

Genetic determinants of telomere length and risk of pancreatic cancer: A PANDoRA study

Daniele Campa ¹, Martina Matarazzi^{1,2}, William Greenhalf³, Maarten Bijlsma ⁴, Kai-Uwe Saum⁵, Claudio Pasquali⁶, Hanneke van Laarhoven⁴, Andrea Szentesi^{7,8}, Francesca Federici⁹, Pavel Vodicka^{10,11,12}, Niccola Funel¹³, Raffaele Pezzilli¹⁴, H. Bas Bueno-de-Mesquita^{15,16,17,18}, Ludmila Vodickova^{10,11,12}, Daniela Basso¹⁹, Ofure Obazee², Thilo Hackert²⁰, Pavel Soucek ¹², Katarina Cuk⁵, Jörg Kaiser²⁰, Cosimo Sperti²², Martin Lovecek²¹, Gabriele Capurso^{23,24}, Beatrice Mohelnikova-Duchonova²⁵, Kay-Tee Khaw²⁶, Anna-Katharina König²⁰, Juozas Kupcinskas²⁷, Rudolf Kaaks ²⁸, Franco Bambi²⁹, Livia Archibugi^{23,24}, Andrea Mambrini⁹, Giulia Martina Cavestro³⁰, Stefano Landi¹, Péter Hegyi^{7,8,31}, Jakob R. Izbicki³², Domenica Gioffreda³³, Carlo Federico Zamboni²², Francesca Tavano ³³, Renata Talar-Wojnarowska³⁴, Krzysztof Jamrozak³⁵, Timothy J. Key³⁶, Gianfranco Delle Fave²³, Oliver Strobel²⁰, Laimas Jonaitis²⁷, Angelo Andriulli³³, Rita T. Lawlor³⁷, Felice Pirozzi³⁸, Verena Katzke²⁸, Chiara Valsuani⁹, Yogesh K. Vashist³², Hermann Brenner^{5,39,40} and Federico Canzian ²

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

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

³Institute for Health Research Liverpool Pancreas Biomedical Research Unit, University of Liverpool, Liverpool, United Kingdom

⁴Medical Oncology, Academic Medical Centre, Amsterdam, The Netherlands

⁵Division of Clinical Epidemiology and Aging Research, German Cancer Research Center (DKFZ), Heidelberg, Germany

⁶Pancreatic and Digestive Endocrine Surgery - Department of Surgery, Oncology and Gastroenterology (DiSCOG), University of Padova, Padova, Italy

⁷Institute for Translational Medicine, University of Pécs, Pécs, Hungary

⁸First Department of Medicine, University of Szeged, Szeged, Hungary

⁹Oncological Department, Azienda USL Toscana Nord Ovest, Oncological Unit of Massa Carrara, Carrara, Italy

¹⁰Department of Molecular Biology of Cancer, Institute of Experimental Medicine, Academy of Science of Czech Republic, Prague, Czech Republic

¹¹Institute of Biology and Medical Genetics, 1st Medical Faculty, Charles University, Prague, Czech Republic

¹²Biomedical Center, Faculty of Medicine in Pilsen, Charles University, Pilsen, Czech Republic

¹³Department of Surgery, Unit of Experimental Surgical Pathology, University of Pisa, Pisa, Italy

¹⁴Pancreas Unit, Department of Digestive System, Sant'Orsola-Malpighi Hospital, Bologna, Italy

¹⁵Department for Determinants of Chronic Diseases (DCD), National Institute for Public Health and the Environment (RIVM), Bilthoven, The Netherlands

¹⁶Department of Gastroenterology and Hepatology, University Medical Centre, Utrecht, The Netherlands

¹⁷Department of Epidemiology and Biostatistics, The School of Public Health, Imperial College London, London, United Kingdom

¹⁸Department of Social and Preventive Medicine, Faculty of Medicine, University of Malaya, Kuala Lumpur, Malaysia

¹⁹Department of Laboratory Medicine, University-Hospital of Padova, Padua, Italy

²⁰Department of General, Visceral and Transplantation Surgery, University Hospital Heidelberg, Heidelberg, Germany

²¹Department of Surgery I, Faculty of Medicine and Dentistry, Palacky University Olomouc and University Hospital Olomouc, Olomouc, Czech Republic

²²Third Surgical Clinic - Department of Surgery, Oncology and Gastroenterology (DiSCOG), University of Padova, Padova, Italy

²³Digestive and Liver Disease Unit, S. Andrea Hospital, 'Sapienza' University, Rome, Italy

Key words: pancreatic ductal adenocarcinoma, genetic polymorphisms, lymphocyte telomere length, Mendelian randomization, association

Abbreviations: LTL: leucocyte telomere length; PDAC: pancreatic ductal adenocarcinoma; SNP: simple nucleotide polymorphism; GWAS: genome-wide association study; PANDoRA: PANcreatic Disease ReseArch consortium; OR: odds ratio; CI: confidence interval; FOX: fork-head box

Additional Supporting Information may be found in the online version of this article.

Conflict of interest: The authors declare no conflicts of interest.

D.C. and F.C. conceived the study. M.M. performed experimental work. D.C., F.C. and M.M. performed data analysis. All other authors contributed to the collection of samples and data. D.C. and F.C. drafted the manuscript and all other authors took part in its critical revision.

Grant sponsor: Statistics Netherlands; **Grant sponsor:** World Cancer Research Fund; **Grant sponsor:** Dutch ZON; **Grant sponsor:** Dutch Prevention Funds; **Grant sponsor:** LK Research Funds; **Grant sponsor:** Netherlands Cancer Registry; **Grant sponsor:** Dutch Ministry of Public Health, Welfare and Sports; **Grant sponsor:** Ministry of Education Youth and Sports of the Czech Republic; **Grant numbers:** 1503, LO1503; **Grant sponsor:** Czech Ministry of Education; **Grant numbers:** UP, 61989592, NPS I LO1304; **Grant sponsor:** Czech Ministry of Health; **Grant numbers:** 16-28375A; **Grant sponsor:** DKFZ

DOI: 10.1002/ijc.31928

History: Received 18 Apr 2018; Accepted 13 Sep 2018; Online 16 Oct 2018

Correspondence to: Federico Canzian, Genomic Epidemiology Group, German Cancer Research Center (DKFZ), Im Neuenheimer Feld 280, 69120 Heidelberg, Germany, E-mail: f.canzian@dkfz.de; Tel. +49-6221-421791; Fax +49-6221-421810

- ²⁴PancreatoBiliary Endoscopy and EUS Division, Pancreas Translational and Clinical Research Center, IRCCS San Raffaele Scientific Institute, Vita Salute San Raffaele University, Milan, Italy
- ²⁵Department of Oncology, Institute of Molecular and Translational Medicine, Faculty of Medicine and Dentistry, Palacky University, Olomouc, Czech Republic
- ²⁶University of Cambridge School of Clinical Medicine Clinical Gerontology Unit, Addenbrooke's Hospital, Cambridge, United Kingdom
- ²⁷Department of Gastroenterology, Lithuanian University of Health Sciences, Kaunas, Lithuania
- ²⁸Division of Cancer Epidemiology, German Cancer Research Center (DKFZ), Heidelberg, Germany
- ²⁹Blood Transfusion Service, Azienda Ospedaliero-Universitaria Meyer, Florence, Italy
- ³⁰Gastroenterology and Gastrointestinal Endoscopy Unit, Vita-Salute San Raffaele University, San Raffaele Scientific Institute, Milan, Italy
- ³¹MTA-SZTE Momentum Translational Gastroenterology Research Group, Szeged, Hungary
- ³²Department of General, Visceral and Thoracic Surgery, University Medical Center Hamburg-Eppendorf, Hamburg, Germany
- ³³Division of Gastroenterology and Molecular Biology Lab, IRCCS Ospedale Casa Sollievo Sofferenza, San Giovanni Rotondo, Italy
- ³⁴Department of Digestive Tract Diseases, Medical University of Lodz, Lodz, Poland
- ³⁵Institute of Hematology and Transfusion Medicine, Warsaw, Poland
- ³⁶Cancer Epidemiology Unit, Nuffield Department of Population Health, University of Oxford, Oxford, United Kingdom
- ³⁷ARC-NET, University and Hospital Trust of Verona, Verona, Italy
- ³⁸Division of Abdominal Surgery, IRCCS Ospedale Casa Sollievo Sofferenza, San Giovanni Rotondo, Italy
- ³⁹Division of Preventive Oncology, German Cancer Research Center (DKFZ) and National Center for Tumor Diseases (NCT), Heidelberg, Germany
- ⁴⁰German Cancer Consortium (DKTK), German Cancer Research Center (DKFZ), Heidelberg, Germany

Telomere deregulation is a hallmark of cancer. Telomere length measured in lymphocytes (LTL) has been shown to be a risk marker for several cancers. For pancreatic ductal adenocarcinoma (PDAC) consensus is lacking whether risk is associated with long or short telomeres. Mendelian randomization approaches have shown that a score built from SNPs associated with LTL could be used as a robust risk marker. We explored this approach in a large scale study within the PANcreatic Disease ReseArch (PANDoRA) consortium. We analyzed 10 SNPs (*ZNF676*-rs409627, *TERT*-rs2736100, *CTC1*-rs3027234, *DHX35*-rs6028466, *PXK*-rs6772228, *NAF1*-rs7675998, *ZNF208*-rs8105767, *OBFC1*-rs9420907, *ACYP2*-rs1125529 and *TERC*-rs10936599) alone and combined in a LTL genetic score (“teloscore”, which explains 2.2% of the telomere variability) in relation to PDAC risk in 2,374 cases and 4,326 controls. We identified several associations with PDAC risk, among which the strongest were with the *TERT*-rs2736100 SNP (OR = 1.54; 95%CI 1.35–1.76; $p = 1.54 \times 10^{-10}$) and a novel one with the *NAF1*-rs7675998 SNP (OR = 0.80; 95%CI 0.73–0.88; $p = 1.87 \times 10^{-6}$, $p_{\text{trend}} = 3.27 \times 10^{-7}$). The association of short LTL, measured by the teloscore, with PDAC risk reached genome-wide significance ($p = 2.98 \times 10^{-9}$ for highest vs. lowest quintile; $p = 1.82 \times 10^{-10}$ as a continuous variable). In conclusion, we present a novel genome-wide candidate SNP for PDAC risk (*TERT*-rs2736100), a completely new signal (*NAF1*-rs7675998) approaching genome-wide significance and we report a strong association between the teloscore and risk of pancreatic cancer, suggesting that telomeres are a potential risk factor for pancreatic cancer.

What's new?

How does lymphocyte telomere length affect pancreatic cancer risk? These authors analyzed 10 SNPs associated with telomere length and their relationship with pancreatic cancer risk, using data from the Pancreatic Disease Research (PANDoRA) consortium. Each patient received a “teloscore” based on the combined SNP data, and it turned out that a low teloscore - predicting a short telomere - was associated with increased pancreatic cancer risk. The researchers also identified for the first time a significant genome-wide association between a SNP, *TERT*-rs2736100, and increased pancreatic cancer risk. They also discovered a completely novel association between a SNP, *NAF1*-rs7675998, and decreased risk.

Introduction

Pancreatic cancer is a relatively rare disease, but it currently ranks as the fourth cause of cancer-related deaths in Europe and USA, and is projected to become the second in a few years.¹ There are several established or suggested environmental risk factors for pancreatic cancer such as smoking, heavy alcohol abuse and predisposing conditions like family history of pancreatic cancer, chronic pancreatitis, obesity, pre-existing

diabetes mellitus.^{2,3} In the last few years genome-wide associations studies (GWAS) and targeted large candidate gene/pathway studies have identified several single nucleotide polymorphisms (SNPs) associated with pancreatic cancer susceptibility and survival.^{4–18} Among these reports several point toward a prominent involvement of the *TERT-CLPTMIL* gene region in the disease etiology.^{5,7,10,15} This region, situated on chromosome 5p15.33, is pleiotropic and there are

overwhelming epidemiologic and molecular evidences on the association of SNPs belonging to it and the risk of various cancer types.¹⁹ The pleiotropy of the region is explained by the central role that TERT exerts in the cell. The *TERT* gene encodes the telomerase reverse transcriptase, and with the telomerase RNA component (*TERC* gene) forms a key part of the telomerase enzymatic complex, which synthesizes telomeric ends.²⁰ Even moderate deregulations of the telomerase activity can jeopardize telomere homeostasis²¹, which in turn can affect chromosomal stability, cell growth and the correct segregation of chromosomes to daughter cells.^{22,23} Interestingly, considerable evidence from molecular cancer biology indicates that telomere length in healthy or nonmalignant tissues, usually studied as lymphocyte telomere length (LTL), also represents a risk marker for a large number of tumor types. Telomere length is highly correlated across tissues^{24,25}, therefore LTL is considered a valid surrogate for the measure of telomere length in specific tissues. For pancreatic cancer, five studies attempted to link LTL with risk of developing the disease. The results were contrasting with two studies reporting an association with shorter telomere length and increased risk^{5,26}, one study reporting longer telomere and increased risk²⁷ and two studies reporting a U-shaped association.^{28,29} The lack of consensus for pancreatic cancer reflects the conflicting results reported for other cancer types and it is at least partially due to the techniques, particularly sensitive to sample handling³⁰ and other confounders³¹ such as age, chemotherapy and the epidemiologic design of the study (retrospective vs. prospective).³² The associations between LTL and various types of cancer and the possible caveats to consider have been reviewed by Hou *et al.*³¹ However, LTL variability is under genetic control. In particular, GWAS have identified 11 SNPs associated with LTL. Recent Mendelian randomization approaches have shown that a score built from these SNPs as a surrogate of LTL could be used as a robust risk marker for several cancer types.^{33–38} Two studies attempted this for pancreatic cancer, and found no association.^{39,40} Given that pancreatic cancer is a rare and very lethal disease, it is crucial to expand our knowledge on risk factors, by conducting a Mendelian randomization analysis of telomere length. This is potentially a better way than measuring LTL directly, given the difficulties in precisely determining this phenotype. We explored this approach in a large scale study within the Pancreatic Disease ReseArch (PANDoRA) consortium, by analyzing 10 telomere-defining SNPs separately or in conjunction computing a score.

Materials and Methods

For our study we used 2,374 pancreatic cancer cases and 4,326 controls belonging to the PANDoRA, EPIC and ESTHER consortia. The PANcreatic Disease ReseArch (PANDoRA) consortium has been described in detail elsewhere.⁴¹ We collected cases and controls from 8 European countries (Italy, Germany, Czech Republic, Hungary, United

Kingdom, Lithuania, Poland, Netherlands). Cases were defined by a confirmed diagnosis of PDAC by histopathology. Controls were collected in the same geographical regions as the cases, mostly in the context of the PANDoRA consortium. Additionally, a part of the German controls was enrolled in ESTHER, a prospective cohort with 9,953 participants recruited during a general health check-up between July 2000 and December 2002 in Saarland (a state in South-western Germany). The remaining German controls and all of the British and Dutch controls were selected from healthy volunteers recruited from the general population in the European Prospective Investigation on Cancer (EPIC), an ongoing prospective cohort study in ten European countries (<http://epic.iarc.fr/>). All subjects signed a written consent form. Ethical approval for the PANDoRA study protocol (that in this report also included controls from ESTHER and EPIC cohorts) was received from the Ethics Commission of the Medical Faculty of the University of Heidelberg.

SNP selection

We selected 11 independent SNPs ($r^2 = 0$ for all pairwise comparisons) that were consistently shown by GWAS to influence telomere length.⁴⁰ Our final selection consisted of: *ZNF676*-rs409627, *TERT*-rs2736100, *CTCF*-rs3027234, *DHX35*-rs6028466, *PXK*-rs6772228, *NAF1*-rs7675998, *ZNF208*-rs8105767, *OBFC1*-rs9420907, *ACY2*-rs11125529, *TERC*-rs10936599 and *ZBTB46*-rs755017. The polymorphic variant reported in the original publication for the *ZNF676* gene was rs412658, but the genotyping assay for this SNP failed quality controls, therefore we genotyped instead rs409627, a proxy in perfect linkage disequilibrium ($r^2 = 1$ in all European populations of the 1,000 Genomes project). A list of the selected SNPs with betas, variance explained and all the relevant information can be found in Table 1.

Genotyping

DNA was extracted from whole blood. Genotyping was carried out at the German Cancer Research Center (DKFZ) in Heidelberg, Germany, using TaqMan (ABI, Applied Biosystems, Foster City, CA) technology. Genotyping was conducted in 384-well plates and for quality control duplicates of 10% of the samples were interspersed throughout the plates. The order of DNA samples from case and control subjects was randomized on plates to ensure that similar numbers of cases and controls were analyzed in each batch. PCR plates were read on a ViiA7 real time instrument (Applied Biosystems). The ViiA7 RUO Software, version 1.2.2 (Applied Biosystems) was used to determine genotypes.

Teloscore computation

For each study subject, a SNP score to estimate telomere length (which we called “teloscore”) was computed as follows: for each SNP the number of alleles associated with longer telomeres (according to the results of the literature reported in Table 1) were counted, and added up, resulting in the

Table 1. SNPs associated with telomere length and genotyped in this study¹

SNPs	Chr ²	Pos ²	Gene	Alleles (M/m) ²	Effect allele ²	Beta ²	SE ²	% variance explained ¹	Base pairs ¹
rs409627 ³	19	22,176,638	ZNF676	G/C	C	0.086	0.010	0.484	103.2
rs2736100	5	1,286,401	TERT	C/A	C	0.085	0.013	0.310	102.0
rs3027234	17	8,232,774	CTC1	C/T	C	0.103	0.012	0.292	123.6
rs6028466	20	39,500,359	DHX35	G/A	A	0.058	0.013	0.041	69.6
rs6772228	3	58,390,292	PXK	T/A	T	0.041	0.014	0.200	49.2
rs7675998	4	163,086,668	NAF1	G/A	G	0.048	0.012	0.190	57.6
rs8105767	19	22,032,639	ZNF208	A/G	G	0.064	0.011	0.090	76.8
rs9420907	10	103,916,707	OBFC1	A/C	C	0.142	0.014	0.171	170.4
rs11125529	2	54,248,729	ACYP2	C/A	A	0.065	0.012	0.080	78.0
rs10936599	3	169,774,313	TERC	C/T	C	0.100	0.011	0.319	120.0

¹Data from Refs. 40,52.

²Chr = chromosome; pos = base-pair position (GRCh38.p3); Effect allele = allele associated with longer telomeres; Beta = standard deviation change in telomere length per copy of the effect allele; SE = standard error; Base pairs = telomere length difference in base pairs associated with each allele.

³Surrogate of rs412658 ($r^2 = 1$).

unweighted score for each subject. Since we finally selected 10 SNPs, the unweighted score can assume any integer value between 0 (shortest telomeres) and 20 (longest telomeres). We then created a weighted score for each study subject. First, we took from the literature estimates of the per-allele effect on LTL in base pairs for each SNP (Table 1). Then, we multiplied at each SNP the number of alleles associated with longer telomeres by the per-allele effect on LTL in base pairs. Finally, we summed up these quantities for each study subject. The weighted score thus represents the estimated difference in telomere length, measured in base pairs, attributable to the SNPs under investigation. Only a subset of the study subjects had a 100% SNP call rate ($N = 1,246$ cases (52.5%), 1945 controls (45.0%), total 3,191 (47.6%)), while the remaining subjects had a call rate between 80% and 100%. Therefore, in order to be able to compute comparable score values for all study subjects, we also considered average values for each score. Supporting Information Table 1 shows examples of how the teloscores were generated.

Statistical analysis

The association between the SNPs and PDAC risk was tested using unconditional logistic regression computing odds ratios (OR) and 95% confidence intervals (CI). We used co-dominant, dominant, recessive and per-allele models of inheritance, calculating also a trend test for the co-dominant model. The threshold for statistical significance was therefore $p = 0.05 / (10 \text{ SNPs} \times 4 \text{ models}) = 0.00125$.

We used each of the teloscores (weighted and unweighted) as continuous variables and as discrete values, calculating quintiles based on the distribution of values of the healthy controls. The association between the teloscores and PDAC risk was tested with logistic regression, computing ORs and 95% CIs.

For a subset of German controls from the ESTHER cohort ($N = 885$), Spearman's correlation coefficients were calculated

between the teloscores and values of relative telomere length previously obtained with a real-time quantitative PCR protocol.⁴²

All analyses were adjusted for age, sex and geographic region of origin. Additional analyses were performed including, as adjustment factors, also tobacco smoking, diabetes diagnosed at least two years before onset of pancreatic cancer and family history of pancreatic cancer, which were available for subsets of cases and controls (Supporting Information Table 2). We also tested the association between the teloscore and smoking and diabetes as endpoints. Egger regression was used to test for possible pleiotropic effects of our genetic instrument. All statistical tests were two-sided.

Bioinformatic tools

We used several bioinformatic tools to assess possible functional relevance for the three SNPs showing the most significant associations with risk of pancreatic cancer. RegulomeDB (<http://regulome.stanford.edu/>)⁴³ and HaploReg⁴⁴ were used to identify the regulatory potential of the region nearby each SNP. The GTEx portal web site was used to identify potential associations between the SNP and expression levels of nearby genes (eQTL).⁴⁵

Results

Data filtering and quality control

Relevant characteristics of the study population are shown in Table 2. All the genotyped SNPs were in Hardy-Weinberg equilibrium when analyzed in controls with the exception of the polymorphic variant ZBTB46-rs755017 that was therefore excluded from the statistical analysis and from the score computations. Subjects with a call rate lower than 80% ($N = 272$ controls, 361 cases, total 633) were excluded from further analyses. This left 2,374 cases and 4,326 controls, for whom the average SNP call rate was 95.7%, with a minimum of 81.81% (ACYP2-rs11125529) and a maximum of 98.99% (CTC1-

Table 2. Description of the study population

	PDAC cases	Controls
Country/region		
Germany	789	1,779
Northern Italy	447	540
Central Italy	382	535
Southern Italy	103	499
Czech Republic	243	156
Poland	74	191
Lithuania	47	172
Netherlands	106	102
Hungary	95	176
United Kingdom	88	176
Total	2,374	4,326
Sex		
Male	1,342	2,178
Female	1,008	2,079
Median age	65.6	59.0
(25th–75th percentile)	57.8–72.3	49.7–66.0

rs3027234). Quality control analysis showed a concordance rate of 98.85%.

SNP main effects

When analyzing the effect of the SNPs on PDAC risk we observed several statistically significant associations. The strongest from a statistical point of view was between the homozygous of the minor allele (A) compared to the carriers of the C allele of the *TERT*-rs2736100 SNP (OR = 1.54; 95% CI 1.35–1.76; $p = 1.54 \times 10^{-10}$). The association with the second lowest p-value was between carriers of the minor A allele of the *NAF1*-rs7675998 SNP and decreased risk of PDAC (OR = 0.80; 95%CI 0.73–0.88; $p = 1.87 \times 10^{-6}$, $p_{trend} = 3.27 \times 10^{-7}$). We observed two additional signals that were very close to the threshold for multiple testing, both assuming a recessive model of inheritance: *ZNF676*-rs409627 (OR = 0.76; 95%CI 0.64–0.91; $p = 0.003$) and *ZNF208*-rs8105767 (OR = 0.69; 95% CI 0.54–0.87; $p = 0.002$). The results of this analysis are presented in Table 3.

Association of the “teloscore” with telomere length measurement and PDAC risk

As a first step we checked whether the computed teloscore was effectively able to predict telomere length. For this purpose we used part of the controls for which we had previously measured telomere length with a real-time quantitative PCR protocol⁴² and we observed a statistically significant association between the teloscore and LTL with a correlation coefficient of 0.122 ($p = 0.0017$), confirming the hypothesized association between the genetic variance in telomeric genes and telomere length. In this subset of controls the 10 SNPs collectively explain 3.35% of the telomere length variation. We subsequently tested the association between

Table 3. Association between the individual SNPs and risk of developing PDAC

Gene	SNP	Alleles (M/m) ¹	MM vs. Mm			Mm vs. Mm + mm			Mm + MM vs. mm			M vs. m			
			Mm	mm	OR (95%CI) ²	p_{het}	p_{hom}	OR (95%CI)	p_{het}	p_{hom}	OR (95%CI)	p_{rec}	OR (95%CI) ²	p	
<i>ZNF676</i>	rs409627 ³	G/C	710/1633	803/1864	230/616	1.04 (0.91–1.19)	0.563	0.011	0.1552	0.97 (0.86–1.10)	0.656	0.76 (0.64–0.91)	0.003	0.92 (0.84–1.01)	0.071
<i>TERT</i>	rs2736100	C/A	598/1226	983/1979	581/818	1.02 (0.89–1.17)	0.778	2.60 × 10 ⁻⁸	1.05 × 10 ⁻⁶	1.17 (1.03–1.33)	0.013	1.54 (1.35–1.76)	1.54 × 10 ⁻¹⁰	1.24 (1.15–1.34)	1.08 × 10 ⁻⁷
<i>CTC1</i>	rs3027234	C/T	1328/2521	869/1528	143/240	1.09 (0.97–1.22)	0.154	0.490	0.1017	1.09 (0.97–1.21)	0.137	1.05 (0.84–1.32)	0.673	1.06 (0.97–1.16)	0.17
<i>DHX35</i>	rs6028466	G/A	2085/3687	215/448	18/24	0.86 (0.72–1.04)	0.118	0.716	0.2118	0.88 (0.73–1.05)	0.156	1.14 (0.60–2.19)	0.684	0.91 (0.77–1.07)	0.233
<i>PXK</i>	rs6772228	T/A	2141/3738	149/224	6/12	1.15 (0.91–1.46)	0.242	0.935	0.2545	1.15 (0.91–1.45)	0.247	1.04 (0.37–2.93)	0.948	1.13 (0.91–1.40)	0.273
<i>NAF1</i>	rs7675998	G/A	1451/2444	753/1548	102/284	0.81 (0.72–0.92)	0.001	1.18 × 10 ⁻⁴	3.27 × 10 ⁻⁷	0.78 (0.70–0.88)	0.001	0.66 (0.51–0.84)	0.001	0.80 (0.73–0.88)	1.87 × 10 ⁻⁶
<i>ZNF208</i>	rs8105767	A/G	1197/1701	912/1391	200/343	0.98 (0.85–1.12)	0.719	0.002	0.0391	0.92 (0.80–1.04)	0.184	0.69 (0.54–0.87)	0.002	0.88 (0.80–0.98)	0.018
<i>OBFC1</i>	rs9420907	A/C	1639/2992	640/1174	68/115	0.99 (0.88–1.12)	0.883	0.403	0.8429	1.00 (0.89–1.13)	0.943	1.16 (0.83–1.62)	0.392	1.02 (0.92–1.13)	0.746
<i>ACYP2</i>	rs11125529	C/A	1732/2441	474/719	39/74	1.04 (0.90–1.20)	0.598	0.441	0.0941	1.02 (0.89–1.178)	0.767	0.84 (0.55–1.28)	0.415	1.00 (0.88–1.13)	0.987
<i>TERC</i>	rs10936599	C/T	1372/2447	782/1554	134/275	0.90 (0.80–1.01)	0.081	0.117	0.0365	0.89 (0.80–1.00)	0.040	0.86 (0.69–1.09)	0.208	0.91 (0.83–0.99)	0.032

¹M = major allele; m = minor allele.

²OR = odds ratio; CI = confidence interval. All analyses were adjusted for age, sex and geographic region of origin. Results in bold are statistically significant ($p < 0.05$).

³Surrogate of rs412658 ($r^2 = 1$).

Table 4. Association between teloscore and PDAC risk

Score ¹	Controls	Cases	OR	95% CI	<i>P</i> value
Quintile 1 (0–47.22)	865	580	Ref.	–	–
Quintile 2 (47.23–55.30)	865	555	0.99	(0.85–1.17)	0.95
Quintile 3 (55.31–61.80)	866	426	0.74	(0.63–0.88)	5.30×10^{-4}
Quintile 4 (61.81–70.56)	864	469	0.80	(0.67–0.94)	7.48×10^{-3}
Quintile 5 (70.57–112.05)	866	344	0.59	(0.49–0.70)	2.98×10^{-9}
Continuous variable	4,326	2,374	0.88	(0.85–0.92)	1.82×10^{-10}

¹Weighted average teloscore, calculated as described in the Methods section and in Supporting Information Table S1. Quintiles were calculated based on the distribution of values of the controls. Numbers in parentheses represent the value in bp that defines the boundaries of each quintile. All analyses were adjusted for age, sex and geographic region of origin. The unit for the "continuous variable" is the increase of one quintile.

the score and PDAC risk. Since not all the individuals were genotyped successfully for all the selected SNPs, in order to increase our statistical power we used the average scores rather than the absolute values (see methods). Considering the average score we observed a strong association between genetically determined long telomere and decreased risk of PDAC when analyzing the score as a categorical variable (OR = 0.59; 95%CI 0.49–0.70; $p = 2.98 \times 10^{-9}$ for highest vs. lowest quintile) and also as a continuous variable (OR = 0.88; 95%CI 0.85–0.92; $p = 1.82 \times 10^{-10}$). The results are shown in Table 4.

We checked possible associations between the teloscore and known risk factors for pancreatic cancer, namely tobacco smoking and diabetes diagnosed before onset of pancreatic cancer. No association was found. Furthermore, we recalculated the association between the teloscore and pancreatic cancer risk by adding the risk factors as adjustment variables, but no substantial difference was observed (data not shown).

In order to explore the possibility that some of the SNPs could have a pleiotropic association with pancreatic cancer risk, we recalculated the teloscore without rs7675998 and rs2736100, that show the most significant associations with pancreatic cancer risk, and the results did not substantially change (data not shown). However Egger regression results were not statistically significant ($p = 0.738$).

Possible functional effects

We used several bioinformatic tools to test for possible functional relevance of the four variants that reached study-wide significance (*TERT*-rs2736100, *NAF1*-rs7675998, *ZNF676*-rs409627, *ZNF208*-rs8105767). RegulomeDB did not reveal any interesting regulatory potential associated with any of the variants. The GTEx portal web site, instead, showed that all the SNPs, with the exception of *TERT*-rs2736100, are multi-tissue eQTLs ($p < 1.1 \times 10^{-4}$). For *TERT*-rs2736100 there were no significant associations with gene expression levels in pancreatic tissue. It is interesting to note that, according to GTEx, *ZNF676*-rs409627 modulates the expression of *ZNF676* in the pancreatic tissue (effect size 0.59, $p = 2.2 \times 10^{-6}$).

Discussion

There are overwhelming epidemiologic and molecular evidences linking telomeres with the etiology of numerous diseases. However, given the capricious nature of association studies and the technical pitfalls in LTL measurement, both short and long telomeres have been associated with the onset of multiple cancer types. The situation is particularly unclear for pancreatic cancer, with five published studies that measured LTL with a real-time quantitative PCR protocol.^{5,26–29} Among these studies, four were conducted in prospective cohorts^{5,26,28,29} and one in a retrospective case-control series²⁷. Two studies found an association between shorter telomeres and increased risk of pancreatic cancer, one found an association with longer telomeres and two found associations with both longer and shorter telomeres (Supporting Information Table 3).

Additionally, two studies that used a genetic risk score reported no evidence for association with pancreatic cancer risk.^{39,40}

The aims of our study were to test whether telomere-related SNPs could modulate pancreatic cancer risk, and to use genetic markers of telomere length in order to understand whether longer or shorter LTL increase the risk of developing PDAC.

We observed a genome-wide significant association ($p = 1.54 \times 10^{-10}$) between the *TERT*-rs2736100 A allele and increased PDAC risk. This SNP is pleiotropic and has been reported to be associated, alongside telomere length, with several cancer types. *TERT*-rs2736100 has been reported by others⁵ and by ourselves⁷ to be associated with PDAC risk, but this is the first time that the association reaches a genome-wide level of significance. This SNP is in very low linkage disequilibrium with the other SNPs in this region that were reported to be associated with pancreatic cancer risk (rs401681 $r^2 = 0.01$, rs2736098 $r^2 = 0.114$) and therefore represents an independent signal. A functional explanation for the consistent associations between this SNP and cancer risk has yet to be found however, since the minor allele is associated both with increased PDAC risk and with decreased LTL.

The association between *NAF1*-rs7675998 SNP and decreased risk of PDAC is novel. The association is close to a

genome-wide significance level ($p_{\text{trend}} = 3.27 \times 10^{-7}$). The *NAF1* (nuclear assembly factor 1) gene product is part of a complex involved in the assembly of telomerase⁴⁶ and is therefore intimately linked to the telomerase activity and telomere length. According to HaploReg *NAF1*-rs7675998 has 43 variants in high LD ($r^2 > 0.8$) and 41 of them (as well as rs7675998 itself) are predicted to alter several regulatory motifs. In particular, rs7675998 is predicted to alter 19 regulatory motifs including those of the forkhead box (FOX) family. According to GTEx this SNP has also two eQTLs affecting *NAF1* expression. However, although these associations are highly significant, they have not been observed in the pancreatic tissue. We observed two other potentially interesting associations between *ZNF676*-rs409627, *ZNF208*-rs8105767 and PDAC risk. The role of these two genes in telomere maintenance has not been established yet, although several hypotheses point to a possible involvement in stabilizing DNA or proteins that bind to DNA.⁴⁷ According to GTEx, rs409627 can modify *ZNF676* expression in the pancreatic tissue while rs8105767 can modify the expression of *ZNF208* in various tissues but not in pancreatic cancer. For both SNPs the allele associated with an increase in risk is the major allele, while the allele associated with telomere shortening is the minor one, indicating that possibly their association with PDAC risk is independent from telomere length.

The most important novel finding of our study is the statistically significant association between genetically determined short LTL (assessed through the teloscore) and increased risk of PDAC. The association reached genome-wide significance both considering the variable as categorical ($p = 2.98 \times 10^{-9}$ for highest vs. lowest quintile) or as continuous ($p = 1.82 \times 10^{-10}$) and does not support a U-shaped association. It should also be noted as a proof of principle that we found a weak but significant correlation between the teloscore and LTL measured by an established method (real-time quantitative PCR) in almost 900 controls belonging to our dataset. In the last couple of years the approach of using SNPs related to telomere shortening as an instrumental mean to infer the effect of telomeres on cancer etiology has been successfully used in different tumor types such as B-cell lymphoma³⁵, adult glioma³⁶, breast cancer³⁴ and squamous cell carcinoma of the head and neck.⁴⁸ The use of genetic markers decreases the risk for reverse causation bias and therefore the differences in the studies (some finding association between cancer risk and longer telomeres, some with shorter telomeres) may reflect tissue-specific effects and activity of TL or a specific regulation of the genes involved in telomere regulation. It is interesting to note that *TERT*-rs2736100 has been consistently associated with several cancer types but the allele increasing the risk is not always the same.¹⁹ Given the strong effect of genetic variants on LTL and given that the allele associated with telomere shortening is always the same, the difference in LTL association with cancer risk may be explained by the different activity of the gene in different tissues.

Two studies previously attempted this analysis in PDAC, but did not find an association.^{39,40} It is difficult to speculate about the reasons for the discordance with our results. It should be noticed that the results reported by Haycock *et al.*, based on the PanScan GWAS (5,105 cases and 8,739 controls), show a nonsignificant associations between shorter telomeres and pancreatic cancer risk (OR = 0.86; 95%CI 0.56–1.32; $p = 0.50$ for PanScan and OR = 0.74, 95%CI 0.53–1.02, $p = 0.0657$ for PanC4), which are compatible with our results.⁴⁰ The results of the other study do not show any association between a teloscore of 8 LTL-associated SNPs and pancreatic cancer risk (OR = 1.04; 95%CI 0.97–1.12; $p = 0.228$), although the sample size was smaller than in our study (1,500 cases and 1,500 controls). Moreover, their score was calculated in a different way from ours (i.e. according to a dominant model, whereby study subjects with one or two copies of the allele associated with shorter telomeres were combined into one group and compared to those who carry two copies of the allele associated with longer telomeres).³⁹

Telomere shortening is known to be present in the first stage of pancreatic onset⁴⁹ and it could be an important determinant of cell progression to malignant state.⁵ Constitutionally shorter telomeres, as determined by germline polymorphisms, may contribute to the very early phases of premalignant transformation of pancreatic cells.

Our study has several obvious advantages: the large scale and the ability to test the teloscore in a group of individuals for which telomere length was measured by RT-PCR homogeneously, in the same laboratory, in samples collected from the same center (the controls belonging to the ESTHER cohort $n = 885$) and using exactly the same procedure for sample handling and storing. A possible drawback is that we tested the teloscore on DNA collected from leukocytes and it is therefore difficult to generalize its ability to be used as a proxy for other tissues. However, there is a growing literature suggesting that telomere shortening is generally consistent in different tissues⁵⁰ and that the variation among different tissues belonging to the same individual is lower than the variability between different individuals.^{49,51} Additionally, an analysis with Egger regression did not yield a significant result, and pointed to high heterogeneity among SNPs, suggesting a possible pleiotropic effect of our SNPs.

In conclusion, here we present a novel genome-wide candidate for PDAC (*TERT*-rs2736100) and a completely new signal for PDAC in *NAF1*-rs7675998 that approaches the genome-wide threshold. In addition, we found a strong association between the teloscore and risk of pancreatic cancer, suggesting that telomeres are a potential risk factor for pancreatic cancer.

Acknowledgements

The authors wish to thank Hartwig Ziegler, Christa Stegmaier, Sonja Wolf, and Volker Herrmann for their outstanding contributions to the conduct of the ESTHER study. This work was partially supported by intramural

funds of the DKFZ; the Czech Ministry of Health (grant number 16-28375A to B. Mohelnikova-Duchonova); the Czech Ministry of Education (NPS I LO1304 and DRO (UP, 61989592); National Sustainability Program I (NPU I) of the Ministry of Education Youth and Sports of the Czech Republic (grant number LO1503 to P. Soucek; grant number 1503

to P. Vodicka and L. Voldickova); Dutch Ministry of Public Health, Welfare and Sports (VWS), Netherlands Cancer Registry (NKR), LK Research Funds, Dutch Prevention Funds, Dutch ZON (Zorg Onderzoek Nederland), World Cancer Research Fund (WCRF), Statistics Netherlands (The Netherlands) to Bas Bueno de Mesquita.

References

- Hidalgo M, Cascinu S, Kleeff J, et al. Addressing the challenges of pancreatic cancer: future directions for improving outcomes. *Pancreatology* 2015;15:8–18.
- Barone E, Corrado A, Gemignani F, et al. Environmental risk factors for pancreatic cancer: an update. *Arch Toxicol* 2016;90:2617–42.
- Maisonneuve P, Lowenfels AB. Risk factors for pancreatic cancer: a summary review of meta-analytical studies. *Int J Epidemiol* 2015;44:186–98.
- Amundadottir L, Kraft P, Stolzenberg-Solomon RZ, et al. Genome-wide association study identifies variants in the ABO locus associated with susceptibility to pancreatic cancer. *Nat Genet* 2009;41:986–90.
- Bao Y, Prescott J, Yuan C, et al. Leucocyte telomere length, genetic variants at the TERT gene region and risk of pancreatic cancer. *Gut* 2016;66:1116–22.
- 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;7:57011–20.
- Campa D, Rizzato C, Stolzenberg-Solomon R, et al. TERT gene harbors multiple variants associated with pancreatic cancer susceptibility. *Int J Cancer* 2015;137:2175–83.
- Childs EJ, Mocci E, Campa D, et al. Common variation at 2p13.3, 3q29, 7p13 and 17q25.1 associated with susceptibility to pancreatic cancer. *Nat Genet* 2015;47:911–6.
- Low SK, Kuchiba A, Zembutsu H, et al. Genome-wide association study of pancreatic cancer in Japanese population. *PLoS One* 2010;5:e11824.
- Petersen GM, Amundadottir L, Fuchs CS, 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.
- Rizzato C, Campa D, Giese N, et al. Pancreatic cancer susceptibility loci and their role in survival. *PLoS One* 2011;6:e27921.
- 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. *Oncol Rep* 2013;29:1637–44.
- Rizzato C, Campa D, Talar-Wojnarowska R, et al. Association of genetic polymorphisms with survival of pancreatic ductal adenocarcinoma patients. *Carcinogenesis* 2016;37:957–64.
- Willis JA, Olson SH, Orlov I, et al. A replication study and genome-wide scan of single-nucleotide polymorphisms associated with pancreatic cancer risk and overall survival. *Clin Cancer Res* 2012;18:3942–51.
- Wolpin BM, Rizzato C, Kraft P, et al. Genome-wide association study identifies multiple susceptibility loci for pancreatic cancer. *Nat Genet* 2014;46:994–1000.
- Wu C, Kraft P, Stolzenberg-Solomon R, et al. Genome-wide association study of survival in patients with pancreatic adenocarcinoma. *Gut* 2012;63:152–60.
- Wu C, Miao X, Huang L, et al. Genome-wide association study identifies five loci associated with susceptibility to pancreatic cancer in Chinese populations. *Nat Genet* 2011;44:62–6.
- 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;7:66328–43.
- Mocellin S, Verdi D, Pooley KA, et al. Telomerase reverse transcriptase locus polymorphisms and cancer risk: a field synopsis and meta-analysis. *J Natl Cancer Inst* 2012;104:840–54.
- Blackburn EH. Switching and signaling at the telomere. *Cell* 2001;106:661–73.
- Armanios M. Telomeres and age-related disease: how telomere biology informs clinical paradigms. *J Clin Invest* 2013;123:996–1002.
- Martinez P, Blasco MA. Telomeric and extra-telomeric roles for telomerase and the telomere-binding proteins. *Nat Rev Cancer* 2011;11:161–76.
- McEachern MJ, Krauskopf A, Blackburn EH. Telomeres and their control. *Annu Rev Genet* 2000;34:331–58.
- Kimura M, Gazitt Y, Cao X, et al. Synchrony of telomere length among hematopoietic cells. *Exp Hematol* 2010;38:854–9.
- Wilson WR, Herbert KE, Mistry Y, et al. Blood leucocyte telomere DNA content predicts vascular telomere DNA content in humans with and without vascular disease. *Eur Heart J* 2008;29:2689–94.
- Lynch SM, Major JM, Cawthon R, et al. A prospective analysis of telomere length and pancreatic cancer in the alpha-tocopherol beta-carotene cancer (ATBC) prevention study. *Int J Cancer* 2013;133:2672–80.
- Skinner HG, Gangnon RE, Litzelman K, et al. Telomere length and pancreatic cancer: a case-control study. *Cancer Epidemiol Biomarkers Prev* 2012;21:2095–100.
- Campa D, Mergarten B, De Vivo I, et al. Leucocyte telomere length in relation to pancreatic cancer risk: a prospective study. *Cancer Epidemiol Biomarkers Prev* 2014;23:2447–54.
- Zhang R, Zhao J, Xu J, et al. Association of peripheral leukocyte telomere length and its variation with pancreatic cancer and colorectal cancer risk in Chinese population. *Oncotarget* 2016;7:38579–85.
- Cunningham JM, Johnson RA, Litzelman K, et al. Telomere length varies by DNA extraction method: implications for epidemiologic research. *Cancer Epidemiol Biomarkers Prev* 2013;22:2047–54.
- Hou L, Zhang X, Gawron AJ, et al. Surrogate tissue telomere length and cancer risk: shorter or longer? *Cancer Lett* 2012;319:130–5.
- Pooley KA, Sandhu MS, Tyrer J, et al. Telomere length in prospective and retrospective cancer case-control studies. *Cancer Res* 2010;70:3170–6.
- Iles MM, Bishop DT, Taylor JC, et al. The effect on melanoma risk of genes previously associated with telomere length. *J Natl Cancer Inst* 2014;106:dju267.
- Luu HN, Long J, Wen W, et al. Association between genetic risk score for telomere length and risk of breast cancer. *Cancer Causes Control* 2016;27:1219–28.
- Machiela MJ, Lan Q, Slager SL, et al. Genetically predicted longer telomere length is associated with increased risk of B-cell lymphoma subtypes. *Hum Mol Genet* 2016;25:1663–76.
- Walsh KM, Codd V, Rice T, et al. Longer genotypically-estimated leukocyte telomere length is associated with increased adult glioma risk. *Oncotarget* 2015;6:42468–77.
- Walsh KM, Codd V, Smirnov IV, et al. Variants near TERT and TERC influencing telomere length are associated with high-grade glioma risk. *Nat Genet* 2014;46:731–5.
- Zhang C, Doherty JA, Burgess S, et al. Genetic determinants of telomere length and risk of common cancers: a Mendelian randomization study. *Hum Mol Genet* 2015;24:5356–66.
- Antwi SO, Bamlet WR, Broderick BT, et al. Genetically predicted telomere length is not associated with pancreatic cancer risk. *Cancer Epidemiol Biomarkers Prev* 2017;26:971–4.
- Haycock PC, Burgess S, Nounu A, et al. Association between telomere length and risk of cancer and non-neoplastic diseases: a Mendelian randomization study. *JAMA Oncol* 2017;3:636–51.
- Campa D, Rizzato C, Capurso G, et al. Genetic susceptibility to pancreatic cancer and its functional characterisation: the PANcreatic disease ReseArch (PANDORA) consortium. *Dig Liver Dis* 2013;45:95–9.
- Campa D, Martino A, Varkonyi J, et al. Risk of multiple myeloma is associated with polymorphisms within telomerase genes and telomere length. *Int J Cancer* 2015;136:E351–8.
- Boyle AP, Hong EL, Hariharan M, et al. Annotation of functional variation in personal genomes using RegulomeDB. *Genome Res* 2012;22:1790–7.
- 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.
- GTEX Consortium. The genotype-tissue expression (GTEx) project. *Nat Genet* 2013;45:580–5.
- Egan ED, Collins K. An enhanced H/ACA RNP assembly mechanism for human telomerase RNA. *Mol Cell Biol* 2012;32:2428–39.
- Klug A. The discovery of zinc fingers and their applications in gene regulation and genome manipulation. *Annu Rev Biochem* 2010;79:213–31.
- Gu Y, Yu C, Miao L, et al. Telomere length, genetic variants and risk of squamous cell

- carcinoma of the head and neck in southeast Chinese. *Sci Rep* 2016;6:20675.
49. van Heek NT, Meeker AK, Kern SE, et al. Telomere shortening is nearly universal in pancreatic intraepithelial neoplasia. *Am J Pathol* 2002;161:1541–7.
50. Daniali L, Benetos A, Susser E, et al. Telomeres shorten at equivalent rates in somatic tissues of adults. *Nat Commun* 2013;4:1597.
51. Hunt SC, Chen W, Gardner JP, et al. Leukocyte telomeres are longer in African Americans than in whites: the National Heart, Lung, and Blood Institute family heart study and the Bogalusa heart study. *Aging Cell* 2008;7:451–8.
52. Codd V, Nelson CP, Albrecht E, et al. Identification of seven loci affecting mean telomere length and their association with disease. *Nat Genet* 2013;45:427e1–2.