**Lack of association for the reported endocrine pancreatic cancer risk *loci* in the PANDoRA consortium**

Campa D1, Obazee O2, Pastore M1,2, Panzuto F3, Liço V4, Greenhalf W5, Katzke V6, Tavano F7, Costello E5, Corbo V8, Talar-Wojnarowska R9, Strobel O10, Zambon CF11, Neoptolemos JP5, Zerboni G3, Kaaks R6, Key TJ12, Lombardo C13,14, Jamroziak K15, Gioffreda D7, Hackert T10, Khaw KT16, Landi S1, Milanetto AC4, Landoni L17, Lawlor RT8, Bambi F18, Pirozzi F7, Basso D19, Pasquali C4, Capurso G3, Canzian F20.

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

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

3 Digestive and Liver Disease Unit, S. Andrea Hospital, "Sapienza" University of Rome, Rome, Italy.

4 Pancreatic and Digestive Endocrine Surgery, Department of Surgery, Oncology and Gastroenterology -DiSCOG, University of Padova, Padua, Italy.

5 Institute of Translational Medicine, Cancer Research UK Liverpool Cancer Trials Unit, Liverpool, United Kingdom.

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

7 Division of Gastroenterology and Research Laboratory, Department of Surgery, IRCCS Scientific Institute and Regional General Hospital "Casa Sollievo della Sofferenza", San Giovanni Rotondo, Italy.

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

9 Department of Digestive Tract Diseases, Medical University of Lodz, Lodz, Poland.

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

11 Department of Medicine - DIMED, University of Padova, Padua, Italy.

12 Cancer Epidemiology Unit, Nuffield Department of Population Health, University of Oxford, United Kingdom.

13 Division of General and Transplant Surgery, University of Pisa, Pisa, Italy.

14 Dipartimento di Ricerca Traslazionale e delle Nuove Tecnologie in Medicina e Chirurgia, University of Pisa, Pisa, Italy.

15 Department of Hematology, Institute of Hematology and Transfusion Medicine, Warsaw, Poland.

16 University of Cambridge School of Clinical Medicine, Cambridge, United Kingdom.

17 The Pancreas Institute, Department of Surgery, University and Hospital Trust of Verona, Verona, Italy.

18 Blood Transfusion Service, Azienda Ospedaliero Universitaria Meyer, Florence, Italy.

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

20 Genomic Epidemiology Group, German Cancer Research Center (DKFZ), Heidelberg, Germany. f.canzian@dkfz.de.

**Keywords**: PNET, PANDoRA, genetic polymorphisms, association

**Running title:** Lack of replication of PNET risk *loci*

**Financial support:** This work was partially supported by

**Corresponding author:** Federico Canzian

Genomic Epidemiology Group

German Cancer Research Center / Deutsches Krebsforschungszentrum (DKFZ)

Im Neuenheimer Feld 280,

69120 Heidelberg,

Tel. +49-6221-421791

Fax +49-6221-421810

f.canzian@dkfz.de

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

**Word count:** abstract 207, main text 821

**Number of figures and tables:** 2

**Abstract**

**Background** Pancreatic neuroendocrine tumors (PNETs) are rare neoplasms for which very little is known about either environmental or genetic risk factors. Only a handful of association studies have been performed so far, suggesting a small number of risk *loci*.

**Methods** In order to replicate the best findings we have selected 16 Single Nucleotide Polymorphisms (SNPs) suggested in previous studies to be relevant in PNET etiogenesis. We genotyped the selected SNPs (rs16944, rs1052536, rs1059293, rs1136410, rs1143634, rs2069762, rs2236302, rs2387632, rs3212961, rs3734299, rs3803258, rs4962081, rs7234941, rs7243091, rs12957119, rs1800629) in 344 PNETs sporadic cases and 2,721 controls in the context of the PANcreatic Disease ReseArch (PANDoRA) consortium.

**Results** After correction for multiple testing we did not observe for any of the SNPs a statistically significant association with PNET risk. We also used three online bioinformatic tools (HaploReg, RegulomeDB and GTEx) to predict a possible functional role of the SNP but we did not observe any clear indication.

**Conclusions** None of the selected SNPs were convincingly associated with PNET risk in the PANDoRA consortium.

**Impact** We can exclude a major role of the selected polymorphisms in PNET etiology and highlight the need of replication of epidemiologic findings in independent populations, especially in rare diseases such as PNETs.

**Introduction**

Pancreatic neuroendocrine tumors (PNETs) are rare neoplasms but their incidence has greatly increased in the last decades ([1](#_ENREF_1)). In comparison with many other solid tumours very little is known about PNETs risk factors, and only a handful of association studies have been performed to uncover the genetic determinants of the disease ([2-4](#_ENREF_2)). Berkovic and colleagues have performed several studies on inflammation-related genes, such as *ILB1* and *TNFA (*[*2*](#_ENREF_2)*,* [*3*](#_ENREF_3)*)*, while Ter-Minassian and collaborators have selected a more comprehensive approach using a custom array containing almost 1500 SNPs ([4](#_ENREF_4)). Both teams suggest several potentially interesting association, however due to the capricious nature of association studies and to the relatively small sample size of the studies it is of the uttermost importance to validate their findings an in independent population such as the Pancreatic Disease Research (PANDoRA) study ([5](#_ENREF_5)).

**Material and Methods**

The PANDoRA consortium has been extensively described elsewhere ([5](#_ENREF_5)). In this study 344 PNETs cases and 2,721 controls have been genotyped (table 1). Cases were sporadic, i.e. not observed in the context of genetic syndromes associated with PNET, such as multiple endocrine noeplasia (MEN)-1, MEN-2, Von Hippel-Lindau (VHL) or tuberous sclerosis (TSC). Controls were selected in the same geographic areas as the cases. We selected 16 SNPs that represent all the polymorphic variants so far identified as risk *loci* for PNETs (with associations reported with p<0.05) but not replicated yet. In addition to the SNPs reported in this manuscript we have recently performed a study on the *CDKN2A* gene variability in relation to PNET susceptibility and therefore all the variants of this gene are not reported here. Genotyping was performed using the KASPar SNP genotyping system (KBiosciences, Hoddesdon, UK). The order of DNA samples from case and control subjects was randomized on plates in order to ensure that similar numbers of cases and controls were analyzed in each batch. For quality control purpose around 8% of the samples were duplicated and the genotype concordance was checked. Statistical analysis was performed using an unconditional logistic regression setting the more common allele for each polymorphisms as reference and adjusting for age (as continous variable), gender and country of origin. Given that in the original article on *IL1B (*[*3*](#_ENREF_3)*)* an analysis was performed combining the alleles of the two SNPs, we also performed it (using the same criteria as Berkovic). We also used online bioinformatic tools such as HaploReg, (http://www.broadinstitute.org/mammals/haploreg/haploreg.php) RegulomeDB (http://www.regulomedb.org/), and GTEx (http://www.gtexportal.org/home/) in order to explore potential functions of the SNPs.

**Results**

The average call rate of the 16 genotyped SNPs was 96.63% (range 92.06%-99.04%). The concordance rate of the duplicated genotypes was higher than 99%. The genotype distributions at all loci were in Hardy-Weinberg equilibrium in controls, with non-significant chi square values. Table 2 shows the frequencies and distribution of the genotypes, the odds ratios (OR) and 95% confidence intervals (CI) for the association with PNET risk. None of the SNPs showed any statistically significant associations even considering a threshold p value of 0.05 The only potentially interesting effect was a trend for the carriers of the minor allele (A) of the *IL1B*-rs1143634 SNP and a decreased risk of developing PNET (Phomozygous=0.082). The combined diplotype analysis for the *IL1B* gene (as done in the original manuscript by Berkovic et al ([3](#_ENREF_3)) did not reveal any statistically significant associations (data not shown). For *IL1B*-rs1143634 HaploReg showed two possible eQTLs (with expression of gene *CHCHD5*, p=4.1x10-5, and of gene *SLC20A1*, p=0.0014), RegulomeDB showed a score of 5 (minimal binding evidence) and GTEx showed no statistically significant results.

**Discussion and Conclusion**

Due to the rarity of PNET, there have been only a small number of studies investigating the genetic susceptibility to this disease. The more potentially interesting findings have been polymorphic variants in genes involved in inflammatory response, in cell cycle control and in DNA repair mechanisms. This is one of the largest study on PNETs to date and our sample size was by far larger than those used in the original studies, giving us more than 95% statistical power to detect the previously reported associations. The *IL1B*-rs1143634 variant is a synonymous, possibly functional, SNP that has been widely studied in relation to a variety of human diseases and conditions, making it an attractive candidate for PNET risk. Our results, however, suggest, at best, only a minor influence of the variant in the etiology of the disease. In addition, we found the minor allele to be associated with decreased risk while in the original publication it was the opposite. The three selected bioinformatic tools did not reveal a clear-cut indication about the functional effect of the SNP. In conclusion in this study we could exclude a major role of the selected polymorphisms in PNET etiology and highlight, on one hand the importance of finding genetic markers for the disease, maybe through a genome-wide association study (GWAS) approach, and on the other hand the need of replication of epidemiologic findings in independent populations.

**References**

1. Viudez A, De Jesus-Acosta A, Carvalho FL, Vera R, Martin-Algarra S, Ramirez N. Pancreatic neuroendocrine tumors: Challenges in an underestimated disease. Critical reviews in oncology/hematology. 2016:193-206.

2. Berkovic MC, Jokic M, Marout J, Radosevic S, Zjacic-Rotkvic V, Kapitanovic S. IL-2 -330 T/G SNP and serum values-potential new tumor markers in neuroendocrine tumors of the gastrointestinal tract and pancreas (GEP-NETs). Journal of molecular medicine (Berlin, Germany). 2010;88:423-9.

3. Cigrovski Berkovic M, Catela Ivkovic T, Marout J, Zjacic-Rotkvic V, Kapitanovic S. Interleukin 1beta gene single-nucleotide polymorphisms and susceptibility to pancreatic neuroendocrine tumors. DNA and cell biology. 2012;31:531-6.

4. Ter-Minassian M, Wang Z, Asomaning K, Wu MC, Liu CY, Paulus JK, et al. Genetic associations with sporadic neuroendocrine tumor risk. Carcinogenesis. 2011;32:1216-22.

5. Campa D, Rizzato C, Capurso G, Giese N, Funel N, Greenhalf W, et al. Genetic susceptibility to pancreatic cancer and its functional characterisation: the PANcreatic Disease ReseArch (PANDoRA) consortium. Dig Liver Dis. 2013;45:95-9.

**Table 1.** Study population.

|  |  |  |
| --- | --- | --- |
|  | **Cases** | **Controls** |
| **Region** |  |  |
| Germany | 32 | 768 |
| Northern Italy | 159 | 520 |
| Central Italy | 68 | 559 |
| Southern Italy | 13 | 509 |
| Poland | 19 | 189 |
| United Kingdom | 53 | 176 |
| Total | 344 | 2721 |
|  |  |  |
| **Sex** |  |  |
| Male | 174 | 1455 |
| Female | 167 | 1239 |
|  |  |  |
| **Median age** | 59 | 59 |
| **(interquartile range)** | (48-68) | (47-68) |

**Table 2**. Associations between endocrine pancreatic cancer risk and the selected SNPs.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **SNP** | **Gene** | **Alleles****M/ma** | **Cases/Controlsb** | **Mm *vs* MMc** | **Phet** | **mm *vs* MMc** | **Phom** | **Original****publication** |
| rs2069762 | *IL2* | A/C | 138/1128 | 148/1099 | 45/285 | 1.01 (0.77-1.34)d | 0.913 | 1.31 (0.89-1.94)d | 0.169 | ([2](#_ENREF_2)) |
| rs16944 | *IL1B* | G/A | 148/1078 | 142/1157 | 47/279 | 0.82(0.63-1.07) | 0.155 | 1.22 (0.82-1.81) | 0.321 | ([3](#_ENREF_3)) |
| rs1143634 | *IL1B* | G/A | 208/1551 | 120/978 | 12/155 | 0.87(0.68-1.12) | 0.289 | 0.58 (0.31-1.07) | 0.082 | ([3](#_ENREF_3)) |
| rs1059293 | *IFNGR2* | C/T | 114/846 | 170/1323 | 54/492 | 0.94 (0.72-1.22) | 0.639 | 0.83 (0.58-1.19) | 0.303 | ([4](#_ENREF_4)) |
| rs1136410 | *ADPRT* | A/G | 246/1753 | 78/611 | 9/63 | 0.86 (0.65-1.15) | 0.311 | 1.03 (0.50-2.15) | 0.934 | ([4](#_ENREF_4)) |
| rs1052536 | *LIG3* | C/T | 114/799 | 151/1230 | 69/459 | 0.87 (0.65-1.16) | 0.330 | 1.12 (0 .78 -1.59) | 0.546 | ([4](#_ENREF_4)) |
| rs2236302 | *MMP14* | C/G | 274/1980 | 62/492 | 4/41 | 0.89 (0.65-1.23) | 0.503 | 0.46 (0.11-1.93) | 0.287 | ([4](#_ENREF_4)) |
| rs2387632 | *VEGFR1* | C/T | 146/1133 | 155/1130 | 34/335 | 1.00 (0.78-1.29) | 0.974 | 0.75 (0.50-1.13) | 0.171 | ([4](#_ENREF_4)) |
| rs3212961 | *ERCC1* | G/T | 244/1368 | 79/410 | 7/38 | 0.99 (0.72-1.36) | 0.938 | 0.98 (0.41-2.36) | 0.967 | ([4](#_ENREF_4)) |
| rs3734299 | *PERP* | T/C | 158/1035 | 144/1117 | 39/348 | 0.80 (0.61 1.04) | 0.104 | 0.75 (0.50-1.12) | 0.169 | ([4](#_ENREF_4)) |
| rs3803258 | *SLC10A2* | T/C | 241/1738 | 89/711 | 7/66 | 0.99 (0.76-1.30) | 0.953 | 0.96 (0.42-2.18) | 0.425 | ([4](#_ENREF_4)) |
| rs4962081 | *TSC1* | C/A | 293/2335 | 49/249 | 2/22 | 1.11 (0.80-1.55) | 0.515 | 0.68 (0.15-3.03) | 0.921 | ([4](#_ENREF_4)) |
| rs7234941 | *BCL2* | C/T | 240/1932 | 87/696 | 10/59 | 0.96 (0.73-1.25) | 0.760 | 1.34 (0.67-2.72) | 0.408 | ([4](#_ENREF_4)) |
| rs7243091 | *BCL2* | G/A | 212/1664 | 110/857 | 15/130 | 0.96 (0.74-1.24) | 0.763 | 0.87 (0.49-1.53) | 0.625 | ([4](#_ENREF_4)) |
| rs12957119 | *BCL2* | A/C | 227/1799 | 86/698 | 18/94 | 0.92 (0.70-1.21) | 0.535 | 1.38 (0.80-2.39) | 0.253 | ([4](#_ENREF_4)) |
| rs1800629 | *TNFA* | G/A | 261/2280 | 66/558 | 6/42 | 0.96 (0.70-1.32) | 0.819 | 0.61 (0.18-2.00) | 0.413 | ([4](#_ENREF_4)) |

a M = major allele; m = minor allele.

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

c Mm *vs* MM *=*heterozygous carriers *vs* common homozygous; mm *vs* MM = rare homozygous *vs* common homozygous (in both cases using a co-dominant model)

d Odds ratio (95% confidence interval). All analyses were adjusted for age at diagnosis/age at recruitment, gender and country of origin.