Hum Genet
DOI 10.1007/s00439-016-1715-1

ORIGINAL INVESTIGATION

Multiple rare variants in high-risk pancreatic cancer-related genes may increase risk for pancreatic cancer in a subset of patients with and without germline *CDKN2A* mutations

Xiaohong R. Yang[1] · Melissa Rotunno[1,2] · Yanzi Xiao[1] · Christian Ingvar[3] · Hildur Helgadottir[4] · Lorenza Pastorino[5,11] · Remco van Doorn[6] · Hunter Bennett[1] · Cole Graham[1] · Joshua N. Sampson[1] · Michael Malasky[1,7] · Aurelie Vogt[1,7] · Bin Zhu[1,7] · Giovanna Bianchi-Scarra[5,11] · William Bruno[5,11] · Paola Queirolo[8] · Giuseppe Fornarini[8] · Johan Hansson[4] · Rainer Tuominen[4] · Laurie Burdett[1,7] · Belynda Hicks[1,7] · Amy Hutchinson[1,7] · Kristine Jones[1,7] · Meredith Yeager[1,7] · Stephen J. Chanock[1] · Maria Teresa Landi[1] · Veronica Höiom[4] · Håkan Olsson[9] · Nelleke Gruis[6] · Paola Ghiorzo[5,11] · Margaret A. Tucker[1] · Alisa M. Goldstein[1,10]

Received: 5 April 2016 / Accepted: 16 July 2016

Abstract The risk of pancreatic cancer (PC) is increased in melanoma-prone families but the causal relationship between germline *CDKN2A* mutations and PC risk is uncertain, suggesting the existence of non-*CDKN2A* factors. One genetic possibility involves patients having mutations in multiple high-risk PC-related genes; however, no systematic examination has yet been conducted. We used next-generation sequencing data to examine 24 putative PC-related genes in 43 PC patients with and 23 PC patients without germline *CDKN2A* mutations and 1001 controls. For each gene and the four pathways in which they occurred, we tested whether PC patients (overall or *CDKN2A*+ and *CDKN2A*− cases separately) had an increased number of rare nonsynonymous variants. Overall, we identified 35 missense variants in PC patients, 14 in *CDKN2A*+ and 21 in *CDKN2A*− PC cases. We found nominally significant associations for mismatch repair genes (*MLH1*, *MSH2*, *MSH6*, *PMS2*) in all PC patients and for *ATM*, *CPA1*, and *PMS2* in *CDKN2A*− PC patients. Further, nine *CDKN2A*+ and four *CDKN2A*− PC patients had rare potentially deleterious variants in multiple PC-related genes. Loss-of-function variants were only observed in *CDKN2A*− PC patients, with *ATM* having the most pathogenic variants. Also, *ATM* variants ($n = 5$) were only

Electronic supplementary material The online version of this article (doi:10.1007/s00439-016-1715-1) contains supplementary material, which is available to authorized users.

✉ Alisa M. Goldstein
goldstea@mail.nih.gov

1 Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892, USA

2 Division of Cancer Control and Population Studies, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA

3 Department of Surgery, Lund University Hospital, Lund, Sweden

4 Department of Oncology Pathology, Karolinska Institutet and Karolinska University Hospital, Solna, Stockholm, Sweden

5 Department of Internal Medicine and Medical Specialties, University of Genoa, Genoa, Italy

6 Department of Dermatology, Leiden University Medical Center, Leiden, The Netherlands

7 Cancer Genomics Research Laboratory, Frederick National Laboratory for Cancer Research, Leidos Biomedical Research, Inc., Frederick, MD, USA

8 Medical Oncology Unit, IRCCS AOU San Martino-IST, Genoa, Italy

9 Department of Oncology, Lund University Hospital, Lund, Sweden

10 9609 Medical Center Dr, Bethesda, MD 20892-9769, USA

11 Genetics of Rare Cancers, IRCCS AOU San Martino-IST, Genoa, Italy

Published online: 23 July 2016

observed in *CDKN2A*− PC patients with a family history that included digestive system tumors. Our results suggest that a subset of PC patients may have increased risk because of germline mutations in multiple PC-related genes.

Introduction

Germline mutations in *CDKN2A*, the major known high-risk melanoma susceptibility gene, have been described in 20–40 % of familial melanoma kindreds (Goldstein et al. 2006). Several features have been shown to be associated with an increased frequency of *CDKN2A* mutations, most notably the occurrence of pancreatic cancer (PC) in a family (Bergman and Gruis 1996; Borg et al. 2000; Ghiorzo et al. 2004; Goldstein et al. 1995, 2006; Lynch et al. 2002; Vasen et al. 2000). At present, though, it is unclear what additional genetic, intrinsic, or extrinsic factors predispose individuals in these families to PC. Further, although specific *CDKN2A* mutations have been associated with PC, many melanoma-prone families with these mutations do not have an excess of PC. These observations suggest that other factors such as mutations in other genes, or tobacco or other carcinogenic exposures may be important in the development of pancreatic cancer (Goldstein 2004; Goldstein et al. 2006; Helgadottir et al. 2014). Further, since PC is such a rapidly fatal cancer, it is important to learn whether additional genetic alterations make individuals more susceptible. Therefore, several research groups that study familial melanoma and have multiple PC families with and without *CDKN2A* mutations collaborated to investigate whether these patients have mutations in multiple high-risk PC-related susceptibility genes.

Twenty-four putative high-risk susceptibility genes for familial PC have been identified (Supplemental Table 1) (Grant et al. 2015; Klein 2013). Many of these genes are part of autosomal dominant (AD) hereditary cancer syndromes such as Peutz–Jeghers, breast-ovarian cancer, and Lynch syndrome [comprising the mismatch repair (MMR) genes] of which PC has been proposed to be a component cancer. In addition, several genes (e.g. *ATM*, *PALB2*, *FANCA*, *FANCC*, *XRCC2*) are mutated in both autosomal recessive (AR) and AD disorders; for example, diseases such as ataxia telangiectasia and Fanconi anemia require two mutations for disease manifestation (bi-allelic) yet later onset cancers such as breast and pancreatic cancer are inherited in a mono-allelic AD pattern. Pancreatic cancer risk is also increased among patients with hereditary/chronic pancreatitis. Since PC patients with *CDKN2A* mutations have not been assessed for mutations in these genes, we used exome sequencing data to systematically examine these 24 susceptibility genes in pancreatic cancer patients with *CDKN2A* mutations. In addition, we also evaluated additional available PC patients without *CDKN2A* mutations but with a family history of cancer, primarily melanoma or digestive system tumors, that were accrued by the same research groups.

Materials and methods

Study population

The PC patients included in this study came from ongoing studies conducted in the United States, Italy, The Netherlands, and Sweden. Details of the source populations, patients, *CDKN2A* mutation status, and study references are presented in Supplemental Table 2. All diagnoses of PC were confirmed by review of pathology reports, medical records, or death certificates. Only deceased PC patients with available blood DNA were included in this study. Each study was approved by its local Institutional Review Board and informed consent was obtained for all participants. All PC patients from a family with a *CDKN2A* mutation carried their respective family's *CDKN2A* mutation. Supplemental Table 3 summarizes the sample population by geographic origin and *CDKN2A* mutation status.

Sequencing methods

Whole exome sequencing

Whole exome sequencing (WES) for the PC patients was performed at the Cancer Genomics Research Laboratory, National Cancer Institute (CGR, NCI), as previously described (Shi et al. 2014). Briefly, SeqCAP EZ Human Exome Library v3.0 (Roche NimbleGen, Madison, WI) was utilized for exome sequence capture. The captured DNA (1.1 μg) was then subject to paired-end sequencing utilizing the Illumina HiSeq 2000 sequencer for 2 × 100 bp sequencing of paired-ends (Illumina, San Diego, CA, USA). For this sample set WES was performed such that 91 % of coding sequence from the University of California, Santa Cruz (UCSC) human genome (hg) 19 transcripts database had >15 reads with average coverage of 71×. The exome data for this paper are archived in the CGR exome build Ensemble_New_Annotation dated 2015-10-28.

Bioinformatics analysis

Alignment and calling of variants

Details of the bioinformatics pipeline for variant alignment and calling used in this study have been previously published (Shi et al. 2014). Briefly, sequencing reads were

first trimmed using the Trimmomatic program (v0.32) (Lohse et al. 2012). Only read pairs with both ends ≥36 bp were used. Reads were then aligned to the hg19 reference genome using Novoalign software (v3.00.05) (http://www.novocraft.com). Duplicate reads were removed using the MarkDuplicates module of Picard software (v1.126) (http://picard.sourceforge.net/). Additionally, two ends of each pair had to map to the reference genome in complementary directions and reflect a reasonable fragment length (300 ± 100 bp). A local realignment around sites of insertion and deletion was performed using the RealignerTargetCreator and IndelRealigner modules from the Genome Analysis Toolkit (DePristo et al. 2011) (GATK v3.1).

Variant discovery and genotype calling of multi-allelic substitutions, insertions and deletions were performed on all individuals globally using the UnifiedGenotyper and HaplotypeCaller modules from GATK as well as the FreeBayes variant caller (v9.9.2). An Ensemble variant calling pipeline (v0.2.2 http://bcb.io/2013/02/06/an-automated-ensemble-method-for-combining-and-evaluating-genomic-variants-from-multiple-callers/) was then implemented to integrate analysis results from the three callers. Subsequently, the Ensemble pipeline applies a Support Vector Machine learning algorithm to identify an optimal decision boundary based on the variant calling results out of multiple variant callers, to produce a more balanced decision between false and true positives.

Annotation of variants Annotation of each variant locus was made via a custom software pipeline based on public data integrated by a CGR in-house script, including Ensembl, refGene, and UCSC KnownGene databases, the dataset from University of Washington's Exome Sequencing Project (ESP6500) (http://evs.gs.washington.edu/EVS/), dbNSFP (Liu et al. 2011): database of human non-synonymous SNPs and function predictions (https://sites.google.com/site/jpopgen/dbNSFP), the Molecular Signatures Database (MSigDB) (http://www.broadinstitute.org/gsea/msigdb/index.jsp), the National Center for Biotechnology Information (NCBI) Clinically Relevant Sequence Variations (ClinVar) and Single Nucleotide Polymorphism database (dbSNP) databases (Sherry et al. 2001) build 137, the 1000 Genomes Project (Abecasis et al. 2010), the Exome Aggregation Consortium (ExAC) database (http://exac.broadinstitute.org/), and the Human Gene Mutation Database (HGMD) (Stenson et al. 2014).

Filtering of variants Supplementary Table 1 shows the 24 PC-related genes evaluated in the current study. The genes were categorized into: MMR, (other) AD disorders, AR/AD disorders, and hereditary/chronic pancreatitis gene sets. Variants were excluded from further evaluation if they did not pass the quality control (QC) filter in the in-house bioinformatics pipeline or if the variant was reported by only one caller. Since the goal of the study was to investigate high-risk variants, the analyses focused on (very) rare exonic variants. Thus, variants were excluded from further evaluation based on the following criteria: (1) allele frequency >0.1 % in the 1000 Genomes Project (overall or European sample) or ESP6500 (European sample); (2) present in >2 families from an in-house database (CGR, NCI) of >900 cancer-prone families (excluding melanoma-prone or PC families); (3) synonymous, intronic, or untranslated region (UTR) variants.

Classification/validation of variants Variants were classified as frameshift, stop-gain, splicing, in-frame deletion/insertion, or nonsynonymous (NS) substitutions (missense). Frameshift and stop-gain variants were defined as loss-of-function (LOF) variants since they are expected to be protein truncating and thus deleterious. To classify NS substitutions as deleterious, we used an ensemble prediction score (Meta Likelihood ratio) [Meta LRP] that incorporates results from nine in silico algorithms [Sorting Intolerant from Tolerant (SIFT), PolyPhen-2, Genome Evolutionary Rate Profiling (GERP++), Mutation Taster, Mutation Assessor, Functional Analysis Through Hidden Markov Models (FATHMM), Likelihood Ratio Test (LRT), SiPhy, and PhyloP] and allele frequency. This ensemble score achieved the highest discriminative power compared to 18 deleterious scoring methods and also showed low false-positive prediction rate for benign yet rare NS variants (Dong et al. 2015). ExAC, HGMD, Leiden Open Variant Databases (LOVD), Align Grantham Variation and Grantham Deviation (GVGD) (Tavtigian et al. 2006), ClinVar (http://www.ncbi.nlm.nih.gov/clinvar/), and ingenuity variant analysis (IVA) were examined to further categorize individual variants.

LOF variants, in-frame deletions/insertions, and selected missense variants including those with the lowest 10 % quality (after filtering) were technically validated using Sanger sequencing (in Genoa, Italy) or Ampliseq (at CGR). For technical validation using Ampliseq, a targeted, multiplexed PCR primer panel was designed using the Ion AmpliSeq Designer v4.4.4 (Life Technologies, Carlsbad, CA, USA). Sample DNA (30 ng) was amplified using this custom AmpliSeq primer panel (average amplicon size = 244 bp), and sequencing libraries were prepared following the Ion AmpliSeq Library Preparation protocol (Life Technologies), using Ion Xpress Barcode Adapters. Individual sample libraries were pooled, then templated and sequenced on the Ion Torrent Personal Genome Machine (PGM) Sequencer using Ion PGM Hi-Q Chef chemistry. Base calling and alignment were performed using Torrent Suite 4.4. Variant calling was done with GATK and Torrent Variant Caller.

Population controls

Data from 1001 European-American/European population controls from two cohort studies [Cancer Prevention Study (CPS)-II, $n = 224$; Prostate, Lung, Colorectal and Ovarian Screening Trial (PLCO), $n = 378$] and one case–control study [Environment and Genetics in Lung Cancer Etiology (EAGLE), $n = 399$] were available for inclusion in the current study to evaluate genetic burden for the PC-related genes (Wang et al. 2012). The sequencing and bioinformatics analysis methods for the population controls followed the same processes as were used for the PC patients. However, the SeqCAP EZ Human Exome Library v3.0 + UTR (Roche NimbleGen, Madison, WI, USA) was utilized for exome sequence capture. Variant calling for the population controls was done together with that for the entire in-house database (CGR, NCI) of >1200 cancer-prone families.

Statistical analyses

We performed a gene- or gene-set level test of association (i.e. genetic burden) for each of the 23 PC-related genes (excluding *CDKN2A*) and the four gene sets (MMR, pancreatitis, AD disorder, AR/AD disorder genes). We defined a rare variant to have frequency ≤ 0.001. Since standard family-based methods (Chen et al. 2013; Svishcheva et al. 2014) relying on asymptotic distributions of the test statistic are inappropriate given our small sample sizes, we used the following approach. For each gene or set, the test statistic was defined as the number of cases with a rare exonic variant. We then calculated a p value, or the probability of observing at least that many cases with a rare variant under the null hypothesis. Specifically, we created a list of the $2n_u$ haplotypes from the n_u controls. In the unlikely situation when a control had ≥ 2 rare variants in a single gene, each haplotype was assumed to carry at least one rare variant. We then used a 2-step "permutation" (i.e. random assignment) procedure. First, we randomly generated identical-by-descent (IBD) patterns for the familial cases using the rules of Mendelian inheritance. Second, we assigned each of the founder chromosomes in the families to carry a haplotype randomly selected from the list of control haplotypes. After 1000 permutations, the p value was the proportion of permutations where the number of family members carrying a rare exonic variant was at least as large as that observed in the actual data. Since the *CDKN2A*− cases were all unrelated, we also used Fisher's exact test for the gene-level association test for comparison.

Results

After excluding seven PC patients because of insufficient DNA, sample failure, or sample mix-up, 66 PC patients (43 *CDKN2A*+ and 23 *CDKN2A*−) were included in the current analyses. Table 1 shows the number of rare variants (total, LOF and missense deleterious or in-frame deletion variants) in the *CDKN2A*+ and *CDKN2A*− PC patients by gene set. Supplemental Table 4 shows the details for the 35 rare variants (in 15 genes) that were found in the *CDKN2A*+ and *CDKN2A*− PC patients after filtering. No variant was observed in all PC patients in a *CDKN2A*+ family with multiple PC patients. Individual *CDKN2A*+ PC patients had variants in MMR, pancreatitis and AD disorder genes but no variants in AR/AD disorder genes. In contrast, *CDKN2A*− PC patients had variants in each gene set with the most variants in AR/AD genes. Further, LOF variants [$n = 3$ plus a known *BRCA2* frameshift (P153) (Ghiorzo et al. 2012)] were only seen in *CDKN2A*− PC patients (Table 2).

There were 14 missense variants in *CDKN2A*+ PC patients, seven in the MMR genes, four in different AD disorder genes (one each in *APC, PALLD, BRCA1*, *BRCA2*), and three in the pancreatitis gene *CFTR*. Seven of these

Table 1 Number of rare variants (total, LOF, deleterious missense or in-frame deletion) in *CDKN2A*+ and *CDKN2A*− pancreatic cancer patients categorized by gene set

Gene set	Number of rare variants in *CDKN2A*+ PC patients ($n = 14$)			Number of rare variants in *CDKN2A*− PC patients ($n = 21$)		
	Total variants	Loss-of-function variants	Deleterious missense variants	Total variants	Loss-of-function variants	Deleterious missense or in-frame deletion variants
MMR	7	0	3	5	0	3
AD disorder	4	0	1	5	1[a]	1
AR/AD disorder	0	0	0	7	3	3
Hereditary/chronic pancreatitis	3	0	3	4	0	1

LOF loss of function, *PC* pancreatic cancer, *MMR* mismatch repair, *AD* autosomal dominant, *AR* autosomal recessive

[a] Previously reported *BRCA2* frameshift (Ghiorzo et al. 2012)

Table 2 Loss-of-function (LOF) and deleterious variants [classified by Meta likelihood ratio prediction (LRP)] in pancreatic cancer patients

Gene	Chrom	Location	rsID	Ref	Var	AA/protein change	Allele freq[a]	*CDKN2A* carrier	History of smoking: yes/no/unk	Family/patient ID
Loss-of-function [stopgain, frameshift] variants										
ATM	chr11	108,155,142		AG		E1313 fs	0	No	No	M595
ATM	chr11	108,183,151	rs587779852	G	T	E1978*	1.05E−04	No	Unk	AJ7379
BRCA2	chr13	32,914,288	rs80359537	AT		H1932 fs	0	No	Yes	P153
PALB2	chr16	23,646,627	rs180177100	G	A	R414*	0	No	Unk	AJ7379
Deleterious missense or in-frame deletion variants classified by Meta LRP										
MSH2	chr2	47,630,331	rs267607911	A	G	M1 V	0	Yes	No[b]	LUMC6_019
MSH2	chr2	47,637,248	rs145649774	C	G	L128 V	0.000349	Yes	No	P147
MSH2	chr2	47,641,430	rs34136999	C	T	A272 V	0.000581	No	No	P149
MSH6	chr2	48,026,852	rs376220212	G	A	R577H	0.000116	Yes	No	P176
PMS2	chr7	6,013,076		G	A	P848L	0	No	Unk	P1
PMS2	chr7	6,022,617	rs587780046	G	A	T671 M	6.34E−04	No	Yes	P282
CFTR	chr7	117,171,037	rs201958172	G	A	A120T	1.36E−04	Yes	No	M402
CFTR	chr7	117,171,152	rs397508725	G	C	S158T	1.73E−05	Yes	No	P61
CFTR	chr7	117,243,784	rs151048781	G	C	M952I	1.20E−04	No	Yes	P115
CFTR	chr7	117,250,625	rs149279509	A	G	Y1014C	0.001	Yes	Unk	K_1001, K_1002
ATM	chr11	108,122,581		T	G	L542 W	0	No	No	P268
ATM	chr11	108,196,797	rs567060474	G	A	A2274T	0.001	No	Yes	P115
FANCA	chr16	89,858,430		CTT		E377del	0	No	Yes	P282
TP53	chr17	7,576,911	rs145151284	G	C	T312S	8.99E−05	No	No	P268
BRCA1	chr17	41,256,266	rs28897673	T	C	Y105C	0.000116	Yes	No Yes	LUMC19_013 LUMC19_016

Chrom chromosome, *Ref* referent allele, *Var* variant allele, *AA* amino acid, *freq* frequency, *unk* unknown

[a] Maximum allele frequency from public databases (see Supplemental Table 4 for details)

[b] Extensive passive smoking exposure

variants were predicted to be deleterious by summary Meta LRP score (Table 2); three of which were in MMR genes (*MSH2*, *MSH6*), three in the pancreatitis gene *CFTR*, and one in *BRCA1*. For most of these variants, however, algorithms such as CLINVAR, IVA, and Align GVGD clinically classified the variants as uncertain (Supplemental Table 4).

Of 21 variants observed in *CDKN2A*− PC patients, 12 (Table 2) were considered deleterious [three LOF variants (frameshift in *ATM*, stop-gain in *ATM* and *PALB2*), the known *BRCA2* frameshift, one nonframe deletion in *FANCA*, and seven missense variants]. One Swedish PC patient had two stop-gains, both technically validated and classified as pathogenic, one each in *ATM* (p.E1978*) and *PALB2* (p.R414*). The third LOF variant, an *ATM* frameshift (p.E1313 fs), was observed in an Italian patient; this variant was not seen in our in-house familial or population controls or reported in 1000 Genomes, ESP, ExAC, or Kaviar (http://db.systemsbiology.net/kaviar/). *ATM* had the greatest number of variants ($n = 5$), all technically validated and most (4/5) with classification as pathogenic or likely pathogenic (Table 2, Supplemental Table 4). Similar to what was observed in *CDKN2A*+ PC patients, most MMR variants were classified as variants of uncertain significance (VUS) (Supplemental Table 4).

Table 3 shows the gene-level association test for genes with rare variants in ≥1 PC patient. For all PC patients, *PMS2* showed a suggestive association ($p = 0.057$). *CDKN2A*− PC patients had a significant increase for rare variants in *ATM* ($p = 0.006$), *CPA1* ($p = 0.021$), and *PMS2* ($p = 0.038$). Evaluation of the MMR gene set showed a significant association for all ($p = 0.033$) and suggestive association in *CDKN2A*+ PC patients ($p = 0.086$). There was no significant gene-set association for pancreatitis, AD disorder (excluding *CDKN2A*), or AR/AD disorder genes for all, *CDKN2A*+, or *CDKN2A*− PC patients (data not shown); however, *CDKN2A*− PC patients showed a suggestive association for AR/AD disorder genes ($p = 0.051$). Restricting the evaluation to LOF/deleterious variants (based on Meta LRP) produced significant differences only in *ATM* ($p = 0.001$) and *PMS2* ($p = 0.013$) for *CDKN2A*− PC patients.

Table 4 shows the number of rare variants in *CDKN2A*− PC patients by cancer family history for LOF/deleterious variants. *ATM* variants were found only in PC patients

Table 3 Results of gene-level association tests in all, *CDKN2A*+, and *CDKN2A*− PC patients vs 1001 population controls

Gene	Number of controls with rare variant	All PC cases ($n = 66$)			*CDKN2A*+ PC cases ($n = 43$)			*CDKN2A*− PC cases ($n = 23$)		
		Number of PC cases with rare variant	No. families with rare variant	*p* value [all PC cases vs controls]	Number of *CDKN2A*+ PC cases with rare variant	No. families with rare variant	*p* value [*CDKN2A*+ PC cases vs controls]	Number of *CDKN2A*− PC cases with rare variant	*p* value [*CDKN2A*− PC cases vs controls]	*p* value, Fisher's exact test [*CDKN2A*− PC cases vs controls][a]
APC	43	1	1	0.931	1	1	0.823	0	–	–
ATM	51	5	5	0.237	0	0	–	5	**0.006**	**0.005**
BRCA1	26	3	2	0.24	2	1	0.317	1	0.459	0.463
BRCA2	52	4	4	0.414	1	1	0.864	3	0.108	0.116
CASR	7	1	1	0.374	0	0	–	1	0.157	0.167
CFTR	61	5	4	0.384	4	3	0.285	1	0.721	0.725
CPA1	12	2	2	0.211	0	0	–	2	**0.021**	**0.037**
FANCA	51	1	1	0.959	0	0	–	1	0.715	1.000
MLH1	19	1	1	0.687	1	1	0.499	0	–	–
MSH2	19	3	3	0.128	2	2	0.2	1	0.357	0.368
MSH6	17	3	3	0.092	2	2	0.179	1	0.303	0.338
PALB2	12	1	1	0.527	0	0	–	1	0.253	0.257
PALLD	19	1	1	0.694	1	1	0.517	0	–	–
PMS2	30	5	5	0.057	2	2	0.367	3	**0.038**	**0.032**
TP53	7	1	1	0.365	0	0	–	1	0.166	0.167

p-values <0.05 are highlighted in bold

PC pancreatic cancer

[a] Gene-level association tests also computed using Fisher's exact test for comparison purposes

Table 4 Number of LOF or potentially deleterious variants in *CDKN2A*− PC patients by cancer family history (includes cancers in studied PC patient)

Family history of cancer[a]	No. of families by cancer family history	No. LOF/deleterious variants			
		MMR	Pancreatitis	*ATM*	Other
PC ($n = 1$) + melanoma ($n \geq 1$)	5	1	0	0	
PC ($n = 1$) + melanoma ($n \geq 1$) + digestive system[d] ($n \geq 1$)	8	1[c]	1[b]	3	1 [*FANCA*]
PC ($n = 1$) + digestive system ($n \geq 1$)	6	0	0	2	1[b] [*PALB2*] 1[b] [*TP53*]
PC ($n = 1$) + melanoma ($n \geq 1$) + non-digestive system ($n \geq 1$)	3	1	0	0	1 [*BRCA2* − QC]

LOF loss of function, *PC* pancreatic cancer, *No*. number, *MMR* mismatch repair, *QC* quality control

[a] Number of family members with PC, melanoma, digestive system cancer, or non-digestive system cancer

[b] Patient has another rare LOF/deleterious variant in *ATM*

[c] Patient has a rare nonframe deletion in *FANCA*

[d] Includes pancreatic cancer, thus family has ≥2 members with PC

with family histories that included digestive system tumors (5/14 = 36 %). Further, in addition to the Swedish patient with two LOF variants, three other *CDKN2A*− PC patients had multiple rare potentially deleterious variants two of which included *ATM*. Similarly, nine *CDKN2A*+ PC patients had multiple rare potentially deleterious variants since they had known *CDKN2A* mutations in addition to potentially deleterious variants in MMR, pancreatitis, or other AD disorder genes (Table 2; Supplemental Table 4).

Discussion

The risk of pancreatic cancer is increased in some families with *CDKN2A* mutations. However, the reason for this increased risk has yet to be determined. We systematically evaluated 24 (including *CDKN2A*) purported PC-related susceptibility genes in PC patients with and without *CDKN2A* mutations that were ascertained by research groups investigating patients/families with *CDKN2A* mutations. Overall, there was a significant increase in genetic burden for carrying rare variants in MMR genes in all PC patients (*CDKN2A*+ and *CDKN2A*−) compared to population controls. Nine *CDKN2A*+ and four *CDKN2A*− PC patients had rare potentially deleterious variants in multiple genes. For the *CDKN2A*− PC patients, *ATM* showed the strongest association with *ATM* variants only observed in PC patients with a family history that included digestive system tumors. Most of the *ATM* variants were predicted to be pathogenic. Further, three of the four *CDKN2A*− patients with potentially deleterious variants in multiple PC-related genes had a variant that involved *ATM*, including one Swedish patient with previously HGMD-reported pathogenic stop-gains in both *ATM* and *PALB2*. However, having two pathogenic LOF mutations in two high-risk PC-related genes did not result in a substantially earlier age at diagnosis, similar to what has been observed in several familial PC (FPC) patients (Roberts et al. 2015) as well as melanoma patients homozygous for the Dutch *CDKN2A* founder mutation (de Snoo et al. 2008; Gruis et al. 1995).

Family history studies of pancreatic adenocarcinoma suggest that 5–10 % of cases have a strong hereditary basis consistent with other adult-onset cancers such as breast cancer and melanoma (Goldstein and Tucker 2001; Zhen et al. 2015). However, in contrast to breast cancer with *BRCA1/2* and melanoma with *CDKN2A*, mutations in individual PC-related predisposition genes do not account for more than a few percent of FPC patients. Genetic causes of PC show extensive genetic heterogeneity with cancer predisposition genes *ATM*, *BRCA1*, *BRCA2*, *CDKN2A*, *PALB2*, and the MMR genes appearing to account for the largest proportion of the known genetic causes of FPC (Klein 2013; Zhen et al. 2015). A recent whole genome sequencing study of FPC patients confirmed the importance of these major FPC susceptibility genes and also proposed some additional candidate genes (*BUB1B*, *CPA1*, *FANCC*, and *FANCG*) for FPC (Roberts et al. 2015). In the current study, we observed no LOF/deleterious variants in *BUB1B*, *FANCG*, or *FANCC* (data not shown) in either *CDKN2A*+ or *CDKN2A*− PC patients. In contrast, there was a significant gene-level association with *CPA1* in *CDKN2A*− PC patients, but neither of the observed *CPA1* variants was classified as deleterious. Roberts et al. (2015) also analyzed 87 predominantly hereditary cancer genes [supplementary table S3 in Roberts et al. (2015)] in depth for protein truncating variants. Review of these hereditary cancer genes in our sample (excluding the genes systematically evaluated) revealed only one very rare (0.0083 % from Kaviar version

150810-Public) LOF variant [c.2011dupA, p.I671fs] in *FANCI* in a *CDKN2A*+ PC patient.

The importance of cancer family histories in PC patients has yet to be fully investigated. A recent examination of 290 PC probands from the population-based Ontario Pancreas Cancer Study who had been selected based on family history of breast and/or ovarian cancer, PC, or no family history of either, found 11 pathogenic mutations (3.8 %) in an investigation of 13 cancer predisposition genes (*APC, ATM, BRCA1, BRCA2, CDKN2A, MLH1, MSH2, MSH6, PALB2, PMS2, PRSS1, STK11, TP53*) (Grant et al. 2015). Of particular interest, carrier status was significantly associated with a personal or first-degree family history of breast or colorectal cancer but not a family history of PC. Our smaller non-population-based study showed a high proportion of patients with a family history of digestive system tumors carrying a deleterious variant which may be etiologically relevant or reflect ascertainment bias. Future much larger studies of PC patients with detailed family history data will be needed to better understand the cancer histories that impact PC risk.

The samples in this study came from research groups who study familial melanoma so most *CDKN2A*+ PC patients were from melanoma-prone families in ongoing studies in the United States, the Netherlands, Italy, or Sweden. In addition, several *CDKN2A*+ and *CDKN2A*− PC patients were part of a PC case–control study conducted in Genoa, Italy. The remaining *CDKN2A*− PC patients included members of melanoma-prone or PC-prone families. Smoking history was not available for many patients and thus this risk factor could not be evaluated. Interestingly, though, about half of the PC patients with LOF/deleterious variants were nonsmokers (see Table 2, Supplemental Table 4). Further, although of interest, small sample size precluded examination of common variants. In addition, the population controls that were sequenced on the same platform as the cases, necessary for conducting the gene-level association tests, were European (Italian) and European-American. Restriction of the gene- and gene-set-level association tests to Italian and American PC patients showed stronger evidence for an association for the MMR gene-set across all PC patient subsets: all ($p = 0.002$), *CDKN2A*+ PC patients ($p = 0.027$), and *CDKN2A*− PC patients ($p = 0.049$). In addition, gene-level associations were stronger for individual MMR genes and *CFTR* for all [*MSH6* ($p = 0.02$), *PMS2* ($p = 0.028$), and *CFTR* ($p = 0.075$)], for *CDKN2A*+ PC patients [*CFTR* ($p = 0.045$) and *MSH6* ($p = 0.057$)], and for *CDKN2A*− PC patients [*PMS2* ($p = 0.012$)]. Finally, in *CDKN2A*− PC patients, *ATM* ($p = 0.005$) showed similar evidence for association but *CPA1* ($p = 0.198$) was no longer significant.

For most PC patients with LOF/deleterious variants, it was not possible to evaluate additional relatives to determine whether they carried the variant seen in their respective families. However, one Italian PC patient proband (P268) who carried the p.L542W *ATM* variant recently had a brother diagnosed with cancer of the epiglottis at age 60 years. After providing a blood sample and consent for genetic testing, the brother was found to carry the same *ATM* variant as the proband. In addition to these two patients, the cancer family history included the siblings' father who had PC, two first cousins of the siblings with PC, and an aunt with colorectal cancer. The median age at diagnosis of the four PC patients was 58 years (range 56–62 years).

In summary, five research groups from the United States, Italy, The Netherlands, and Sweden that have been studying melanoma, PC, and the *CDKN2A* gene and have some of the largest samples of PC patients with *CDKN2A* mutations contributed to this study. Nevertheless, after combining material from these five research groups, the sample size remained limited with sequencing data available for only 43 PC patients with and 23 PC patients without *CDKN2A* mutations. However, even with this limited sample size, we found a nominally significant gene-set level association for MMR genes in all PC patients with stronger evidence for this association in the Italian and American PC patients, particularly in *CDKN2A*+ PC patients, and for *ATM*, *CPA1*, and *PMS2* in *CDKN2A*− PC patients. Further, numerous PC patients had rare likely deleterious variants in more than one PC-related gene suggesting that a subset of *CDKN2A*+ and *CDKN2A*− PC patients may have an increased risk of pancreatic cancer because of mutations in multiple PC-related susceptibility genes. However, many *CDKN2A*+ PC patients did not have deleterious variants in any of the PC-related genes and therefore other genes, exposures, and/or alternative mechanisms that involve *CDKN2A* likely influence risk of PC in many of these families. Additional research is needed to confirm these findings and to more fully evaluate specific variants/genes that may play a role in pancreatic cancer in patients and families with and without *CDKN2A* mutations.

Acknowledgments We are indebted to the participating families, whose generosity and cooperation have made this study possible. We acknowledge the research nurse contributions to this work that were made by Virginia Pichler (NCI), Deborah Zametkin (NCI), Mary Fraser (NCI), and Anita Zander (Lund). We wish to thank Dr. Chiara Baldo from the Galliera Genetic Bank-Network of Telethon Genetic Biobanks (Project No. GTB12001, Telethon, Italy) for providing patients' lymphoblastoid cell lines. We acknowledge the contribution of members of the NCI DCEG Cancer Sequencing Working Group: Lynn R. Goldin, Mary L. McMaster, Neil E. Caporaso, Bari Ballew, Sharon Savage, Mark H. Greene, Allan Hildesheim, Nan Hu, Jennifer Loud, Phuong Mai, Lisa Mirabello, Lindsay Morton, Dilys Parry, Douglas R. Stewart, Philip R. Taylor, Geoffrey S. Tobias, and Guoqin Yu and members of the NCI DCEG Cancer Genomics Research Laboratory: Sarah Bass, Joseph Boland, Salma Chowdhury, Michael Cullen, Casey Dagnall, Herbert Higson, Sally Larson, Kerry Lashley, Hyo Jung Lee, Wen Luo, Michelle Manning, Jason Mitchell, David

Roberson, Mingyi Wang. We acknowledge the contributions of the Genoa Pancreatic Cancer Study Group: Virginia Andreotti, Claudia Martinuzzi, Luca Mastracci, Federica Grillo, Vincenzo Savarino, Pietro Dulbecco, Stefania Sciallero, Francesco Spagnolo, Virginia Picasso, Franco DeCian, Barbara Pasini, Paola Ogliara, Daniela Turchetti, Elena Sala.

Compliance with ethical standards

Funding This work was supported by the Intramural Research Program of the National Cancer Institute, the National Institutes of Health, the Division of Cancer Epidemiology and Genetics. This work was also supported in part by the Swedish Cancer Society, Kamprad Foundation, Gunnar Nilsson Foundation and the ERC advanced Grant 294576-risk factors cancer; the Swedish Medical Research Council, the Swedish Cancer Society, Radiumhemmets research funds, the Stockholm County Council (ALF-project), Karolinska Institutet Research funds; The Paulsson Trust (Lund); AIRC IG 15460 to PG, Italian Ministry of Health 5 × 1000 to IRCCS AOU San Martino-IST to PG and GBS; the work of NAG and RvD was in part supported by the Dutch Cancer Society (UL 2012-5489).

Conflict of interest The authors declare that they have no conflict of interest.

References

Abecasis GR et al (2010) A map of human genome variation from population-scale sequencing. Nature 467:1061–1073. doi:10.1038/nature09534

Bergman W, Gruis N (1996) Familial melanoma and pancreatic cancer. N Engl J Med 334:471–472

Borg A et al (2000) High frequency of multiple melanomas and breast and pancreas carcinomas in *CDKN2A* mutation-positive melanoma families. J Natl Cancer Inst 92:1260–1266

Chen H, Meigs JB, Dupuis J (2013) Sequence kernel association test for quantitative traits in family samples. Genet Epidemiol 37:196–204. doi:10.1002/gepi.21703

de Snoo FA et al (2008) Increased risk of cancer other than melanoma in CDKN2A founder mutation (p16-Leiden)-positive melanoma families Clinical cancer research. Off J Am Assoc Cancer Res 14:7151–7157. doi:10.1158/1078-0432.CCR-08-0403

DePristo MA et al (2011) A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet 43:491–498. doi:10.1038/ng.806

Dong C, Wei P, Jian X, Gibbs R, Boerwinkle E, Wang K, Liu X (2015) Comparison and integration of deleteriousness prediction methods for nonsynonymous SNVs in whole exome sequencing studies. Hum Mol Genet 24:2125–2137. doi:10.1093/hmg/ddu733

Ghiorzo P et al (2004) INK4/ARF germline alterations in pancreatic cancer patients. Ann Oncol Off J Eur Soc Med Oncol/ESMO 15:70–78

Ghiorzo P et al (2012) Contribution of germline mutations in the BRCA and PALB2 genes to pancreatic cancer in Italy. Fam Cancer 11:41–47. doi:10.1007/s10689-011-9483-5

Goldstein AM (2004) Familial melanoma, pancreatic cancer and germline *CDKN2A* mutations. Hum Mutat 23:630. doi:10.1002/humu.9247

Goldstein AM, Tucker MA (2001) Genetic epidemiology of cutaneous melanoma: a global perspective. Arch Dermatol 137:1493–1496

Goldstein AM et al (1995) Increased risk of pancreatic cancer in melanoma-prone kindreds with p16INK4 mutations. N Engl J Med 333:970–974

Goldstein AM et al (2006) High-risk melanoma susceptibility genes and pancreatic cancer, neural system tumors, and uveal melanoma across. GenoMEL Cancer Res 66:9818–9828. doi:10.1158/0008-5472.CAN-06-0494

Grant RC et al (2015) Prevalence of germline mutations in cancer predisposition genes in patients with pancreatic cancer. Gastroenterology 148:556–564. doi:10.1053/j.gastro.2014.11.042

Gruis NA et al (1995) Homozygotes for CDKN2 (p16) germline mutation in Dutch familial melanoma kindreds. Nat Genet 10:351–353. doi:10.1038/ng0795-351

Helgadottir H et al (2014) High risk of tobacco-related cancers in *CDKN2A* mutation-positive melanoma families. J Med Genet 51:545–552. doi:10.1136/jmedgenet-2014-102320

Klein AP (2013) Identifying people at a high risk of developing pancreatic cancer. Nat Reviews Cancer 13:66–74. doi:10.1038/nrc3420

Liu X, Jian X, Boerwinkle E (2011) dbNSFP: a lightweight database of human nonsynonymous SNPs and their functional predictions. Hum Mutat 32:894–899. doi:10.1002/humu.21517

Lohse M, Bolger AM, Nagel A, Fernie AR, Lunn JE, Stitt M, Usadel B (2012) RobiNA: a user-friendly, integrated software solution for RNA-Seq-based transcriptomics. Nucleic Acids Res 40:W622–W627. doi:10.1093/nar/gks540

Lynch HT et al (2002) Phenotypic variation in eight extended CDKN2A germline mutation familial atypical multiple mole melanoma-pancreatic carcinoma-prone families: the familial atypical mole melanoma-pancreatic carcinoma syndrome. Cancer 94:84–96

Roberts NJ et al (2015) Whole genome sequencing defines the genetic heterogeneity of familial pancreatic cancer. Cancer Discov. doi:10.1158/2159-8290.CD-15-0402

Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM, Sirotkin K (2001) dbSNP: the NCBI database of genetic variation. Nucleic Acids Res 29:308–311

Shi J et al (2014) Rare missense variants in POT1 predispose to familial cutaneous malignant melanoma. Nat Genet 46:482–486. doi:10.1038/ng.2941

Stenson PD, Mort M, Ball EV, Shaw K, Phillips A, Cooper DN (2014) The Human Gene Mutation Database: building a comprehensive mutation repository for clinical and molecular genetics, diagnostic testing and personalized genomic medicine. Hum Genet 133:1–9. doi:10.1007/s00439-013-1358-4

Svishcheva GR, Belonogova NM, Axenovich TI (2014) FFBSKAT: fast family-based sequence kernel association test. PLoS ONE 9:e99407. doi:10.1371/journal.pone.0099407

Tavtigian SV et al (2006) Comprehensive statistical study of 452 BRCA1 missense substitutions with classification of eight recurrent substitutions as neutral. J Med Genet 43:295–305. doi:10.1136/jmg.2005.033878

Vasen HF, Gruis NA, Frants RR, van Der Velden PA, Hille ET, Bergman W (2000) 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 87:809–811. doi:10.1002/1097-0215(20000915)87:6<809:AID-IJC8>3.0.CO;2-U

Wang Z et al (2012) Improved imputation of common and uncommon SNPs with a new reference set. Nat Genet 44:6–7. doi:10.1038/ng.1044

Zhen DB et al (2015) BRCA1, BRCA2, PALB2, and *CDKN2A* mutations in familial pancreatic cancer: a PACGENE study. Genet Med Off J Am Coll Med Genet 17:569–577. doi:10.1038/gim.2014.153