were excluded from further analysis. After imputation the four datasets were merged using only the SNPs with imputation quality (INFO score r^2) higher than 0.7 (N= 24,735,918 SNPs). The pooled dataset was filtered, removing the variants with call rate<98% (N=11,699,683 SNPs), MAF<1% (N=5,524,684 SNPS) or departure from HWE ($P<1x10^{-5}$, N=2,206 SNPs). The discovery dataset consisted therefore of 8,738 PDAC case and 7,034 controls that were analysed for 7,509,345 SNPs. The "inflation factor" calculated in each dataset did not show evidence of systematic inflation (λ=1.000 for PanScan I, λ=1.015 for PanScan II, λ=1.026 for PanScan III, λ=1.000 for PanC4, and λ=1.000 for the aggregate dataset).

Replication phase

In the replication phase, data obtained from four independent populations were analysed, using the summary statistics of three studies: (I) PanGenEU, (II) FinnGen, and (III) JaPAN. Additionally, the SNPs to be validated were genotyped in the PANDoRA consortium. All populations have been described in detail elsewhere [23,24,41,42,44]. A brief description is also given in **supplementary material 1**.

Accepted Manuscript

Downloaded from https://academic.oup.com/mutage/advance-article/doi/10.1093/mutage/geae012/7644568 by Szegedi Tudomanyegyetem / University of Szeged user on 22 April 2024

Summary statistics of PanGenEU were used as first replication (Replication 1), then the variants that showed a statistically significant association (P<0.05) in PanGenEU were genotyped in PANDoRA and looked up in FinnGen and JaPAN (Replication 2) for a total of 7,317 cases and 212,142 controls comprised in the four populations.

PANDoRA sample preparation and genotyping

DNA of cases and controls from PANDoRA was extracted from whole blood, using the QIamp® 96 DNA QIAcube® HT Kit (Qiagen, Hilden, Germany). Genotyping was performed using TaqMan technology (ThermoFisher Applied Biosystems, Waltham MA, USA) in 384-well plates. A similar number of cases and controls was distributed in each plate, and duplicate samples (8%) were added for quality control purposes. Genotypes were determined using the QuantStudio™ 5 Real-Time PCR system (Thermofisher, USA).

Statistical analysis

The association between SNPs and risk of developing PDAC was evaluated through unconditional logistic regression analysis adjusted for sex, age and the top eight principal components for the discovery phase (PanScan I, II, III and PanC4) and for sex, age, and country of origin in the replication phase (PANDoRA). All the statistical analyses were conducted using PLINK 2.0 and R software. The details on the statistical analyses adopted in the PanGenEU, JaPAN and FinnGen GWASs are reported elsewhere [23,41,42]. A meta-analysis was performed for all the variants that showed a statistically significant association in PanGenEU using all the populations (PanScan I, II, III, PanC4, PanGenEU, PANDoRA, JaPAN and FinnGen). Stratified analysis including only European individuals was also performed to avoid confounding bias due to the different ethnic groups. To account for multiple testing, we considered linkage disequilibrium (LDr^2>0.8; 1000 genomes, Europeans) among the SNPs used in the discovery phase to obtain a list of independent variants (N=37,435). The threshold for statistical significance was, therefore, set to $P=0.05/37,435=1.34\times10^{-6}$ using Bonferroni's correction.

Results

Figure 1 shows the flowchart of the study design and replication across the populations. During the discovery phase 73,700 unique SNPs associated with at least one human trait were identified in GWAS Catalog and analysed in the Pancreatic Cancer Cohort Consortium (PanScan I, II and III), and the Pancreatic Cancer Case-Control Consortium (PanC4) datasets, for a total of 8,738 PDAC cases and 7,034 controls; 12,728 variants and their LD proxies (r^2>0.8) were not present in the dataset. Among the 61,052 remaining variants, 428 showed a statistically significant association with PDAC risk (P<0.05) (**Supplementary Table 1**). Among these, 164 SNPs were in LD (r^2>0.8) with known PDAC risk loci, and the remaining 264 variants were pruned (r^2>0.6) to eliminate residual LD and to identify independent SNPs

Accepted Manuscript

Downloaded from https://academic.oup.com/mutage/advance-article/doi/10.1093/mutage/geae012/7644568 by Szegedi Tudomanyegyetem / University of Szeged user on 22 April 2024

to be validated. The final list of SNPs to be further validated consisted of 113 SNPs (**Supplementary Table 1**).

These 113 SNPs were tested in PanGenEU, where seven variants showed an association (P<0.05). One SNP (11p14.2-rs117551578) showed a statistically significant association with the risk of developing PDAC in both the discovery phase and in PanGenEU, but in the opposite direction, therefore it was excluded from the subsequent analyses. **Table 2** shows the seven SNPs associated in PanGenEU, their P-value of association with PDAC risk and the P-value of association with the trait for which they were originally selected.

The remaining six SNPs (6q22.32-rs6919397; 8p23.1-rs2980752; 7q36.3-rs288762; 10q21.1-rs4948550; 12q12-rs12427164; 17q23.2-rs9903801) were analysed in replication 2 (described in the methods), that included genotypes of PANDoRA and summary statistics from FinnGen and JaPAN and then meta-analysed. None of the SNPs showed a statistically significant association in the studies belonging to replication 2. Furthermore, the overall meta-analysis, performed including a total of 16,055 PDAC cases and 212,149 controls, did not show any statistically significant association, considering Bonferroni correction for multiple testing (Table 3). Excluding non-European (Brazilian from PANDoRA and Japanese from JaPAN) individuals from the meta-analysis, lower P-values compared to the discovery phase were observed for 7q36.3-rs288762 (OR=1.08, 95%CI=1.04-1.12, $P=3.03\times10^{-5}$) and for 10q21.1-rs4948550 (OR=0.92, 95%CI=0.89-0.96, $P=6.52\times10^{-5}$).

Discussion

Among the six SNPs selected for the replication phase in PANDoRA, JaPAN, and FinnGen, none showed significant associations (P<0.05) in the studies taken individually; however, 10q21.1-rs4948550 and 7q36.3-rs288762 showed a lower P-value in the meta-analysis compared with the discovery phase. In addition, removing individuals of non-European or of admixed ancestry from the analysis, the significance level of the results improved further for 7q36.3-rs288762 ($P_{value}=3.03\times10^{-5}$) and for 10q21.1-rs4948550 ($P_{value}=6.52\times10^{-5}$). This improvement could be explained by the fact that the two SNPs are just risk markers, not directly responsible for the disease and, therefore, the presence of a different LD architecture across populations could dilute the results when considering different ethnicities together. The causative SNP could be linked to the markers in central Europeans but not in Asians, or Brazilians, or in LD with a variant that was not genotyped and that is in different LD blocks in the various populations. However, none of the variants reached the P-value threshold set for this study, considering the correction for multiple testing ($P=1.34\times10^{-6}$).

The best association was observed for 10q21.1-rs4948550 which is located in the BicC Family RNA Binding Protein 1 (*BICC1*) gene, that encodes an RNA-binding protein that regulates cell proliferation and apoptosis [45]. This SNP is a missense variant that leads to the aminoacidic variation Ser943Pro, for which a benign clinical significance is reported on NCBI dbSNP portal (www.ncbi.nlm.nih.gov). The 10q21.1-rs4948550 region shows a high degree of pleiotropy since the SNP (or the SNPs in LD with it) is associated with cardiovascular disease, bilateral cleft lip, morningness (*i.e.,* the individual preference of

Accepted Manuscript

Downloaded from https://academic.oup.com/mutage/advance-article/doi/10.1093/mutage/geae012/7644568 by Szegedi Tudomanyegyetem / University of Szeged user on 22 April 2024

waking up early) and colorectal cancer [46–48]. In particular, 10q21.1-rs4948550 is in strong LD with rs4948317 (r^2=0.82, D'=0.98 in Europeans) which was identified to be associated with risk of colorectal cancer (CRC, P=7x10^{-8}) in a GWAS study carried out in Est Asian individuals. The connection between CRC and PDAC is intriguing because the two tumours share several risk loci, for example 5p15.33-*TERT*, 16q24.1-*LINC01081/LINC00917*, 7p12.3-*TNS3* and the region of *ABO* on 9q34 [9,12,49–52]. Therefore, the results of our study may suggest another potential pleiotropic locus shared by PDAC and CRC, highlighting the importance of pleiotropy in human neoplastic diseases and a possible overlap in pathways and mechanisms that lead to the development of the two diseases.

The other potentially interesting SNP, 7q36.3-rs288762, is situated 28,558 base pairs upstream of the sonic hedgehog (*SHH*) gene, that is involved in cell differentiation. *SHH* has been observed to be overexpressed in cancer patients, thus creating a favourable environment for metastasis, proliferation, and drug resistance [11,53,54]. Three SNPs located in the *SHH* gene (rs167020, rs172310, rs288746) are already known to be associated with the risk of developing PDAC [12], highlighting the importance of the genetic variability of this region in PDAC. These three SNPs are in weak LD withrs288762, (r^2<0.35 in the European population). This region is particularly interesting due to the different results obtained with cohort and case-control studies. Amundadottir and colleagues observed a strong association for rs167020 (P=1.76x10^{-7}), rs172310 (P=2.01x10^{-7}) and rs288746 (P=1.35x10^{-4}) in prospective cohorts included in PanScan-I, but the associations were not confirmed in the retrospective studies used for replication (P=0.122, P=0.095 and P=0.108, respectively) [12]. Similarly, in our study, 7q36.3-rs288762 was observed to be associated with PDAC risk in the discovery phase (P=1.09×10^{-4}), that included mainly prospective studies, among which those analysed in 2009 by Amundadottir and colleagues, but not in the European retrospective populations used in the replication phase (PanGenEU P=0.03, PANDoRA P=0.66 and FinnGen P=0.37). Interestingly, Regan et al. observed that non canonical expression of SHH pathways positively regulates WNT signalling and may be crucial for colon cancer stem cell survival once again highlighting a connection between PDAC and CRC. [55]. Moreover, the T allele of rs288762, associated with an increased risk of developing PDAC, is also significantly associated with a low estimated glomerular filtration rate. This trait was already found to be associated with one (rs9903801) of the six SNPs analysed in Replication 2, suggesting a possible correlation between the risk of developing PDAC and a low estimated glomerular filtration rate. Additionally, it is interesting to note that C-reactive protein/albumin ratio is a predictor of pancreatic cancer survival, however further studies are warranted to characterise the mechanism linking these two traits.

A clear strength of this study is the large sample size, with 16,055 cases of PDAC and 212,149 controls, and a rigorous multi-phase process to eliminate spurious findings.

A possible limitation is the lack of a prospective cohort to replicate the finding on 7q36.2 to shed more light on the association of rs288762, located at this locus, since it appears to have different effect in prospective and retrospective studies. Another possible limitation is the fact that in the FinnGen study exocrine and endocrine pancreatic cancers are not divided. However, considering the very low prevalence of the latter (less than 2% of all pancreatic cancers), it is unlikely that this would have changed our results.

Accepted Manuscript

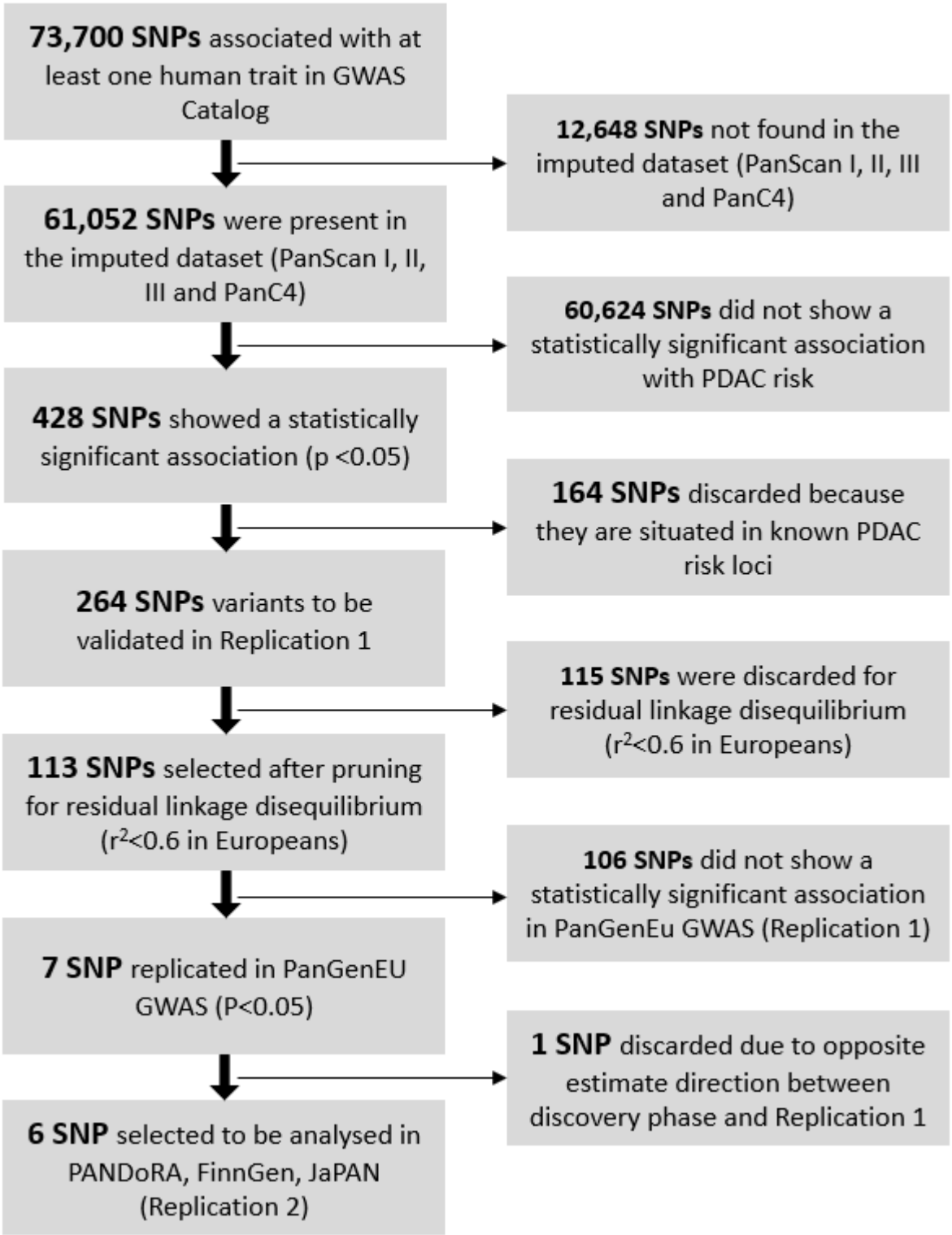
Downloaded from https://academic.oup.com/mutage/advance-article/doi/10.1093/mutage/geae012/7644568 by Szegedi Tudomanyegyetem / University of Szeged user on 22 April 2024

In conclusion, none of the SNPs showed a formally statistically significant association after correction for multiple testing. However, due to their pleiotropic nature and their connection with CRC, the two SNPs showing the best associations with PDAC risk merit further investigation through fine mapping and ad hoc functional studies.

Accepted Manuscript

Downloaded from https://academic.oup.com/mutage/advance-article/doi/10.1093/mutage/geae012/7644568 by Szegedi Tudomanyegyetem / University of Szeged user on 22 April 2024

Figure 1: Flowchart of SNP selection.



Downloaded from https://academic.oup.com/mutage/advance-article/doi/10.1093/mutage/geae012/7644568 by Szegedi Tudomanyegyetem / University of Szeged user on 22 April 2024

Table 1. Description of the study populations.

	PanScan I II III & PanC4	***PanGenEU***	***PANDoRA***	***JaPAN[1]***	***FinnGen[2]***	***Total***
Study phase	Discovery	Replication 1	Replication 2	Replication 2	Replication 2	
Number of subjects						
Cases	8,738	1,317	3,442	2,039	519	16,055
Controls	7,034	1,616	3,928	32,592	174,006	212,149
Total	15,772	2,933	7,370	34,631	174,525	235,231
Median age (25%-75% percentiles)						
Cases	65 (55-75)	66 (57-73)	66(58-73)	62.7| 66.3	-	-
Controls	65 (55-75)	65 (55-75)	59(50-66)	43.6| 56.3	-	-
Sex						
Female	46%	42%	47%	-	-	-
Male	54%	58%	53%	-	-	-

[1]The information about sex and age are reported as minimum | maximum value.

[2] We used the FinnGen documentation of R4 Release.

"-": Information not available in the original database.

Accepted Manuscript

Downloaded from https://academic.oup.com/mutage/advance-article/doi/10.1093/mutage/geae012/7644568 by Szegedi Tudomanyegyetem / University of Szeged user on 22 April

Table 2: SNPs associated with PDAC risk in discovery phase (PanScan and PanC4) that were found associated also in replication 1 (PanGenEU GWAS).

SNP	Locus	M/m	MAF in CEU	Results of replication 1		Associated traits in GWAS Catalog		
				OR (95%CI)	Pvalue	Traits	OR (95%CI)	Pvalue
rs6919397	6q22.32	T/G	0.409	0.80 (0.69-0.92)	0.002	Neuroticism	NR (NR)	$4x10^{-8}$
						Type 1 diabetes (rs9388489, r^2=0.99, D'=1.00) #	1.17 (1.10-1.24)	$4x10^{-13}$
						Type 2 diabetes (rs4897182, r^2=0.97, D'=0.99) #	1.05 (1.02-1.11)	$3x10^{-8}$
rs288762	7q36.3	C/T	0.384	1.17 (1.01-1.36)	0.034	Estimated glomerular filtration rate in non-diabetics	1.44 (1.11-2.30)	$2x10^{-11}$
rs2980752	8p23.1	C/A	0.294	0.84 (0.72-0.99)	0.034	Heel bone mineral density	1.03 (1.00-1.07)	$3x10^{-50}$
						Triglyceride levels (rs2980755, r^2=0.41, D'=0.91) #	0.98 (0.95-0.99)	$1x10^{-10}$
rs4948550	10q21.1	C/T	0.328	0.85 (0.73-1.00)	0.049	Cardiovascular disease	NR (NR)	$7x10^{-8}$
						Colorectal cancer (rs4948317, r^2=0.82, D'=0.97) #	1.10 (1.06-1.13)	$7x10^{-8}$
						Morning person (rs2893787, r^2=0.91, D'=0.99) #	NR (NR)	$2x10^{-8}$
rs35138700*	12q12	C/T	0.394	1.18 (1.02-1.37)	0.029	Morning person	1.05 (NR)	$4x10^{-43}$
						Type 2 diabetes (rs7315028, r^2=0.57, D'=0.96) #	1.13 (1.04-1.33)	$2x10^{-8}$

Accepted Manuscript

Downloaded from https://academic.oup.com/mutage/advance-article/doi/10.1093/mutage/geae012/7644568 by Szegedi Tudomanyegyetem / University of Szeged user on 22 April

rs7214227*	17q23.2	C/T	0.146	0.76 (0.63-0.93)	0.014	Estimated glomerular filtration rate	1.01 (1.00-1.01)	$2x10^{-36}$
						Renal function-related traits (rs11868441, r^2=0.61, D'=0.89) #	1.01 (1.00-1.01)	$2x10^{-9}$

*SNPs selected as proxy (r^2=1, D'≥ 0.94).

Traits reported in GWAS Catalog for variants in LD (r^2>0.4) with the SNPs analysed in replication 1.

M/m: major/minor allele; OR (95%CI): odds ratio and its 95% confidence interval; NR: data nor reported in GWAS Catalog, the Pvalue refers to the association with the reported trait.

Accepted Manuscript

Downloaded from https://academic.oup.com/mutage/advance-article/doi/10.1093/mutage/geae012/7644568 by Szegedi Tudomanyegyetem / University of Szeged user on 22 April

Table 3: Associations between selected SNPs and PDAC risk in the individual studies and in the meta-analysis.

SNP, locus, M/m		*PanScan/PanC 4*	*PanGenEU*	*PANDoRA*[1]	*PANDoRA*[2]	*JaPAN*	*FinnGen*	*Meta-analysis*[1]	*Meta-analysis*[2]
rs6919397	OR (95%CI)	0.92 (0.88-0.96)	0.80 (0.69-0.92)	098 (0.92-1.05)	0.98 (0.91-1.05)	-	1.06 (0.93-1.2)	0.94 (0.87-1.02)	0.95 (0.88-1.03)
6q22.32	P_{value}	1.31×10^{-4}	2.44×10^{-3}	6.17×10^{-1}	5.40×10^{-1}	-	3.84×10^{-1}	1.35×10^{-1}	2.17×10^{-1}
T/G	P_{value} Het.							1.18×10^{-2}	9.70×10^{-3}
rs288762	OR (95%CI)	1.10 (1.05-1.15)	1.17 (1.01-1.36)	1.02 (0.95-1.10)	1.02 (0.95-1.09)	1.02 (0.94-1.11)	1.09 (0.96-1.24)	1.07 (1.04-1.11)	1.08 (1.04-1.12)
7q36.3	P_{value}	1.09×10^{-4}	3.38×10^{-2}	5.04×10^{-1}	6.60×10^{-1}	6.13×10^{-1}	1.79×10-[1]	3.98×10^{-5}	3.03×10^{-5}
C/T	P_{value} Het.							2.52×10^{-1}	2.13×10^{-1}
rs2980752	OR (95%CI)	0.90 (0.85-0.95)	0.84 (0.72-0.99)	0.98 (0.91-1.05)	0.98 (0.91-1.06)	1.02 (0.95-1.10)	0.97 (0.85-1.10)	0.95 (0.89-1.01)	0.92 (0.89-0.96)
8p23.1	P_{value}	2.92×10^{-4}	3.38×10^{-2}	5.29×10^{-1}	5.40×10^{-1}	5.33×10^{-1}	6.28×10^{-1}	7.98×10^{-2}	1.38×10^{-4}
C/A	P_{value} Het.							2.55×10^{-2}	1.71×10^{-1}
rs4948550	OR (95%CI)	0.91 (0.87-0.96)	0.85 (0.73-1.00)	0.97 (0.89-1.04)	0.95 (0.88-1.03)	0.99 (0.91-1.07)	0.99 (0.86-1.14)	0.94 (0.91-0.97)	0.92 (0.89-0.96)
10q21.1	P_{value}	2.92×10^{-4}	4.99×10^{-2}	3.64×10^{-1}	2.40×10^{-1}	7.74×10^{-1}	8.84×10^{-1}	3.91×10^{-4}	6.52×10^{-5}
C/T	P_{value} Het.							2.58×10^{-1}	4.47×10^{-1}
*rs35138700**	OR (95%CI)	1.11 (1.04-1.18)	1.18 (1.02-1.37)	1.01 (0.94-1.09)	1.02 (0.93-1.12)	1.03 (0.93-1.14)	0.91 (0.93-1.20)	1.05 (0.97-1.12)	1.05 (0.96-1.16)
12q12	P_{value}	9.32×10^{-5}	2.86×10^{-2}	7.11×10^{-1}	6.40×10^{-1}	5.29×10^{-1}	1.48×10^{-1}	2.19×10^{-1}	2.92×10^{-1}
C/T	P_{value} Het.							1.24×10^{-2}	1.06×10^{-2}
*rs7214227**	OR (95%CI)	0.9 (0.85-0.96)	0.76 (0.63-0.93)	1.03 (0.94-1.14)	1.01 (0.92-1.11)	1.03 (0.93-1.14)	0.92 (0.77-1.10)	0.94 (0.86-1.03)	0.92 (0.88-0.97)
17q23.2	P_{value}	1.14×10^{-3}	6.68×10^{-3}	4.74×10^{-1}	8.60×10^{-1}	5.87×10^{-1}	3.72×10^{-1}	1.93×10^{-1}	8.89×10^{-4}
C/T	P_{value} Het.							1.06×10^{-2}	5.18×10^{-2}

[1] result with all subjects; [2] results excluding non-European subjects.

* SNPs selected as proxy (r^2=1, D'≥ 0.94).

m= minor allele; M= major allele; "P_{value} Het" = P_{value} of heterogeneity. When the studies were heterogeneous, we performed the meta-analysis with the random effect model, while when studies were not heterogenous, we used the fixed-effect model.

Accepted Manuscript

Downloaded from https://academic.oup.com/mutage/advance-article/doi/10.1093/mutage/geae012/7644568 by Szegedi Tudomanyegyetem / University of Szeged user on 22 April

Competing interests:

M.F. Bijlsma. has received research funding from Celgene and LeadPharma, and has acted as a consultant to Servier.

Funding:

this work was partially supported by intramural funds of University of Pisa and DKFZ; by Fondazione Tizzi (www.fondazionetizzi.it) and by Fondazione Arpa (www.fondazionearpa.it). Tavano F. was supported by Italian Ministry of Health grants to Fondazione "Casa Sollievo della Sofferenza" IRCCS Hospital, San Giovanni Rotondo (FG), Italy and by the "5x1000" voluntary contribution**;** by Ministery of Health, Czech Republic, NV19-09-00088 (to Lovecek M.), NV19-03-00097 (to B. Mohelníková-Duchoňová), NV19-08-00113 (to P. Soucek), NU21-07-00247 (to P. Vodicka). G. Capurso received founding from AIRC (IG 2021 ID 26201). The PanGenEU team was supported by Fondo de Investigaciones Sanitarias (FIS), Instituto de Salud Carlos III, Spain (#PI061614, #PI11/01542, #PI0902102, #PI12/01635, #PI12/00815, #PI15/01573, #PI18/01347), Red Temática de Investigación Cooperativa en Cáncer, Spain (#RD12/0036/0034, #RD12/0036/0050, #RD12/0036/0073), EU-6FP Integrated Project (#018771-MOLDIAG-PACA), EU-FP7-HEALTH (#259737-CANCERALIA, #256974-EPC-TM-Net), Pancreatic Cancer Collective (PCC): Lustgarten Foundation & Stand-Up to Cancer (SU2C #6179).

Acknowledgements:

We would like to acknowledge the contribution of the late Dr. Bas Bueno-de-Mesquita.

This article/publication is based upon work from COST Action Transpan supported by COST (European Cooperation in Science and Technology).

Members of the PanGenEu consortium are listed at **supplementary Material: Annex S1**. We would like to thank Prof. Key T.J. (Epidemiology Unit Nuffield Department of Population Health, University of Oxford, Oxford, United Kingdom) and Prof. Wareham N.J. (MRC Epidemiology Unit, Institute of Metabolic Science, University of Cambridge, Cambridge, United Kingdom) and Prof. Vermeulen R.C.H. (University of Utrecht) for the EPIC genotyping data used in the PANDoRA replication. The EPIC-Norfolk study (DOI 10.22025/2019.10.105.00004) has received funding from the Medical Research Council (MR/N003284/1 and MC-UU_12015/1) and Cancer Research UK (C864/A14136). We are grateful to all the participants who have been part of the project and to the many members of the study teams at the University of Cambridge who have enabled this research. The authors would also thank all the participants of the PanScan GWAS Study and Pancreatic Cancer Case Control Association Study, and dbGaP for providing cancer genotyping dataset. For the Humanitas group, the biological material of the individuals included in the study was provided by the Center for Biological Resources at Humanitas Research Hospital.

Accepted Manuscript

Downloaded from https://academic.oup.com/mutage/advance-article/doi/10.1093/mutage/geae012/7644568 by Szegedi Tudomanyegyetem / University of Szeged user on 22 April 2024

Ethics statement: each participating study obtained approval from the responsible institutional review board (IRB) and IRB certification permitting data sharing in accordance with the NIH Policy for sharing of Data Obtained in NIH-Supported or NIH-Conducted Genome Wide Association Studies. The PANDoRA study protocol was approved by the Ethics Commission of the Medical Faculty of the University of Heidelberg. In accordance with the Declaration of Helsinki, written informed consent was obtained from each participant. IRB ethical approval and written informed consent was obtained by all participating centres contributing to PanGenEU Study. The FinnGen study was approved by the ethical Review Board of the Hospital District of Helsinki and Uusimaa. FinnGen participants provided written, informed consent. For JaPAN, written informed consent was obtained from all study participants, and the study protocol was approved by the Ethical Review Board of Aichi Medical University, the Institutional Ethics Committee of Aichi Cancer Center, the Human Genome and Gene Analysis Research Ethics Committee of Nagoya University, and the ethics committees of all participating hospitals.

Accepted Manuscript

Downloaded from https://academic.oup.com/mutage/advance-article/doi/10.1093/mutage/geae012/7644568 by Szegedi Tudomanyegyetem / University of Szeged user on 22 April 2024

Data availability: the PanScan and PanC4 genotyping data are available from the database of Genotypes and Phenotypes (dbGaP, study accession numbers phs000206.v5.p3 and phs000648.v1.p1). JaPAN data are available from the JaPAN consortium website (www.aichi-med-u.ac.jp/JaPAN). FinnGen summary statistics are available from the FinnGen study website (www.finngen.fi). The PANDoRA primary data for this work will be made available to researchers who submit a reasonable request to the corresponding author, conditional to approval by the PANDoRA Steering Committee and Ethics Commission of the Medical Faculty of the University of Heidelberg. Data will be stripped from all information allowing identification of study participants. PanGenEU GWAS summary statistics are available in GWAS catalog repository.

Author contribution: D.C. conceived and designed the study. M.Gi., M.R., performed the lab work, M.Ge., M.R. analysed the data. All the authors contributed with the interpretation of the data. M.Gi., M.Ge., D.C. wrote the first draft of the manuscript and all authors contributed to the writing and approve of the final version of the manuscript.

Accepted Manuscript

Downloaded from https://academic.oup.com/mutage/advance-article/doi/10.1093/mutage/geae012/7644568 by Szegedi Tudomanyegyetem / University of Szeged user on 22 April 2024

References

1. Sung, H., Ferlay, J., Siegel, R.L., Laversanne, M., Soerjomataram, I., Jemal, A., and Bray, F. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA: a cancer journal for clinicians*, 2021 **71**,209–249. doi:10.3322/caac.21660.

2. Maisonneuve, C., I, G., P, V., T, W., P, A., T, B., M, B., and DB, C. Bicaudal C, a novel regulator of Dvl signaling abutting RNA-processing bodies, controls cilia orientation and leftward flow. *Development (Cambridge, England)*, 2009 **136**,3019–3030. doi:10.1242/DEV.038174.

3. Lu, Y., Gentiluomo, M., Lorenzo-Bermejo, J., Morelli, L., Obazee, O., Campa, D., and Canzian, F. Mendelian randomisation study of the effects of known and putative risk factors on pancreatic cancer. *Journal of medical genetics*, 2020 **57**,820–828. doi:10.1136/jmedgenet-2019-106200.

4. Rizzato, C., Campa, D., Pezzilli, R., et al. ABO blood groups and pancreatic cancer risk and survival: results from the PANcreatic Disease ReseArch (PANDoRA) consortium. *Oncology reports*, 2013 **29**,1637–1644. doi:10.3892/or.2013.2285.

5. Peduzzi, G., Felici, A., Pellungrini, R., et al. Analysis of exposome and genetic variability suggests stress as a major contributor for development of pancreatic ductal adenocarcinoma. *Digestive and Liver Disease*, 2023. Accessed 17 January 2024.

6. Chang, J., Tian, J., Zhu, Y., et al. Exome-wide analysis identifies three low-frequency missense variants associated with pancreatic cancer risk in Chinese populations. *Nature communications*, 2018 **9**,3688. doi:10.1038/s41467-018-06136-x.

7. Nakatochi, M., Lin, Y., Ito, H., et al. Prediction model for pancreatic cancer risk in the general Japanese population. *PloS one*, 2018 **13**,e0203386. doi:10.1371/journal.pone.0203386.

8. Zhang, M., Wang, Z., Obazee, O., et al. Three new pancreatic cancer susceptibility signals identified on chromosomes 1q32.1, 5p15.33 and 8q24.21. *Oncotarget*, 2016 **7**,66328–66343. doi:10.18632/oncotarget.11041.

9. Wolpin, B.M., Rizzato, C., Kraft, P., et al. Genome-wide association study identifies multiple susceptibility loci for pancreatic cancer. *Nature Genetics*, 2014 **46**,994–1000. doi:10.1038/ng.3052.

10. Low, S.-K., Kuchiba, A., Zembutsu, H., et al. Genome-wide association study of pancreatic cancer in Japanese population. *PloS one*, 2010 **5**,e11824. doi:10.1371/journal.pone.0011824.

Accepted Manuscript

Downloaded from https://academic.oup.com/mutage/advance-article/doi/10.1093/mutage/geae012/7644568 by Szegedi Tudomanyegyetem / University of Szeged user on 22 April 2024

11. Petersen, G.M., Amundadottir, L., Fuchs, C.S., et al. A genome-wide association study identifies pancreatic cancer susceptibility loci on chromosomes 13q22.1, 1q32.1 and 5p15.33. *Nature genetics*, 2010 **42**,224–228. doi:10.1038/ng.522.

12. Amundadottir, L., Kraft, P., Stolzenberg-Solomon, R.Z., et al. Genome-wide association study identifies variants in the ABO locus associated with susceptibility to pancreatic cancer. *Nature Genetics*, 2009 **41**,986–990. doi:10.1038/ng.429.

13. Wu, C., Miao, X., Huang, L., et al. Genome-wide association study identifies five loci associated with susceptibility to pancreatic cancer in Chinese populations. *Nature Genetics*, 2011 **44**,62–66. doi:10.1038/ng.1020.

14. Childs, E.J., Mocci, E., Campa, D., et al. Common variation at 2p13.3, 3q29, 7p13 and 17q25.1 associated with susceptibility to pancreatic cancer. *Nature Genetics*, 2015 **47**,911–916. doi:10.1038/ng.3341.

15. Gentiluomo, M., Canzian, F., Nicolini, A., Gemignani, F., Landi, S., and Campa, D. Germline genetic variability in pancreatic cancer risk and prognosis. *Seminars in cancer biology*, 2020. doi:10.1016/j.semcancer.2020.08.003.

16. Corradi, C., Gentiluomo, M., Gajdán, L., et al. Genome-wide scan of long noncoding RNA single nucleotide polymorphisms and pancreatic cancer susceptibility. *International journal of cancer*, 2021 **148**,2779–2788. doi:10.1002/ijc.33475.

17. Campa, D., Pastore, M., Gentiluomo, M., et al. Functional single nucleotide polymorphisms within the cyclin-dependent kinase inhibitor 2A/2B region affect pancreatic cancer risk. *Oncotarget*, 2016 **7**,57011–57020. doi:10.18632/oncotarget.10935.

18. Campa, D., Rizzato, C., Stolzenberg-Solomon, R., et al. TERT gene harbors multiple variants associated with pancreatic cancer susceptibility. *International Journal of Cancer*, 2015 **137**,2175–2183. doi:10.1002/ijc.29590.

19. Campa, D., Matarazzi, M., Greenhalf, W., et al. Genetic determinants of telomere length and risk of pancreatic cancer: A PANDoRA study. *International Journal of Cancer*, 2019 **144**,1275–1283. doi:10.1002/ijc.31928.

20. Gentiluomo, M., Peduzzi, G., Lu, Y., Campa, D., and Canzian, F. Genetic polymorphisms in inflammatory genes and pancreatic cancer risk: a two-phase study on more than 14 000 individuals. *Mutagenesis*, 2019 **34**,395–401. doi:10.1093/mutage/gez040.

21. Peduzzi, G., Gentiluomo, M., Tavano, F., et al. Genetic Polymorphisms Involved in Mitochondrial Metabolism and Pancreatic Cancer Risk. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*, 2021 **30**,2342–2345. doi:10.1158/1055-9965.EPI-21-0353.

Accepted Manuscript

Downloaded from https://academic.oup.com/mutage/advance-article/doi/10.1093/mutage/geae012/7644568 by Szegedi Tudomanyegyetem / University of Szeged user on 22 April 2024

22. Campa, D., Gentiluomo, M., Obazee, O., et al. Genome-wide association study identifies an early onset pancreatic cancer risk locus. *International Journal of Cancer*, 2020,ijc.33004. doi:10.1002/ijc.33004.

23. López de Maturana, E., Rodríguez, J.A., Alonso, L., et al. A multilayered post-GWAS assessment on genetic susceptibility to pancreatic cancer. *Genome medicine*, 2021 **13**,15. doi:10.1186/s13073-020-00816-4.

24. Campa, D., Gentiluomo, M., Stein, A., et al. The PANcreatic Disease ReseArch (PANDoRA) consortium: Ten years' experience of association studies to understand the genetic architecture of pancreatic cancer. *Critical Reviews in Oncology/Hematology*, 2023 **186**.

25. Lindström, S., Finucane, H., Bulik-Sullivan, B., et al. Quantifying the Genetic Correlation between Multiple Cancer Types. *Cancer Epidemiology and Prevention Biomarkers*, 2017 **26**,1427–1435. doi:10.1158/1055-9965.EPI-17-0211.

26. Rashkin, S.R., Graff, R.E., Kachuri, L., et al. Pan-cancer study detects genetic risk variants and shared genetic basis in two large cohorts. *Nature Communications 2020 11:1*, 2020 **11**,1–14. doi:10.1038/s41467-020-18246-6.

27. Loos, R.J.F. 15 years of genome-wide association studies and no signs of slowing down. *Nature Communications 2020 11:1*, 2020 **11**,1–3. doi:10.1038/s41467-020-19653-5.

28. Sivakumaran, S., Agakov, F., Theodoratou, E., et al. Abundant pleiotropy in human complex diseases and traits. *American Journal of Human Genetics*, 2011 **89**,607–618. doi:10.1016/j.ajhg.2011.10.004.

29. Gratten, J. and Visscher, P.M. Genetic pleiotropy in complex traits and diseases: implications for genomic medicine. *Genome Medicine 2016 8:1*, 2016 **8**,1–3. doi:10.1186/S13073-016-0332-X.

30. Solovieff, N., Cotsapas, C., Lee, P.H., Purcell, S.M., and Smoller, J.W. Pleiotropy in complex traits: challenges and strategies. *Nature Reviews Genetics 2013 14:7*, 2013 **14**,483–495. doi:10.1038/nrg3461.

31. Fehringer, G., Kraft, P., Pharoah, P.D., et al. Cross-Cancer Genome-Wide Analysis of Lung, Ovary, Breast, Prostate, and Colorectal Cancer Reveals Novel Pleiotropic Associations. *Cancer research*, 2016 **76**,5103–5114. doi:10.1158/0008-5472.CAN-15-2980.

32. Chung, C.C. and Chanock, S.J. Current status of genome-wide association studies in cancer. *Human Genetics*, 2011 **130**,59–78. doi:10.1007/s00439-011-1030-9.

33. Crocco, P., Barale, R., Rose, G., et al. Population-specific association of genes for telomere-associated proteins with longevity in an Italian population. *Biogerontology*, 2015 **16**,353–364. doi:10.1007/s10522-015-9551-6.

Accepted Manuscript

Downloaded from https://academic.oup.com/mutage/advance-article/doi/10.1093/mutage/geae012/7644568 by Szegedi Tudomanyegyetem / University of Szeged user on 22 April 2024

34. Bonàs-Guarch, S., Guindo-Martínez, M., Miguel-Escalada, I., et al. Re-analysis of public genetic data reveals a rare X-chromosomal variant associated with type 2 diabetes. *Nature communications*, 2018 **9**,321. doi:10.1038/s41467-017-02380-9.

35. Gentiluomo, M., Luddi, A., Cingolani, A., et al. Telomere Length and Male Fertility. *International journal of molecular sciences*, 2021 **22**. doi:10.3390/ijms22083959.

36. Chong, M., Mohammadi-Shemirani, P., Perrot, N., et al. GWAS and ExWAS of blood mitochondrial DNA copy number identifies 71 loci and highlights a potential causal role in dementia. *eLife*, 2022 **11**. doi:10.7554/eLife.70382.

37. Richardson, T.G., Sanderson, E., Palmer, T.M., Ala-Korpela, M., Ference, B.A., Davey Smith, G., and Holmes, M. V. Evaluating the relationship between circulating lipoprotein lipids and apolipoproteins with risk of coronary heart disease: A multivariable Mendelian randomisation analysis. *PLoS medicine*, 2020 **17**,e1003062. doi:10.1371/journal.pmed.1003062.

38. Pierce, B.L. and Ahsan, H. Genome-wide "Pleiotropy Scan" Identifies HNF1A Region as a Novel Pancreatic Cancer Susceptibility Locus. *Cancer research*, 2011 **71**,4352. doi:10.1158/0008-5472.CAN-11-0124.

39. Molina-Montes, E., Coscia, C., Gómez-Rubio, P., et al. Deciphering the complex interplay between pancreatic cancer, diabetes mellitus subtypes and obesity/BMI through causal inference and mediation analyses. *Gut*, 2021 **70**,319–329. doi:10.1136/gutjnl-2019-319990.

40. Nagai, A., Hirata, M., Kamatani, Y., et al. Overview of the BioBank Japan Project: Study design and profile. *J Epidemiol*, 2017 **27**,S2–S8. doi:10.1016/j.je.2016.12.005.

41. Lin, Y., Nakatochi, M., Hosono, Y., et al. Genome-wide association meta-analysis identifies GP2 gene risk variants for pancreatic cancer. *Nature communications*, 2020 **11**,3175. doi:10.1038/s41467-020-16711-w.

42. FinnGen. FinnGen Documentation of R5 release. 2021.

43. 1000 Genomes Project Consortium., Abecasis, G.R., Altshuler, D., et al. A map of human genome variation from population-scale sequencing. *Nature*, 2010 **467**,1061–1073. doi:10.1038/nature09534.

44. Campa, D., Rizzato, C., Capurso, G., et al. Genetic susceptibility to pancreatic cancer and its functional characterisation: the PANcreatic Disease ReseArch (PANDoRA) consortium. *Digestive and liver disease : official journal of the Italian Society of Gastroenterology and the Italian Association for the Study of the Liver*, 2013 **45**,95–99. doi:10.1016/j.dld.2012.09.014.

Accepted Manuscript

Downloaded from https://academic.oup.com/mutage/advance-article/doi/10.1093/mutage/geae012/7644568 by Szegedi Tudomanyegyetem / University of Szeged user on 22 April 2024

45. Zhao, R., Peng, C., Song, C., et al. BICC1 as a novel prognostic biomarker in gastric cancer correlating with immune infiltrates. *International Immunopharmacology*, 2020 **87**,106828. doi:10.1016/J.INTIMP.2020.106828.

46. Jones, S.E., Lane, J.M., Wood, A.R., et al. Genome-wide association analyses of chronotype in 697,828 individuals provides insights into circadian rhythms. *Nature communications*, 2019 **10**,343. doi:10.1038/s41467-018-08259-7.

47. Kichaev, G., Bhatia, G., Loh, P.-R., et al. Leveraging Polygenic Functional Enrichment to Improve GWAS Power. *American journal of human genetics*, 2019 **104**,65–75. doi:10.1016/j.ajhg.2018.11.008.

48. Zhang, B., Jia, W.-H., Matsuda, K., et al. Large-scale genetic study in East Asians identifies six new loci associated with colorectal cancer risk. *Nature genetics*, 2014 **46**,533–542. doi:10.1038/ng.2985.

49. Zhang, B.-L., He, N., Huang, Y.-B., Song, F.-J., and Chen, K.-X. ABO blood groups and risk of cancer: a systematic review and meta-analysis. *Asian Pacific journal of cancer prevention : APJCP*, 2014 **15**,4643–4650. doi:10.7314/apjcp.2014.15.11.4643.

50. Schmit, S.L., Edlund, C.K., Schumacher, F.R., et al. Novel Common Genetic Susceptibility Loci for Colorectal Cancer. *Journal of the National Cancer Institute*, 2019 **111**,146–157. doi:10.1093/jnci/djy099.

51. Klein, A.P., Wolpin, B.M., Risch, H.A., et al. Genome-wide meta-analysis identifies five new susceptibility loci for pancreatic cancer. *Nature Communications*, 2018 **9**,1–11. doi:10.1038/s41467-018-02942-5.

52. Law, P.J., Timofeeva, M., Fernandez-Rozadilla, C., et al. Association analyses identify 31 new risk loci for colorectal cancer susceptibility. *Nature communications*, 2019 **10**,2154. doi:10.1038/s41467-019-09775-w.

53. Jeng, K.S., Chang, C.F., and Lin, S.S. Sonic hedgehog signaling in organogenesis, tumors, and tumor microenvironments. *International Journal of Molecular Sciences*, 2020 **21**. doi:10.3390/ijms21030758.

54. Li, X., Ma, Q., Duan, W., Liu, H., Xu, H., and Wu, E. Paracrine sonic hedgehog signaling derived from tumor epithelial cells: A key regulator in the pancreatic tumor microenvironment. *Critical Reviews in Eukaryotic Gene Expression*, 2012 **22**,97–108. doi:10.1615/CritRevEukarGeneExpr.v22.i2.20.

55. Zhao, X., Cao, D., Ren, Z., et al. Dipeptidyl peptidase like 6 promoter methylation is a potential prognostic biomarker for pancreatic ductal adenocarcinoma. *Bioscience Reports*, 2020 **40**. doi:10.1042/BSR20200214.

Accepted Manuscript

Downloaded from https://academic.oup.com/mutage/advance-article/doi/10.1093/mutage/geae012/7644568 by Szegedi Tudomanyegyetem / University of Szeged user on 22 April 2024