

Functional single nucleotide polymorphisms within the cyclin-dependent kinase inhibitor 2A/2B region affect pancreatic cancer risk

Daniele Campa^{1,2}, Manuela Pastore^{1,2}, Manuel Gentiluomo^{1,2}, Renata Talar-Wojnarowska³, Juozas Kupcinskas⁴, Ewa Malecka-Panas³, John P. Neoptolemos⁵, Willem Niesen⁶, Pavel Vodicka^{7,8}, Gianfranco Delle Fave⁹, H. Bas Bueno-de-Mesquita^{10,11,12}, Maria Gazouli¹³, Paola Pacetti¹⁴, Milena Di Leo¹⁵, Hidemi Ito¹⁶, Harald Klüter¹⁷, Pavel Soucek^{18,19}, Vincenzo Corbo²⁰, Kenji Yamao²¹, Satoyo Hosono¹⁶, Rudolf Kaaks²², Yogesh Vashist²³, Domenica Gioffreda²⁴, Oliver Strobel⁶, Yasuhiro Shimizu²⁵, Frederike Dijk²⁶, Angelo Andriulli²⁴, Audrius Ivanauskas⁴, Peter Bugert¹⁷, Francesca Tavano²⁴, Ludmila Vodickova^{8,27}, Carlo Federico Zambon²⁸, Martin Lovecek²⁹, Stefano Landi¹, Timothy J. Key³⁰, Ugo Boggi³¹, Raffaele Pezzilli³², Krzysztof Jamrozak³³, Beatrice Mohelnikova-Duchonova^{18,34}, Andrea Mambrini¹⁴, Franco Bambi³⁵, Olivier Busch³⁶, Valerio Pazienza²⁴, Roberto Valente⁹, George E. Theodoropoulos³⁷, Thilo Hackert⁶, Gabriele Capurso⁹, Giulia Martina Cavestro¹⁵, Claudio Pasquali³⁸, Daniela Basso³⁹, Cosimo Sperti³⁸, Keitaro Matsuo⁴⁰, Markus Büchler⁶, Kay-Tee Khaw⁴¹, Jakob Izbicki²³, Eithne Costello⁵, Verena Katzke²², Christoph Michalski⁶, Anna Stepień⁴², Cosmeri Rizzato⁴³, Federico Canzian²

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

²Genomic Epidemiology Group, German Cancer Research Center (DKFZ), Heidelberg, Germany

³Department of Digestive Tract Diseases, Medical University of Lodz, Poland

⁴Department of Gastroenterology, Lithuanian University of Health Sciences, Kaunas, Lithuania

⁵Institute for Health Research Liverpool Pancreas Biomedical Research Unit, University of Liverpool, Liverpool, United Kingdom

⁶Department of General, Visceral and Transplantation Surgery, Heidelberg University Hospital, Heidelberg, Germany

⁷Institute of Experimental Medicine, Czech Academy of Science, Prague Czech Republic

⁸Institute of Biology and Medical Genetics, 1st Medical Faculty, Charles University, Prague, Czech Republic

⁹Digestive and Liver Disease Unit, S. Andrea Hospital, 'Sapienza' University of Rome, Rome, Italy

¹⁰Department for Determinants of Chronic Diseases (DCD), National Institute for Public Health and the Environment (RIVM), Bilthoven, The Netherlands

¹¹Department of Epidemiology and Biostatistics, The School of Public Health, Imperial College London, London, United Kingdom

¹²Department of Social & Preventive Medicine, Faculty of Medicine, University of Malaya, Kuala Lumpur, Malaysia

¹³Department of Basic Medical Sciences, Laboratory of Biology, Medical School, National and Kapodistrian University of Athens, Athens Greece

¹⁴Oncological Department Massa Carrara Azienda USL Toscana Nord Ovest, Carrara, Italy

¹⁵Gastroenterology and Gastrointestinal Endoscopy Unit, Vita-Salute San Raffaele University, IRCCS San Raffaele Scientific Institute Milan, Italy

¹⁶Division Epidemiology and Prevention, Aichi Cancer Center Research Institute, Nagoya, Japan

¹⁷Institute of Transfusion Medicine and Immunology, German Red Cross Blood Service Baden-Württemberg – Hessen gGmbH, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany

¹⁸Laboratory of Toxicogenomics, National Institute of Public Health, Prague, Czech Republic

¹⁹Laboratory of Pharmacogenomics, Biomedical Center, Faculty of Medicine in Pilsen, Charles University in Prague, Pilsen, Czech Republic

²⁰ARC-Net Research Centre, and Department of Diagnostics and Public Health University and Hospital Trust of Verona, Verona, Italy

- ²¹Department of Gastroenterology, Aichi Cancer Center Hospital
- ²²Division of Cancer Epidemiology, German Cancer Research Center (DKFZ), Heidelberg, Germany
- ²³Department of General, Visceral and Thoracic Surgery, University Medical Center Hamburg-Eppendorf, Hamburg, Germany
- ²⁴Division of Gastroenterology and Research Laboratory, IRCCS Scientific Institute and Regional General Hospital "Casa Sollievo della Sofferenza", San Giovanni Rotondo, Italy
- ²⁵Department of Gastroenterological Surgery, Aichi Cancer Center Hospital, Nagoya, Japan
- ²⁶Department of Pathology, Academic Medical Centre, Amsterdam, the Netherlands
- ²⁷Biomedical Center, Faculty of Medicine in Pilsen, Charles University in Prague, Czech Republic
- ²⁸Department of Medicine - DIMED, University of Padova, Italy
- ²⁹Department of Surgery I, Faculty of Medicine and Dentistry, Palacky University Olomouc and University Hospital Olomouc, Czech Republic
- ³⁰Epidemiology Unit Nuffield Department of Population Health University of Oxford, Oxford, UK
- ³¹Division of General and Transplant Surgery, Pisa University Hospital Pisa, Italy
- ³²Pancreas Unit, Department of Digestive System, Dant'Orsola-Malpighi Hospital, Bologna, Italy
- ³³Department of Hematology, Institute of Hematology and Transfusion Medicine, Warsaw, Poland
- ³⁴Department of Oncology, Faculty of Medicine and Dentistry, Palacky University Olomouc and University Hospital Olomouc, Czech Republic
- ³⁵Blood Transfusion Service, Azienda Ospedaliero Universitaria Meyer, Florence, Italy
- ³⁶Department of Surgery, Academic Medical Centre, Amsterdam, the Netherlands
- ³⁷Colorectal Unit, First Department of Propaedeutic Surgery, Athens Medical School, National and Kapodistrian University of Athens, Athens, Greece
- ³⁸Department of Surgery, Oncology and Gastroenterology-DiSCOG, University of Padova, Italy
- ³⁹Department of Laboratory Medicine, University-Hospital of Padova, Italy
- ⁴⁰Div. Molecular Medicine, Aichi Cancer Center Research Institute, Nagoya, Japan
- ⁴¹University of Cambridge School of Clinical Medicine Clinical Gerontology Unit Box 251 Addenbrooke's Hospital, Cambridge, UK
- ⁴²Laboratory of Clinical, Transplant Immunology and Genetics, Copernicus Memorial Hospital, Lodz, Poland
- ⁴³Department of Translational Research and New Technologies in Medicine and Surgery, University of Pisa, Pisa, Italy

Correspondence to: Federico Canzian, **email:** f.canzian@dkfz.de

Keywords: pancreatic cancer, CDKN2A, single nucleotide polymorphisms, miRSNP, association study

Received: May 09, 2016

Accepted: July 13, 2016

Published: July 29, 2016

ABSTRACT

The *CDKN2A* (*p16*) gene plays a key role in pancreatic cancer etiology. It is one of the most commonly somatically mutated genes in pancreatic cancer, rare germline mutations have been found to be associated with increased risk of developing familial pancreatic cancer and *CDKN2A* promoter hyper-methylation has been suggested to play a critical role both in pancreatic cancer onset and prognosis. In addition several unrelated SNPs in the 9p21.3 region, that includes the *CDNK2A*, *CDNK2B* and the *CDNK2B-AS1* genes, are associated with the development of cancer in various organs. However, association between the common genetic variability in this region and pancreatic cancer risk is not clearly understood. We sought to fill this gap in a case-control study genotyping 13 single nucleotide polymorphisms (SNPs) in 2,857 pancreatic ductal adenocarcinoma (PDAC) patients and 6,111 controls in the context of the Pancreatic Disease Research (PANDORA) consortium. We found that the A allele of the rs3217992 SNP was associated with an increased pancreatic cancer risk ($OR_{het} = 1.14$, 95% CI 1.01-1.27, $p = 0.026$, $OR_{hom} = 1.30$, 95% CI 1.12-1.51, $p = 0.00049$). This pleiotropic variant is reported to be a mir-SNP that, by changing the binding site of one or more miRNAs, could influence the normal cell cycle progression and in turn increase PDAC risk. In conclusion, we observed a novel association in a pleiotropic

region that has been found to be of key relevance in the susceptibility to various types of cancer and diabetes suggesting that the *CDKN2A/B* locus could represent a genetic link between diabetes and pancreatic cancer risk.

INTRODUCTION

The majority of pancreatic cancer patients die within a year of diagnosis [1]. The poor prognosis is caused by various factors, including the lack of appropriate markers for early detection, the aggressiveness of the disease and the dearth of effective treatment possibilities available. One of the best strategies to reduce the mortality of the disease is to improve early diagnosis, and it is, therefore, important to identify individuals at high risk in the population and subject them to enhanced surveillance.

Only a few epidemiologic risk factors have been established for pancreatic cancer, including cigarette smoking, heavy alcohol intake, diabetes mellitus, obesity, chronic pancreatitis and family history of pancreatic cancer [2, 3]. Even less is known about the genetic contribution to the disease, since only a rather small number of susceptibility *loci* have been identified through genome-wide association studies (GWAS) [4–9] and confirmed by follow-up studies [10]. Moreover it has been shown that a small proportion of pancreatic tumors arises as a result of high penetrance germline mutations in genes such as *BRC1A1*, *BRC1A2*, *p16/CDKN2A*, *STK11/LKB*, *ATM*, *PALB2*, and DNA mismatch repair genes, usually in the context of familial cancer syndromes [2, 3, 11–15]. However, the very low frequency of those mutations cannot explain the bulk of genetic susceptibility to pancreatic cancer.

There are compelling epidemiologic and molecular evidences pointing to a key role for the *CDKN2A* gene in pancreatic cancer etiology. *CDKN2A* is one of the most commonly somatically mutated genes in pancreatic cancer [16], rare germline mutations have been found to be associated with increased risk of developing familial pancreatic cancer [15, 17], and also *CDKN2A* promoter hyper-methylation has been suggested to play a critical role both in pancreatic cancer onset and prognosis [18]. Additionally, the 9p21.3 region, that includes in addition to *CDKN2A* also *CDKN2B* and *CDKN2B-AS1*, is pleiotropic and several polymorphic variants in the region spanning the three genes are susceptibility markers for several cancer types [19–24]. In addition Li and colleagues performed an association study and meta-analysis across multiple cancers corroborating the pleiotropic role of the locus [25]. Several polymorphic variants in the region are also associated with type two diabetes mellitus (T2DM), which is a predisposing factor for pancreatic cancer, suggesting a possible role of the 9p21.3 region as a genetic link between the two diseases [26–28]. The pleiotropic role of the variants of this region is probably due to the central importance of the genes situated in it in cell cycle regulation. For example *CDKN2A* codes, by alternative splicing, for the two oncosuppressors p16INK4a and

p14ARF [29, 30]. Despite all the hints pointing towards an association between common genetic variability in the *CDKN2A/B* gene region and pancreatic cancer risk no one has attempted to directly relate them so far. We sought to fill this gap in a case-control study genotyping 13 single nucleotide polymorphisms (SNPs) in the context of the PANcreatic Disease ReseArch (PANDoRA) consortium.

RESULTS

Data filtering and quality control

The characteristics of the population enrolled in the study are shown in Table 1. All analyzed SNPs were in HWE in controls ($P > 0.0038$) with the exception of rs3731246 in the controls from Mannheim (Germany) and Southern Italy and rs2811710 in the controls from Mannheim. The populations not respecting the HWE were not included in the statistical analyses for the relevant SNPs. Starting from a population of 9,796 subjects, 828 subjects with a call rate $< 75\%$ were removed after genotyping, leaving 8,968 for further analysis. The average SNP call rate was 96% with a minimum of 79.95% (rs3731246) and a maximum of 99.19% (rs3218009). The analysis of the random duplicate samples showed a concordance rate of 99.68%. For the Japanese cases it was not possible to correctly genotype rs3731246 and therefore this SNP was not used in the risk analysis for the Japanese population. Supplementary Table S1 shows the call rate and the HWE equilibrium for each SNP.

SNP main effects

We analyzed separately Caucasian and Japanese individuals. For the individuals of Caucasian origin we observed that 6 SNPs showed a statistically significant association ($p < 0.05$) with increased or decreased PDAC risk. The strongest association with an increased risk of PDAC was observed with the A allele of rs3217992 SNP ($OR_{het} = 1.14$, 95% CI 1.01-1.27, $p = 0.026$, $OR_{hom} = 1.30$, 95% CI 1.12-1.51, $p = 0.00049$, unadjusted p-trend = 0.0002, adjusted p-trend = 0.32). Other, less significant associations were observed with rs3731249, rs2811708, rs3731211, and rs1063192. The frequencies and distributions of the genotypes (for the Caucasian group) and the OR for the association of each polymorphism with PDAC are described in Table 2. In the Japanese population the G allele of rs2811708 was associated with a decreased risk of PDAC ($OR_{het} = 0.58$, 95% CI 0.36-0.95, $p = 0.029$) and the A allele of rs1063192 was associated with a decreased risk of PDAC ($OR_{het} = 0.49$, 95% CI 0.30-0.80, $p = 0.005$, p-trend = 0.03). The frequencies and distribution of the Japanese genotypes

Table 1: Characteristics of the studied population

	Cases	Controls	Total
Median age (25%-75% percentiles)	65 (57-72)	57 (48-64)	
Gender			
Males	1637	3371	5008
Females	1182	2677	3859
Unknown	38	63	101
Geographic origin			
Germany	1066	2282	3348
Czech Republic	251	518	769
Greece	80	175	255
Italy Center	480	549	1029
Italy North	361	595	956
Italy South	114	500	614
Lithuania	57	190	247
Poland	87	335	422
Netherlands	117	102	219
UK	99	176	275
Japan	145	689	834
Total	2857	6111	8968

and the OR for the association of each polymorphism with PDAC are described in Table 3.

Possible functional effects

We used several bioinformatic tools to predict possible functional relevance for the SNPs showing the most significant associations. Using Genevar, we observed that the C allele of rs3217992 was associated with increased expression of the interferon alpha 4 (*IFNA4*) gene ($P=0.047$). This association, however, was not below the threshold suggested by Genevar for significance ($P<10^{-3}$). GTEx did not show any significant eQTL for any of the SNPs associated with pancreatic cancer risk. RegulomeDB showed a score of 5, indicating the possible presence of a transcription factor binding motif or a DNase sensitivity peak for rs2811708, rs3217992 and rs1063192 and a score of 4 suggesting the presence of a transcription factor binding motif and a DNase sensitivity peak for rs3731211 and rs3731249. HaploReg did not suggest any relevant signals for the significant SNPs.

DISCUSSION

Several unrelated SNPs in the 9p21.3 region, that includes the *CDNK2A*, *CDNK2B* and the *CDNK2B-AS1*

genes, are associated with the development of cancer in various organs [19-21, 23, 24, 31-33] and with T2D [26-28]. In this study we have performed an in depth analysis of the region and found several promising associations between the common genetic variability and the disease onset. However, considering the correction for multiple testing only rs3217992 showed a statistically significant association with increased risk of developing PDAC.

This finding is interesting for two reasons. The first is that rs3217992 shows a plethora of associations with several human traits such as primary open-angle glaucoma [34], aggressive periodontitis, chronic periodontitis [35] and myocardial infarction [36]. These phenotypes are very different from each other, but all the studies, including the one presented here, have in common the fact that is always the A allele to be associated with the increased risk of the disease. This observation strongly suggests a pleiotropic role for rs3217992 and also highlights that the polymorphisms alters the function of the protein in a way that affects the related phenotypes. In recent years several pleiotropic SNPs have been identified to be associated with multiple human phenotypes or multiple cancer types. These pleiotropic stretches of DNA that are densely packed with risk alleles have been defined Nexus regions [37] and intensely studied in relation to cancer risk. PDAC

Table 2: Association between the selected SNPs and PDAC risk in the Caucasian population

SNP	Alleles ^a (M/m)	Cases ^b			Controls ^b			MM vs Mm		p	MM vs mm		p	p-trend adj	p-trend unadj	SNP Annotation ^d
		MM	Mm	mm	MM	Mm	mm	OR ^c	95% CI ^c		OR	95% CI				
rs3731257	C/T	1,362	1,082	211	2,567	2,007	354	0.99	0.89-1.10	0.83	1.07	0.88-1.31	0.476	0.419	0.30	5' flanking
rs11515	C/G	1,875	700	50	3,568	1,297	118	1.00	0.89-1.13	0.92	0.70	0.48-1.01	0.058	0.70	0.81	3' UTR
rs2518719	G/A	2,057	542	46	3,779	1,124	100	0.90	0.79-1.01	0.09	0.97	0.66-1.42	0.87	0.15	0.03	Intronic
rs3731249	C/T	2,496	159	5	4,140	341	9	0.75	0.60-0.92	0.007	0.64	0.19-2.07	0.45	0.14	0.01	Missense
rs3731246	C/G	1,814	450	37	1,842	515	30	0.85	0.73-0.99	0.04	1.10	0.65-1.85	0.72	0.39	0.32	Intronic
rs2811708	T/G	1,382	894	145	2,770	1,985	359	0.90	0.80-1.00	0.057	0.76	0.61-0.95	0.015	0.03	0.009	Intronic
rs3731239	C/T	1,066	1,193	313	2,237	2,365	673	1.06	0.96-1.19	0.25	0.99	0.84-1.18	0.92	0.47	0.82	Intronic
rs3731211	A/T	1,387	997	190	2,220	1,701	380	0.96	0.86-1.08	0.51	0.75	0.62-0.92	0.007	0.29	0.02	Intronic
rs2811710	T/C	1,087	1,207	333	1,413	1,636	504	1.00	0.88-1.12	0.97	0.88	0.74-1.05	0.17	0.26	0.08	Intronic
rs3218009	C/G	2,188	437	23	4,071	888	66	0.96	0.84-1.10	0.54	0.75	0.45-1.24	0.26	0.52	0.04	Intronic
rs3217992	G/A	821	1,332	504	1,807	2,543	868	1.14	1.01-1.27	0.026	1.30	1.12-1.51	0.0005	0.32	0.0002	3'UTR
rs1063192	G/A	1,011	1,219	341	1,756	2,291	722	0.89	0.80-0.99	0.047	0.82	0.70-0.97	0.019	0.55	0.009	3'UTR
rs3217986	C/A	2,243	392	17	4,511	810	28	0.97	0.84-1.12	0.71	1.48	0.78-2.82	0.23	0.14	0.89	3'UTR

^a M = major allele (i.e. more common in controls); m = minor allele (less common in controls).

^b Numbers may not add up to 100% due to genotyping failure, DNA depletion or covariate missing values.

^c Odds ratio (95% confidence interval).

^d as shown in Haploreg

All analyses were adjusted for age at diagnosis/age at recruitment, gender and country of origin.

Table 3: Association between the selected SNPs and PDAC risk in the Japanese population

SNP	Alleles ^a (M/m)	Cases ^b			Controls ^b			MM vs Mm		p	MM vs mm		p	p-trend adj	p-trend unadj
		MM	Mm	mm	MM	Mm	mm	OR	95% CI		OR	95% CI			
rs3731257	C/T	24	68	49	128	336	204	1.10	0.66-1.83	0.71	1.30	0.76-2.23	0.33	0.41	0.32
rs11515	C/G	131	7	0	643	17	0	1.97	0.80-4.85	0.14	-	-	-	1.00	0.12
rs2518719	G/A	145	0	0	667	0	0	-	-	-	-	-	-	-	-
rs3731249	C/T	136	0	0	683	0	0	-	-	-	-	-	-	-	-
rs2811708	T/G	101	23	5	452	176	18	0.58	0.36-0.95	0.029	1.23	0.45-3.41	0.68	0.59	0.15
rs3731239	C/T	109	27	2	507	164	14	0.76	0.48-1.21	0.25	0.66	0.15-2.96	0.59	0.24	0.22
rs3731211	A/T	97	38	9	429	218	20	0.77	0.51-1.16	0.21	2.01	0.89-4.55	0.09	0.85	0.97
rs2811710	T/C	82	52	9	376	272	36	0.87	0.59-1.27	0.48	1.15	0.53-2.47	0.73	0.41	0.81
rs3218009	C/G	143	0	0	687	0	0	-	-	-	-	-	-	-	-
rs3217992	G/A	28	44	50	128	289	218	0.70	0.41-1.17	0.17	1.05	0.63-1.75	0.85	0.44	0.59
rs1063192	G/A	99	22	9	432	198	32	0.49	0.30-0.80	0.005	1.23	0.57-2.66	0.60	0.03	0.11
rs3217986	C/A	130	14	0	622	66	0	1.00	0.55-1.85	0.98	-	-	-	0.06	0.96

^a M = major allele (i.e. more common in controls); m = minor allele (less common in controls).

^b Numbers may not add up to 100% due to genotyping failure, DNA depletion or covariate missing values.

^c Odds ratio (95% confidence interval).

All analyses were adjusted for age at diagnosis/age at recruitment, gender and country of origin.

has at least another such region in the *TERT-CLPTMIL* locus as widely demonstrated by several authors [6, 9, 38].

The second reason is that in addition to the strong statistical association and the putative pleiotropic effect, rs3217992 SNP lies in a miRNA target region

(miR-138-2-3p). In a very recent study Ghanbari and collaborators showed, using cardiac cell lines, that the rs3217992 SNP might have an effect on the miRNA-mRNA interactions [39]. Specifically, the G allele increased the miRNA-dependent regulation of *CDKN2B*.

Given the role of CDKN2B in the cell cycle this result seems counterintuitive since the A allele is associated with increased risk in our study and in others as mentioned before. The 9p21.3 locus has, however, a very complex regulation, and it is possible that the binding between the G allele and miR-138-2-3p might be specific for the cardiac tissue. Moreover the 9p21.3 locus hosts several genes that are key cell cycle regulators and therefore the fact that rs3217992 changes the binding site of one or more miRNAs could influence the normal cell cycle progression and in turn increase PDAC risk. There are growing evidences that mir-SNPs could be involved in various pathologies including cancer and diabetes given their ability to affect gene regulation. However, the functional significance of this SNP needs thorough scrutiny in the pancreatic tissue to address its possible role in PDAC susceptibility.

Another finding of potential significance is the association between the A allele of rs1063192 with decreased PDAC risk. This SNP also lies in the *CDKN2B* 3'UTR, is a putative mir-SNP and shows a weak linkage disequilibrium (LD) with rs3217992 ($r^2=0.474$ in CEU and $r^2=0.170$ CHBJPT, 1000 Genomes Project). Li and colleagues found this SNP to be associated with esophageal squamous cell carcinoma [25], while in a recent report [40] the A allele of rs1063192 was found to be associated with a decreased risk of developing gestational diabetes mellitus (GDM) suggesting a direct genetic link between diabetes and pancreatic cancer though the genetic variants analyzed in both studies.

The strongest association we observed in the Japanese population was between the A allele of rs1063192 and a decreased PDAC risk. This finding gives rise to two considerations: the first is that even though the association does not reach the threshold of significance considering Bonferroni's correction, it is unlikely to be a false finding giving the fact that it was found in two different populations. The second is that since in the Caucasian and in the Japanese the leading SNP is not the same it is possible that the real causal variant is yet untyped and mildly in LD with rs1063192 and rs3217992.

We also observed an association close to statistical significance (taking into account multiple testing) between the missense SNP rs3731249 and PDAC susceptibility. This variation is predicted by PolyPhen to be possibly damaging with a score of 0.487 and has been widely investigated in childhood acute lymphocytic leukemia [23, 24, 41], once again highlighting the pleiotropic nature of this genomic region. Through dbGaP we could perform a GWAS look-up for the results of our candidate SNPs. We found results only for Caucasians (PanScan2, [6]) and only for three of the significant SNPs rs2811708, rs3217992 and rs1063192. All the SNPs showed ORs that were in the same direction with our results but none showed a statistically significant association, the best was $p=0.08$ for rs1063192. In 2 out of 3 populations (PANDoRA Japanese

and PanScan) tested the best association was observed for rs1063192 and in the last population (PANDoRA Caucasians) for rs3217992. The two SNPs are in weak LD. This suggests the hypothesis that another variant (possibly a low frequency/rare one) may be associated with both, that was not typed in either population and that has an LD pattern with rs1063192 and rs3217992 that varies in the different populations. The difference in the results between PANDoRA and PanScan may be explained by the fact that in PANDoRA the strongest statistical association was observed for the rare homozygous carriers compared with the common allele carriers suggesting a recessive model of inheritance, while in the PanScan data only the allelic model is shown. Additionally, the results from PanScan arise partly from a prospective cohort while ours are from a case-control study. Although the biological explanation of this phenomenon is difficult to understand, discrepancies between the two study designs have been observed several times in different neoplastic diseases [4, 21, 42, 43].

One of the major strengths of this study is its size, since with a total of 8,968 subjects this is one the largest genetic analysis of pancreatic cancer risk published to-date. Additionally, our selection of SNPs provides an extensive coverage of genetic variability in the regions of interest. In addition we could analyze simultaneously two ethnic groups, which helps to generalize the findings. A possible limitation of the study is that patients and controls in PANDoRA were recruited in various centers across Europe and therefore there might be some population-based differences. Moreover we did not include rare variants and therefore we cannot exclude to have missed associations due to alleles with a $MAF < 0.05$.

In conclusion we observed several novel associations in a pleiotropic region that has been found to be of key relevance in the susceptibility to various types of cancer and diabetes, confirming a key role for *CDKN2A/B* in pancreatic cancer and suggesting a possible involvement of the common genetic variability at this locus in PDAC risk.

MATERIALS AND METHODS

Study population

In this study 2,857 pancreatic ductal adenocarcinoma (PDAC) patients and 6,111 controls were collected in nine countries (Germany, Czech Republic, Greece, Italy, Lithuania, Poland, Netherlands, UK and Japan) belonging to the PANDoRA consortium that has been described in detail elsewhere [44]. Briefly, cases were retrospectively collected and are defined by a confirmed diagnosis of PDAC. Controls were recruited in the same hospitals, or at least geographical region from where the cases were recruited. British and Dutch controls (N=176 and 102, respectively) were selected from healthy volunteers recruited from the general population in the

European Prospective Investigation on Cancer (EPIC), an ongoing prospective cohort being carried out in ten European countries [45]. The German controls used were partly blood donors from the blood transfusion center in Mannheim and partly healthy volunteers selected among EPIC subjects collected in Heidelberg. Patients provided written informed consent and the study was approved by Ethical Review Board of the University of Heidelberg (Medizinische Fakultät Heidelberg) the study was performed in accordance with Declaration of Helsinki.

SNP selection

To survey the common genetic variability in the *CDKN2A/B* locus we selected tagging SNPs using the Tagger tool of the Haploview software using the Caucasian population in the HapMap web site (International HapMap Project, version 28; <http://www.hapmap.org>) as reference (<http://www.broad.mit.edu/mpg/haploview/>; <http://www.broad.mit.edu/mpg/tagger/>). We considered a region centered on the *CDKN2A* gene region and we added 5k bp at 5' and 5k at 3' of the gene resulting in a window of around 40k base pair. We used a pairwise tagging method with a minimum r^2 of 0.8 and a minor frequency allele of 0.05 to select tagging SNPs inside the region. The first tagging SNP at the 5' end, rs3731257, is situated at 21956221 (Hg18) while the last tagging SNP at the 3' end, rs3217986, is situated at 21995330 (Hg18). In addition we added two functional SNPs (rs1063192 and rs3217992) that are putative miR-SNPs, i.e. predicted to alter the binding of one or more microRNAs to their target. The final selection resulted in 13 SNPs in the 9p21.3 region. We checked if the tagging selection used was valid also for the Asian ethnicities and we obtained that the tagging set for Chinese and Japanese is a subset of the one used for Caucasians.

Sample preparation and genotyping

For each sample DNA was extracted from whole blood or from paraffin-embedded pancreatic tissues of patients and controls using the AllPrep Isolation Kit (Qiagen, Hilden, Germany) or the Qiagen-mini kit (Qiagen, Hilden, Germany), according to the manufacturer's protocol. Genotyping was performed using TaqMan (ABI, Applied Biosystems, Foster City, CA, USA) and KASPar (KBioscience, Hoddesdon, UK) technologies. The order of DNA samples was randomized on plates in order to ensure that similar numbers of cases and controls were analyzed in each batch. Detection was performed using an ABI PRISM 7900 HT or Viia7 sequence detection system with SDS 2.2 or Viia7 software (Applied Biosystems, Foster City, CA, USA). Genotyping for British and Dutch controls was performed in the context of a genome-wide association study using the Human 660W-Quad BeadChip array according to manufacturer's instructions (Illumina,

San Diego, CA, USA). For quality control, duplicates of 10% of the samples were interspersed throughout the plates. In addition, we discarded all the samples that had a call rate < 75%.

Statistical analysis

The observed genotype frequencies of all SNPs in the control subjects were tested for deviation from Hardy-Weinberg equilibrium (HWE) using Pearson's chi-square test. The association between the genotypes of all polymorphisms and PDAC risk was estimated using an unconditional logistic regression computing odds ratios (OR), 95% confidence intervals (95% CIs) and p values. The more common allele among the controls was assigned as the reference category and the co-dominant model inheritance model was assessed. All analyses were adjusted for age, gender and geographic origin (among the European countries). European and Japanese individuals were kept separate in the analyses. All analyses were adjusted for multiple testing using the Bonferroni correction whereby a significance threshold of 0.0038 (0.05/13) was set.

Bioinformatic analysis

To assess the possible functional relevance for the SNPs showing the most significant associations with risk of developing PDAC several bioinformatic tools were used. RegulomeDB (<http://regulome.stanford.edu/>) [46] and HaploReg v2B [47] were used to identify the regulatory potential of the region nearby each SNP. Genevar (<http://www.sanger.ac.uk/resources/software/genevar/>) [48] and GTEEx (<http://www.gtexportal.org/home/>) [49] were used to identify potential associations between the SNP and expression levels of nearby genes (eQTL).

ACKNOWLEDGMENTS

The authors would like to express their thanks to the PancoBank (EPZ/Heidelberg, Germany; Prof. M.W. Büchler, Dr. N.A. Giese, E. Soyka, M. Stauch, M. Meinhardt) from where the cases from Heidelberg were obtained.

CONFLICTS OF INTEREST

John P. Neoptolemos has the following conflicts of interest. Payment for Lectures: Amgen, Mylan Research Grants: Taiho Pharma (Japan); KAEL GemVax (Korea); AstraZeneca; Clovis Oncology and Ventana; Pharma Nord. Consultancy: Boehringer Ingelheim Pharma GmbH & Co. KG; Novartis Pharma AG; KAEL GemVax; Astellas. Educational Travel Grants: NUCANA.

The other authors do not have any conflict of interest to declare.

THIS PROJECT WAS PARTIALLY SUPPORTED BY:

National Institutes of Health Research Liverpool Pancreas Biomedical Research Unit, UK.

Professor John Neoptolemos, the Owen and Ellen Evans Chair of Surgery, FMedSci, is a National Institutes of Health Research Senior Investigator.

BMBF grants (01GS08114,01ZX1305C, 01KT1506), Heidelberger Stiftung Chirurgie and Biomaterial Bank Heidelberg (Prof. P. Schirmacher; BMBF grant 01EY1101).

Ministry of Health of the Czech Republic (grant no. 16-28375A to B.M-D.) and the Czech Science Foundation (grant no. P301/12/1734 to P.S.)

Italian Cancer Genome Project (FIRB RBAP10AHJB), Associazione Italiana Ricerca Cancro (AIRC grant n. 12182).

REFERENCES

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2015. *CA Cancer J Clin.* 2015; 65: 5-29. doi: 10.3322/caac.21254.
2. Hidalgo M. Pancreatic cancer. *N Engl J Med.* 2010; 362: 1605-17. doi: 10.1056/NEJMra0901557.
3. Wolfgang CL, Herman JM, Laheru DA, Klein AP, Erdek MA, Fishman EK, Hruban RH. Recent progress in pancreatic cancer. *CA Cancer J Clin.* 2013; 63: 318-48. doi: 10.3322/caac.21190.
4. Amundadottir L, Kraft P, Stolzenberg-Solomon RZ, Fuchs CS, Petersen GM, Arslan AA, Bueno-de-Mesquita HB, Gross M, Helzlsouer K, Jacobs EJ, LaCroix A, Zheng W, Albanes D, et al. Genome-wide association study identifies variants in the ABO locus associated with susceptibility to pancreatic cancer. *Nat Genet.* 2009; 41: 986-90. doi: ng.429 [pii] 10.1038/ng.429.
5. Low SK, Kuchiba A, Zembutsu H, Saito A, Takahashi A, Kubo M, Daigo Y, Kamatani N, Chiku S, Totsuka H, Ohnami S, Hirose H, Shimada K, et al. Genome-wide association study of pancreatic cancer in Japanese population. *PLoS One.* 2010; 5: e11824.
6. Petersen GM, Amundadottir L, Fuchs CS, Kraft P, Stolzenberg-Solomon RZ, Jacobs KB, Arslan AA, Bueno-de-Mesquita HB, Gallinger S, Gross M, Helzlsouer K, Holly EA, Jacobs EJ, 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. doi: ng.522 [pii] 10.1038/ng.522.
7. Wu C, Miao X, Huang L, Che X, Jiang G, Yu D, Yang X, Cao G, Hu Z, Zhou Y, Zuo C, Wang C, Zhang X, et al. Genome-wide association study identifies five loci associated with susceptibility to pancreatic cancer in Chinese populations. *Nat Genet.* 2011; 44: 62-6.
8. Childs EJ, Mocci E, Campa D, Bracci PM, Gallinger S, Goggins M, Li D, Neale RE, Olson SH, Scelo G, Amundadottir LT, Bamlet WR, Bijlsma MF, et al. Common variation at 2p13.3, 3q29, 7p13 and 17q25.1 associated with susceptibility to pancreatic cancer. *Nat Genet.* 2015; 47: 911-6. doi: 10.1038/ng.3341.
9. Wolpin BM, Rizzato C, Kraft P, Kooperberg C, Petersen GM, Wang Z, Arslan AA, Beane-Freeman L, Bracci PM, Buring J, Canzian F, Duell EJ, Gallinger S, et al. Genome-wide association study identifies multiple susceptibility loci for pancreatic cancer. *Nat Genet.* 2014; 46: 994-1000. doi: 10.1038/ng.3052.
10. Rizzato C, Campa D, Giese N, Werner J, Rachakonda PS, Kumar R, Schanne M, Greenhalf W, Costello E, Khaw KT, Key TJ, Siddiq A, Lorenzo-Bermejo J, et al. Pancreatic cancer susceptibility loci and their role in survival. *PLoS One.* 2011; 6: e27921.
11. Goggins M, Schutte M, Lu J, Moskaluk CA, Weinstein CL, Petersen GM, Yeo CJ, Jackson CE, Lynch HT, Hruban RH, Kern SE. Germline BRCA2 gene mutations in patients with apparently sporadic pancreatic carcinomas. *Cancer Res.* 1996; 56: 5360-4.
12. Hassan MM, Bondy ML, Wolff RA, Abbruzzese JL, Vauthey JN, Pisters PW, Evans DB, Khan R, Chou TH, Lenzi R, Jiao L, Li D. Risk factors for pancreatic cancer: case-control study. *Am J Gastroenterol.* 2007; 102: 2696-707.
13. Kastrinos F, Mukherjee B, Tayob N, Wang F, Sparr J, Raymond VM, Bandipalliam P, Stoffel EM, Gruber SB, Syngal S. Risk of pancreatic cancer in families with Lynch syndrome. *Jama.* 2009; 302: 1790-5. doi: 10.1001/jama.2009.1529.
14. Tischkowitz MD, Sabbaghian N, Hamel N, Borgida A, Rosner C, Taherian N, Srivastava A, Holter S, Rothenmund H, Ghadirian P, Foulkes WD, Gallinger S. Analysis of the gene coding for the BRCA2-interacting protein PALB2 in familial and sporadic pancreatic cancer. *Gastroenterology.* 2009; 137: 1183-6.
15. Vasen HF, Gruis NA, Frants RR, van Der Velden PA, Hille ET, Bergman W. Risk of developing pancreatic cancer in families with familial atypical multiple mole melanoma associated with a specific 19 deletion of p16 (p16-Leiden). *Int J Cancer.* 2000; 87: 809-11.
16. Biankin AV, Waddell N, Kassahn KS, Gingras MC, Muthuswamy LB, Johns AL, Miller DK, Wilson PJ, Patch AM, Wu J, Chang DK, Cowley MJ, Gardiner BB, et al. Pancreatic cancer genomes reveal aberrations in axon guidance pathway genes. *Nature.* 2012; 491: 399-405. doi: 10.1038/nature11547.
17. Ghiorzo P. Genetic predisposition to pancreatic cancer. *World J Gastroenterol.* 2014; 20: 10778-89. doi: 10.3748/wjg.v20.i31.10778.
18. Tang B, Li Y, Qi G, Yuan S, Wang Z, Yu S, Li B, He S. Clinicopathological Significance of CDKN2A Promoter Hypermethylation Frequency with Pancreatic Cancer. *Sci Rep.* 2015; 5: 13563. doi: 10.1038/srep13563.

19. Al Olama AA, Kote-Jarai Z, Berndt SI, Conti DV, Schumacher F, Han Y, Benlloch S, Hazelett DJ, Wang Z, Saunders E, Leongamornlert D. A meta-analysis of 87,040 individuals identifies 23 new susceptibility loci for prostate cancer. *Nat Genet.* 2014; 46: 1103-9. doi: 10.1038/ng.3094.
20. Bishop DT, Demenais F, Iles MM, Harland M, Taylor JC, Corda E, Randerson-Moor J, Aitken JF, Avril MF, Azizi E, Bakker B, Bianchi-Scarra G, Bressac-de Paillerets B, et al. Genome-wide association study identifies three loci associated with melanoma risk. *Nat Genet.* 2009; 41: 920-5. doi: 10.1038/ng.411.
21. Campa D, Barrdahl M, Gaudet MM, Black A, Chanock SJ, Diver WR, Gapstur SM, Haiman C, Hankinson S, Hazra A, Henderson B, Hoover RN, Hunter DJ, et al. Genetic risk variants associated with in situ breast cancer. *Breast Cancer Res.* 2015; 17: 82. doi: 10.1186/s13058-015-0596-x.
22. Stacey SN, Sulem P, Masson G, Gudjonsson SA, Thorleifsson G, Jakobsdottir M, Sigurdsson A, Gudbjartsson DF, Sigurgeirsson B, Benediktsdottir KR, Thorisdottir K, Ragnarsson R, Scherer D, et al. New common variants affecting susceptibility to basal cell carcinoma. *Nat Genet.* 2009; 41: 909-14. doi: 10.1038/ng.412.
23. Vijayakrishnan J, Henrion M, Moorman AV, Fiege B, Kumar R, Inacio da Silva Filho M, Holroyd A, Koehler R, Thomsen H, Irving JA, Allan JM, Lightfoot T, Roman E, et al. The 9p21.3 risk of childhood acute lymphoblastic leukaemia is explained by a rare high-impact variant in CDKN2A. *Sci Rep.* 2015; 5: 15065. doi: 10.1038/srep15065.
24. Walsh KM, de Smith AJ, Hansen HM, Smirnov IV, Gonseth S, Endicott AA, Xiao J, Rice T, Fu CH, McCoy LS, Lachance DH, Eckel-Passow JE, Wiencke JK, et al. A Heritable Missense Polymorphism in CDKN2A Confers Strong Risk of Childhood Acute Lymphoblastic Leukemia and Is Preferentially Selected during Clonal Evolution. *Cancer Res.* 2015; 75: 4884-94. doi: 10.1158/0008-5472.can-15-1105.
25. Li WQ, Pfeiffer RM, Hyland PL, Shi J, Gu F, Wang Z, Bhattacharjee S, Luo J, Xiong X, Yeager M, Deng X, Hu N, Taylor PR, et al. Genetic polymorphisms in the 9p21 region associated with risk of multiple cancers. *Carcinogenesis.* 2014; 35: 2698-705. doi: 10.1093/carcin/bgu203.
26. Mahajan A, Go MJ, Zhang W, Below JE, Gaulton KJ, Ferreira T, Horikoshi M, Johnson AD, Ng MC, Prokopenko I, Saleheen D, Wang X, Zeggini E, et al. Genome-wide trans-ancestry meta-analysis provides insight into the genetic architecture of type 2 diabetes susceptibility. *Nat Genet.* 2014; 46: 234-44. doi: 10.1038/ng.2897.
27. Scott LJ, Mohlke KL, Bonnycastle LL, Willer CJ, Li Y, Duren WL, Erdos MR, Stringham HM, Chines PS, Jackson AU, Prokunina-Olsson L, Ding CJ, Swift AJ, et al. A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. *Science.* 2007; 316: 1341-5. doi: 10.1126/science.1142382.
28. Zeggini E, Scott LJ, Saxena R, Voight BF, Marchini JL, Hu T, de Bakker PI, Abecasis GR, Almgren P, Andersen G, Ardlie K, Bostrom KB, Bergman RN, et al. Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. *Nat Genet.* 2008; 40: 638-45. doi: 10.1038/ng.120.
29. Serrano M, Hannon GJ, Beach D. A new regulatory motif in cell-cycle control causing specific inhibition of cyclin D/CDK4. *Nature.* 1993; 366: 704-7. doi: 10.1038/366704a0.
30. Zhang Y, Xiong Y, Yarbrough WG. ARF promotes MDM2 degradation and stabilizes p53: ARF-INK4a locus deletion impairs both the Rb and p53 tumor suppression pathways. *Cell.* 1998; 92: 725-34.
31. Michailidou K, Hall P, Gonzalez-Neira A, Ghoussaini M, Dennis J, Milne RL, Schmidt MK, Chang-Claude J, Bojesen SE, Bolla MK, Wang Q, Dicks E, Lee A, et al. Large-scale genotyping identifies 41 new loci associated with breast cancer risk. *Nat Genet.* 2013; 45: 353-61. doi: 10.1038/ng.2563.
32. Stacey SN, Helgason H, Gudjonsson SA, Thorleifsson G, Zink F, Sigurdsson A, Kehr B, Gudmundsson J, Sulem P, Sigurgeirsson B, Benediktsdottir KR, Thorisdottir K, Ragnarsson R, et al. New basal cell carcinoma susceptibility loci. *Nat Commun.* 2015; 6: 6825. doi: 10.1038/ncomms7825.
33. Stacey SN, Sulem P, Gudbjartsson DF, Jonasdottir A, Thorleifsson G, Gudjonsson SA, Masson G, Gudmundsson J, Sigurgeirsson B, Benediktsdottir KR, Thorisdottir K, Ragnarsson R, Fuentelsaz V, et al. Germline sequence variants in TGM3 and RGS22 confer risk of basal cell carcinoma. *Hum Mol Genet.* 2014; 23: 3045-53. doi: 10.1093/hmg/ddt671.
34. Pasquale LR, Loomis SJ, Kang JH, Yaspan BL, Abdrabou W, Budenz DL, Chen TC, Delbono E, Friedman DS, Gaasterland D, Gaasterland T, Grosskreutz CL, Lee RK, et al. CDKN2B-AS1 genotype-glaucoma feature correlations in primary open-angle glaucoma patients from the United States. *Am J Ophthalmol.* 2013; 155: 342-53. doi: 10.1016/j.ajo.2012.07.023.
35. Schaefer AS, Richter GM, Dommisch H, Reinartz M, Nothnagel M, Noack B, Laine ML, Folwaczny M, Groessner-Schreiber B, Loos BG, Jepsen S, Schreiber S. CDKN2BAS is associated with periodontitis in different European populations and is activated by bacterial infection. *J Med Genet.* 2011; 48: 38-47. doi: 10.1136/jmg.2010.078998.
36. Yang XC, Zhang Q, Chen ML, Li Q, Yang ZS, Li L, Cao FF, Chen XD, Liu WJ, Jin L, Wang XF. MTAP and CDKN2B genes are associated with myocardial infarction in Chinese Hans. *Clin Biochem.* 2009; 42: 1071-5. doi: 10.1016/j.clinbiochem.2009.02.021.
37. Chung CC, Chanock SJ. Current status of genome-wide association studies in cancer. *Hum Genet.* 2011.
38. Campa D, Rizzato C, Stolzenberg-Solomon R, Pacetti P, Vodicka P, Cleary SP, Capurso G, Bueno-de-Mesquita HB, Werner J, Gazouli M, Butterbach K, Ivanauskas A, Giese N, et al. TERT gene harbors multiple variants associated with pancreatic cancer susceptibility. *Int J Cancer.* 2015; 137: 2175-83. doi: 10.1002/ijc.29590.
39. Ghanbari M, Franco OH, de Looper HW, Hofman A, Erkeland SJ, Dehghan A. Genetic Variations in

- MicroRNA-Binding Sites Affect MicroRNA-Mediated Regulation of Several Genes Associated With Cardio-metabolic Phenotypes. *Circ Cardiovasc Genet.* 2015; 8: 473-86. doi: 10.1161/circgenetics.114.000968.
40. Wang X, Li W, Ma L, Gao J, Liu J, Ping F, Nie M. Association study of the miRNA-binding site polymorphisms of CDKN2A/B genes with gestational diabetes mellitus susceptibility. *Acta Diabetol.* 2015; 52: 951-8. doi: 10.1007/s00592-015-0768-2.
 41. Xu H, Zhang H, Yang W, Yadav R, Morrison AC, Qian M, Devidas M, Liu Y, Perez-Andreu V, Zhao X, Gastier-Foster JM, Lupo PJ, Neale G, et al. Inherited coding variants at the CDKN2A locus influence susceptibility to acute lymphoblastic leukaemia in children. *Nat Commun.* 2015; 6: 7553. doi: 10.1038/ncomms8553.
 42. Barrdahl M, Canzian F, Joshi AD, Travis RC, Chang-Claude J, Auer PL, Gapstur SM, Gaudet M, Diver WR, Henderson BE, Haiman CA, Schumacher FR, Le Marchand L, et al. Post-GWAS gene-environment interplay in breast cancer: results from the Breast and Prostate Cancer Cohort Consortium and a meta-analysis on 79,000 women. *Hum Mol Genet.* 2014; 23: 5260-70. doi: 10.1093/hmg/ddu223.
 43. Campa D, Kaaks R, Le Marchand L, Haiman CA, Travis RC, Berg CD, Buring JE, Chanock SJ, Diver WR, Dostal L, Fournier A, Hankinson SE, Henderson BE, et al. Interactions between genetic variants and breast cancer risk factors in the breast and prostate cancer cohort consortium. *J Natl Cancer Inst.* 2011; 103: 1252-63. doi: 10.1093/jnci/djr265.
 44. Campa D, Rizzato C, Capurso G, Giese N, Funel N, Greenhalf W, Soucek P, Gazouli M, Pezzilli R, Pasquali C, Talar-Wojnarowska R, Cantore M, Andriulli A, et al. Genetic susceptibility to pancreatic cancer and its functional characterisation: The PANcreatic Disease ReseArch (PANDoRA) consortium. *Dig Liver Dis.* 2012. doi: <http://dx.doi.org/10.1016/j.dld.2012.09.014>.
 45. Riboli E, Hunt KJ, Slimani N, Ferrari P, Norat T, Fahey M, Charrondiere UR, Hemon B, Casagrande C, Vignat J, Overvad K, Tjonneland A, Clavel-Chapelon F, et al. European Prospective Investigation into Cancer and Nutrition (EPIC): study populations and data collection. *Public Health Nutr.* 2002; 5: 1113-24. doi: 10.1079/phn2002394.
 46. Boyle AP, Hong EL, Hariharan M, Cheng Y, Schaub MA, Kasowski M, Karczewski KJ, Park J, Hitz BC, Weng S, Cherry JM, Snyder M. Annotation of functional variation in personal genomes using RegulomeDB. *Genome Res.* 2012; 22: 1790-7. doi: 10.1101/gr.137323.112.
 47. Ward LD, Kellis M. HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucleic Acids Res.* 2012; 40: D930-4. doi: 10.1093/nar/gkr917.
 48. Grundberg E, Small KS, Hedman AK, Nica AC, Buil A, Keildson S, Bell JT, Yang TP, Meduri E, Barrett A, Nisbett J, Sekowska M, Wilk A, et al. Mapping cis- and trans-regulatory effects across multiple tissues in twins. *Nat Genet.* 2012; 44: 1084-9. doi: 10.1038/ng.2394.
 49. Consortium. G. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science.* 2015; 348: 648-60. doi: 10.1126/science.1262110.