Evaluation of Association of HNF1B Variants with Diverse Cancers: Collaborative Analysis of Data from 19 Genome-Wide Association Studies


Published Version doi:10.1371/journal.pone.0010858

Citable link http://nrs.harvard.edu/urn-3:HUL.InstRepos:4621713

Terms of Use This article was downloaded from Harvard University’s DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA
Evaluation of Association of \textit{HNF1B} Variants with Diverse Cancers: Collaborative Analysis of Data from 19 Genome-Wide Association Studies

Katherine S. Elliott$^{1}$, Eleftheria Zeggini$^{1,2}$, Mark I. McCarthy$^{1,3}$, Julius Gudmundsson$^{4}$, Patrick Sulem$^{4}$, Simon N. Stacey$^{4}$, Steinunn Thorlacus$^{4}$, Laufey Amundadottir$^{5}$, Henrik Grönberg$^{6}$, Jianfeng Xu$^{7,8}$, Valerie Gaborieau$^{9}$, Rosalind A. Eeles$^{10}$, David E. Neal$^{11}$, Jenny L. Donovan$^{12}$, Freddie C. Hamdy$^{13}$, Kenneth Muir$^{14}$, Shih-Jen Hwang$^{15}$, Margaret R. Spitz$^{16}$, Brent Zanke$^{17,18}$, Luis Carvajal-Carmona$^{1}$, Kevin M. Brown$^{19}$, Australian Melanoma Family Study Investigators$^{a}$, Nicholas K. Hayward$^{20}$, Stuart Macgregor$^{20}$, Ian P. M. Tomlinson$^{1}$, Mathieu Lemire$^{17}$, Christopher I. Amos$^{16}$, Joanne M. Murabito$^{21}$, William B. Isaacs$^{22}$, Douglas F. Easton$^{23}$, Paul Brennan$^{9}$, The PanScan Consortium$^{9}$, Rosa B. Barkdottir$^{24,25}$, Daniel F. Gudbjartsson$^{9}$, Thorunn Rafnar$^{4}$, David J. Hunter$^{26}$, Stephen J. Chanock$^{5}$, Kari Stefansson$^{4}$, John P. A. Ioannidis$^{27,28}$

$^{1}$Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, United Kingdom, $^{2}$Applied Statistical Genetics, Wellcome Trust Sanger Institute, University of Cambridge, Cambridge, United Kingdom, $^{3}$Oxford Centre for Diabetes, Endocrinology and Metabolism, University of Oxford, Oxford, United Kingdom, $^{4}$deCODE Genetics, Reykjavik, Iceland, $^{5}$Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, Maryland, United States of America, $^{6}$Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden, $^{7}$Center for Cancer Genomics, Wake Forest University School of Medicine, Winston-Salem, North Carolina, United States of America, $^{8}$Center for Human Genomics, Wake Forest University School of Medicine, Winston-Salem, North Carolina, United States of America, $^{9}$Genetic Epidemiology Group, International Agency for Research on Cancer (IARC), Lyon, France, $^{10}$Oncogenetics Team, The Institute of Cancer Research, Sutton, United Kingdom, $^{11}$Department of Oncology, University of Cambridge, Cambridge, United Kingdom, $^{12}$Department of Social Medicine, University of Bristol, Bristol, United Kingdom, $^{13}$Nuffield Department of Surgery, University of Oxford, Oxford, United Kingdom, $^{14}$Health Sciences Research Institute, University of Warwick, Coventry, United Kingdom, $^{15}$Framingham Study, National Heart, Lung, and Blood Institute, Bethesda, Maryland, United States of America, $^{16}$Department of Epidemiology, M.D. Anderson Cancer Center, Houston, Texas, United States of America, $^{17}$Ontario Institute for Cancer Research, MaRS Centre, Toronto, Ontario, Canada, $^{18}$Ottawa Health Research Institute, University of Ottawa, Ottawa, Ontario, Canada, $^{19}$Integrated Cancer Genomics Division, The Translational Genomics Research Institute, Phoenix, Arizona, United States of America, $^{20}$Queensland Institute of Medical Research, Royal Brisbane Hospital, Brisbane, Queensland, Australia, $^{21}$Section of General Internal Medicine, Boston University School of Medicine, Boston, Massachusetts, United States of America, $^{22}$The Brady Urological Institute, Johns Hopkins Medical Institutions, Baltimore, Maryland, United States of America, $^{23}$Cancer Research UK Genetic Epidemiology Unit, University of Cambridge, Cambridge, United Kingdom, $^{24}$Department of Pathology, Landspitali-University Hospital of Iceland, Reykjavik, Iceland, $^{25}$Faculty of Medicine, University of Iceland, Reykjavik, Iceland, $^{26}$Department of Epidemiology, Harvard School of Public Health, Boston, Massachusetts, United States of America, $^{27}$Department of Hygiene and Epidemiology, University of Ioannina School of Medicine and Biomedical Research Institute, Foundation for Research and Technology-Hellas, Ioannina, Greece, $^{28}$Center for Genetic Epidemiology and Modelling, Tufts University School of Medicine, Boston, Massachusetts, United States of America

Abstract

\textbf{Background:} Genome-wide association studies have found type 2 diabetes-associated variants in the \textit{HNF1B} gene to exhibit reciprocal associations with prostate cancer risk. We aimed to identify whether these variants may have an effect on prostate cancer risk in general versus a specific effect on prostate cancer only.

\textbf{Methodology/Principal Findings:} In a collaborative analysis, we collected data from GWAS of cancer phenotypes for the frequently reported variants of \textit{HNF1B}, rs4430796 and rs7501939, which are in linkage disequilibrium ($r^2 = 0.76$, HapMap CEU). Overall, the analysis included 16 datasets on rs4430796 with 19,640 cancer cases and 21,929 controls; and 21 datasets on rs7501939 with 26,923 cases and 49,085 controls. Malignancies other than prostate cancer included colorectal, breast, lung and pancreatic cancers, and melanoma. Meta-analysis showed large between-dataset heterogeneity that was driven by different effects in prostate cancer and other cancers. The per-T2D-risk-allele odds ratios (95% confidence intervals) for rs4430796 were 0.79 (0.76, 0.83) per G allele for prostate cancer ($p<10^{-15}$ for both); and 1.03 (0.99, 1.07) for all other cancers. Similarly for rs7501939 the per-T2D-risk-allele odds ratios (95% confidence intervals) were 0.80 (0.77, 0.83) per T allele for prostate cancer ($p<10^{-15}$ for both); and 1.00 (0.97, 1.04) for all other cancers. No malignancy other than prostate cancer had a nominally statistically significant association.

\textbf{Conclusions/Significance:} The examined \textit{HNF1B} variants have a highly specific effect on prostate cancer risk with no apparent association with any of the other studied cancer types.


\textbf{Editor:} Florian Kronenberg, Innsbruck Medical University, Austria

\textbf{Received} April 8, 2010; \textbf{Accepted} April 28, 2010; \textbf{Published} May 28, 2010

\textbf{Copyright:} © 2010 Elliott et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.
**Introduction**

A large number of epidemiological studies have suggested correlations between type 2 diabetes (T2D) and various cancers,[1,2,3] Most evidence suggests an inverse correlation between T2D and prostate cancer[4,5,6] although not all studies agree on this[7]. Several studies also suggest positive correlations between other cancers and T2D[1,2,3]. It is unclear whether these correlations, if true, represent casual relationships and whether they may also reflect some shared genetic background. Recently, with the advent of genome-wide association studies (GWAS), a large number of genetic variants have been identified that confer susceptibility to T2D or specific types of cancer[8]. An interesting observation has been that specific variants in the HNF1B gene (formerly TCF2) have been demonstrated to be associated both with the risk of prostate cancer[9,10,11] and the risk of T2D[9,1,2] with the effects being in the opposite direction for these two phenotypes.

HNF1B was previously known to be mutated in individuals with maturity-onset diabetes of the young type 5 (MODY 5)[13], but a biological explanation of the impact of the identified common variation on T2D and prostate cancer risk remains elusive. The identified genetic effects are small in magnitude even for prostate cancer and T2D, representing odds ratios [ORs] per allele in the range of 1.2 [9,11] and 0.9 [9,12], respectively. Therefore, small effects for other cancer types would not be readily detectable, unless very large studies were performed or data were combined from several studies.

A definitive answer on whether HNF1B variants modulate also the risk of other malignancies, or show specificity for prostate cancer, requires large sample sizes. Here we present the results of a large collaborative meta-analysis of HNF1B, rs4430796 and rs7501939, which have the most consistent associations with both prostate cancer and T2D. Relevant data were collected on the two variants from GWAS on cancer phenotypes in Caucasian populations in order to examine whether they have an effect on cancer risk in general, on few specific cancers, or only on prostate cancer.

**Results**

**Database of contributed information**

All the originally contacted investigators of cancer-related GWA studies agreed to participate in this collaborative analysis, with the exception of the investigators of 3 GWA studies [14,15,16] (1 on breast cancer, 1 on colorectal cancer and 1 on neuroblastoma), 1 of which had no data on the requested variants, as they had used an Affymetrix platform[15]. Investigators who agreed to participate in the collaborative analysis contributed data on 13 datasets for rs4430796 and 19 datasets for rs7501939 [11,17,18,19,20,21,22,23,24,25,26,27,28,29,30,31,32,33]. For 5 datasets, data were available only for the latter polymorphism either because the polymorphism was not available on the platform used or the SNP failed quality control criteria.

The contributing teams and datasets are shown in Table 1 with data on the number of cases and controls for each polymorphism and for each type of cancer. Datasets from the Framingham cohort contained imputed data for both polymorphisms since an Affymetrix platform had been used, rs4430796 data from the M.D. Anderson Cancer Center was used since this SNP had not been directly genotyped, and melanoma data from AMFS and Q-MEGA contained counts from pooling experiments, otherwise all other datasets had direct genotyping on individual participants. Detailed demographic and other characteristics of the study populations can be found in the respective primary publications of these GWA studies [14,15,16,17,18,20,21,22,23,24,25,26,27,28,29,30,31,32,34]. Overall, the collaborative analysis included data on rs4430796 for 19,640 cancer cases and 21,929 controls; for prostate cancer there were 11,145 cases and 9,650 controls, while for all other cancers there were 8,495 cases and 12,279 controls. The collected data on rs7501939 included 26,923 cases and 49,085 controls; for prostate cancer there were 12,898 cases and 40,371 controls, while for the other cancers there were 14,025 cases and 43,893 controls. Malignancies other than prostate cancer in these datasets included colorectal, breast, lung and pancreatic cancers, and melanoma (Table 1). deCODE contributed data on 4 different cancers and had a common population control group for all 4 of them. The Framingham Heart Study (FHS) contributed data on 4 different cancers and had a common population control group [subjects ≥65 years at the last contact who are not nuclear family member of the cancer cases] for all 4 studies with the exception of prostate and breast cancer which used male and female only controls respectively. The common control groups for deCODE and FHS are only counted once in the total sample sizes above.

The meta-analysis of all datasets (Table 2, Figure 1) showed a per T2D risk allele association with both rs4430796 (G allele OR 0.91 [95% CI: 0.88, 0.94] p = 3 × 10\(^{-16}\)) and rs7501939 (T allele
**Table 1.** Characteristics of datasets included in the collaborative meta-analysis.

<table>
<thead>
<tr>
<th>Study Centre</th>
<th>Cancer</th>
<th>Genotyping platform(s)</th>
<th>rs4430796 cases</th>
<th>rs4430796 controls</th>
<th>rs7501939 cases</th>
<th>rs7501939 controls</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>ARCTIC</em> colorectal[23]</td>
<td>Sequenom homogenous MassExtend (in house)</td>
<td>1,079</td>
<td>1,089</td>
<td>1,075</td>
<td>1,087</td>
<td></td>
</tr>
<tr>
<td><em>AMFS</em> melanoma[17,24]</td>
<td>Illumina 550K (pooled)</td>
<td>490 $^p$</td>
<td>427 $^p$</td>
<td>490 $^p$</td>
<td>427 $^p$</td>
<td></td>
</tr>
<tr>
<td>Cambridge breast[33]</td>
<td>Perlegen</td>
<td>387</td>
<td>363</td>
<td>387</td>
<td>363</td>
<td></td>
</tr>
<tr>
<td><em>CAPS</em> prostate[26]</td>
<td>Sequenom (in house)</td>
<td>2,874</td>
<td>1,708</td>
<td>2,865</td>
<td>1,707</td>
<td></td>
</tr>
<tr>
<td><em>CORGI</em> colorectal[27]</td>
<td>Illumina 550K</td>
<td>n/a</td>
<td>n/a</td>
<td>900</td>
<td>908</td>
<td></td>
</tr>
<tr>
<td>deCODE breast[28]</td>
<td>Illumina 300K</td>
<td>n/a</td>
<td>n/a</td>
<td>1,815</td>
<td>30,742</td>
<td></td>
</tr>
<tr>
<td>deCODE colorectal[29]</td>
<td>Illumina 300K</td>
<td>n/a</td>
<td>n/a</td>
<td>988</td>
<td>30,742</td>
<td></td>
</tr>
<tr>
<td>deCODE lung[29,30]</td>
<td>Illumina 300K</td>
<td>n/a</td>
<td>n/a</td>
<td>651</td>
<td>30,742</td>
<td></td>
</tr>
<tr>
<td>deCODE prostate[9,31,32]</td>
<td>Illumina 300K</td>
<td>n/a</td>
<td>n/a</td>
<td>1,619</td>
<td>30,742</td>
<td></td>
</tr>
<tr>
<td><em>FHS</em> breast[34]</td>
<td>Affymetrix 500K and MIPS 50K combined</td>
<td>182 $^t$</td>
<td>852 $^t$</td>
<td>182 $^t$</td>
<td>852 $^t$</td>
<td></td>
</tr>
<tr>
<td><em>FHS</em> colorectal[34]</td>
<td>Affymetrix 500K and MIPS 50K combined</td>
<td>108 $^t$</td>
<td>1,498 $^t$</td>
<td>108 $^t$</td>
<td>1,498 $^t$</td>
<td></td>
</tr>
<tr>
<td><em>FHS</em> lung[34]</td>
<td>Affymetrix 500K and MIPS 50K combined</td>
<td>90 $^t$</td>
<td>1,498 $^t$</td>
<td>90 $^t$</td>
<td>1,498 $^t$</td>
<td></td>
</tr>
<tr>
<td><em>FHS</em> prostate[34]</td>
<td>Affymetrix 500K and MIPS 50K combined</td>
<td>190 $^t$</td>
<td>646 $^t$</td>
<td>190 $^t$</td>
<td>646 $^t$</td>
<td></td>
</tr>
<tr>
<td><em>IARC</em> lung[20,21]</td>
<td>Illumina 300K</td>
<td>641</td>
<td>2,435</td>
<td>1,797</td>
<td>2,378</td>
<td></td>
</tr>
<tr>
<td><em>JHH</em> prostate[26]</td>
<td>Sequenom (in house)</td>
<td>1,512</td>
<td>478</td>
<td>1,521</td>
<td>479</td>
<td></td>
</tr>
<tr>
<td><em>MDACC</em> lung[22]</td>
<td>Illumina 317K</td>
<td>1,152 $^t$</td>
<td>1,137 $^t$</td>
<td>1,152</td>
<td>1,137</td>
<td></td>
</tr>
<tr>
<td><em>PANSCAN</em> pancreatic Stage 1[19,48]</td>
<td>Illumina 550K and 610K</td>
<td>1,754</td>
<td>1,796</td>
<td>1,757</td>
<td>1,796</td>
<td></td>
</tr>
<tr>
<td><em>PANSCAN</em> pancreatic Stage 2[19,48]</td>
<td>Illumina 550K and 610K</td>
<td>1,748</td>
<td>1,818</td>
<td>1,769</td>
<td>1,841</td>
<td></td>
</tr>
<tr>
<td><em>Q-MEGA</em> melanoma[24]</td>
<td>Illumina 550K (pooled)</td>
<td>864 $^p$</td>
<td>864 $^p$</td>
<td>864 $^p$</td>
<td>864 $^p$</td>
<td></td>
</tr>
<tr>
<td><em>UKGPCS</em> prostate[18]</td>
<td>Illumina 550K</td>
<td>1,609</td>
<td>1,797</td>
<td>1,834</td>
<td>1,867</td>
<td></td>
</tr>
</tbody>
</table>

Unless otherwise indicated all data is from direct genotyping. *ARCTIC* (Assessment of Risk for Colorectal Tumors in Canada), AMFS (Australian Melanoma Family Study), CGEMS (Cancer Genetics Markers of Susceptibility), CAPS (Cancer of the Prostate in Sweden), CORGI (Colorectal Tumour Gene Identification), FHS (Framingham Heart Study), IARC (International Agency for Research on Cancer), JHH (Johns Hopkins Hospital), MDACC (M.D. Anderson Cancer Center, Texas), PANSCAN (Pancreatic Cancer Cohort Consortium), Q-MEGA (Queensland study of Melanoma: Environment and Genetic Associations), UKGPCS (UK Genetic Prostate Cancer Study). n/a: no available data; $^t$: imputed; $^p$: pooled.

doi:10.1371/journal.pone.0010858.t001

OR 0.91 [95% CI: 0.88, 0.94] $p = 5 \times 10^{-10}$ according to fixed effects calculations, while by random effects calculations there was no significant $OR \neq 0.94$ [95% CI: 0.88, 1.00] $p = 0.033$ for rs4430796 and 0.93 [95% CI: 0.86, 1.01] $p = 0.07$ for rs7501939. The reason for this diversity is that there was very large between-study heterogeneity in the effect sizes ($I^2$ of 82% [95% CI: 73–89%] and 80% [95% CI: 70–86%], respectively, for the two polymorphisms; Q-test p-value <0.001 for both polymorphisms), and this makes the fixed effects calculations less reliable. Results were qualitatively similar when we increased the variance in effect sizes by random effects, and this makes the fixed effects calculations less reliable. Results were consistent across studies except for the Framingham study where the number of prostate cancer cases was more limited.

Conversely, the results for all other cancers suggested no significant association and results were consistent across studies. The summary OR was 1.03 and 1.00 for the two polymorphisms respectively ($p = 0.14$ and 0.81 by fixed effects) and the 95% CIs excluded ORs deviating more than 7% from the null (OR = 1.00) for rs4430796 and more than 4% from the null for rs7501939 (Table 2). The Q-test p-value was 0.99 and 0.45 for the two polymorphisms respectively and random effects estimates were thus identical to fixed effects estimates.

There was also no convincing evidence for an association between either of the two polymorphisms and any of the other cancers (besides prostate cancer), when each cancer type was evaluated separately. Point estimates were in the opposite direction and there was no nominally statistically significant association (Table 2). The heterogeneity between the prostate cancer and other cancers’ effect estimates was beyond chance ($p<0.05$) for both polymorphisms.
Discussion

The current collaborative analysis documents that both rs4430796 and rs7501939 have robust support for association with prostate cancer, while we did not observe any convincing evidence for an association of any of the other cancers examined with either polymorphism. When data from all other cancers, excluding prostate cancer, were combined the summary effects had 95% CIs that excluded even subtle associations. Apart from prostate cancer, when other datasets for each individual cancer type was combined, the 95% CIs consistently excluded associations with modest effects. This would suggest that the effects mediated by these polymorphisms are specific to T2D and prostate cancer and they do not involve any other cancer types.

The HNF1B gene encodes a transcription factor and it was initially identified as a MODY gene[13]. Subsequent studies have suggested that mutations in this gene may also be related to renal disease[35] and chromophobe renal cell carcinoma[36]. No GWAS evaluating kidney cancer were included in our analysis, on fixed effects calculations. When point estimates or confidence intervals differ by over 1% in random effects calculations, random effects results are mentioned in the Table 2.

### Table 2. Summary of results for association between rs4430796 and rs7501939 and diverse cancer types.

<table>
<thead>
<tr>
<th>Cancer type</th>
<th>rs4430796 Studies (cases, controls) OR (95% CI) I² (95% CI)</th>
<th>rs4430796 Studies (cases, controls) OR (95% CI) I² (95% CI)</th>
<th>rs4430796 Studies (cases, controls) OR (95% CI) I² (95% CI)</th>
<th>rs7501939 Studies (cases, controls) OR (95% CI) I² (95% CI)</th>
<th>rs7501939 Studies (cases, controls) OR (95% CI) I² (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>All cancers</td>
<td>16 (19,640, 21,929)* 0.91 (0.88, 0.94) 82 (73, 89)</td>
<td>21 (26,923, 49,085)* 0.92 (0.90, 0.95) 80 (70, 86)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prostate</td>
<td>5 (11,145, 9,650) 0.79 (0.76, 0.83)</td>
<td>6 (12,898, 40,371) 0.80 (0.77, 0.83)</td>
<td>56 (0, 82)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>All Others</td>
<td>11 (8,495, 12,279)* 1.03 (0.99, 1.07)</td>
<td>0 (0, 60)</td>
<td>15 (14,025, 43,893)* 1.00 (0.97, 1.04)</td>
<td>0 (0, 54)</td>
<td></td>
</tr>
<tr>
<td>Breast</td>
<td>2 (569, 1,215) 1.00 (0.84, 1.20)</td>
<td>n/a</td>
<td>3 (2,384, 31,957) 0.97 (0.91, 1.04)</td>
<td>0 (0, 90)</td>
<td></td>
</tr>
<tr>
<td>Lung</td>
<td>3 (1,883, 5,070) 1.05 (0.98, 1.13)</td>
<td>0 (0, 90)</td>
<td>4 (3,690, 35,755) 1.03 (0.96, 1.10)</td>
<td>0 (0, 85)</td>
<td></td>
</tr>
<tr>
<td>Colorectal</td>
<td>2 (1,187, 2,587) 1.01 (0.90, 1.14)</td>
<td>n/a</td>
<td>4 (3,071, 34,235) 1.01 (0.94, 1.08)</td>
<td>0 (0, 85)</td>
<td></td>
</tr>
<tr>
<td>Melanoma</td>
<td>2 (1,354, 1,291) 0.98 (0.87, 1.01)</td>
<td>n/a</td>
<td>2 (1,354, 1,291) 1.01 (0.90, 1.13)</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td>Pancreatic</td>
<td>2 (3,502, 3,614) 1.04 (0.98, 1.11)</td>
<td>n/a</td>
<td>2 (3,526, 3,637) 1.03 (0.97, 1.10)</td>
<td>n/a</td>
<td></td>
</tr>
</tbody>
</table>

OR: odds ratio, CI: confidence interval, n/a: not applicable (heterogeneity I² confidence intervals are not calculated when there are only 2 studies). Odds ratios are based on fixed effects calculations. When point estimates or confidence intervals differ by over 1% in random effects calculations, random effects results are mentioned in the text. * the common control groups of deCODE and FHS are counted only once.

http://dx.doi.org/10.1371/journal.pone.0010858 (600)

In conclusion, while the two examined HNF1B variants conclusively have pleiotropic effects on both T2D and prostate cancer, the pleiotropy apparently does not extend to other cancer types. Genetic effects may offer a way to dissect comorbidity between specific cancers and metabolic phenotypes. Besides HNF1B, other gene loci have started appearing where variants are identified that modulate susceptibility to both T2D and some malignancy, e.g. prostate cancer for the JAZF1 locus gene[11,42] and melanoma for the CDKN2A locus[43], although different, unlinked variants are implicated in the susceptibility to the malignancy and T2D, respectively. The elucidation of correlated pleiotropic effects on diverse phenotypes will require very large studies, given the generally subtle effects involved. Collaborative efforts between multiple teams, as in the current study, may offer a suitable approach to answer such questions.

### Methods

**Eligible GWA investigations and data**

We used the NHGRI catalogue of published GWA studies[44], a comprehensive database of GWA investigations to identify GWA studies on cancer phenotypes published as of May 20, 2008. We also performed additional PubMed searches to identify whether any additional GWA studies on cancer phenotypes had been published until then. We focused on solid cancers, excluding hematologic malignancies. Given that these GWAS did not include any studies on pancreatic cancer (of special interest, given the association with T2D), we also identified GWAS on pancreatic cancer that had not been published by that time, so as to ensure their inclusion.
Figure 1. Association of rs4430796 and rs7501939 with diverse cancer types. Panel A shows results for rs4430796 and panel B shows results for rs7501939. Each study is shown by its odds ratio and 95% confidence intervals. Prostate cancer studies appear on the top and other cancer studies follow in alphabetical order. For the abbreviations of the names of the studies see Table 1. The summary diamond at the bottom corresponds to the fixed effects summary. Weight indicates the relative proportion of the total evidence found in each study (the weight is inversely proportional to the variance).

doi:10.1371/journal.pone.0010858.g001
We communicated with the corresponding and principal investigators of all of these studies to request their participation in the collaborative meta-analysis. The investigators of these studies were asked to contribute relevant data on genotype frequencies in cancer cases and non-cancer controls for the HNF1B variants, rs4430796 and rs7501939 respectively. The risk alleles for prostate cancer are A and C for rs4430796 and rs7501939 respectively. The risk alleles for T2D are G and T for rs4430796 and rs7501939 respectively. The two SNPs have modestly high LD in Caucasians, but low LD in Africans (\(r^2 = 0.77\) and 0.22 in CEU and YRI respectively). Investigators were requested to provide all GWA data that they had obtained for evaluation of any cancer phenotype, including any additional unpublished datasets. Additional genotyping for the two specific variants was encouraged, when a GWA platform had been used that did not directly genotype these polymorphisms (e.g., Affymetrix or Perlegen rather than Illumina). When a study had data on more than one cancer type, data were requested to be provided separately for each cancer type. Investigators were asked to provide also information and clarifications about the design of their studies, and to ensure that population stratification and cryptic relatedness had been appropriately addressed and appropriate quality controls were available for the genotyping. All GWAS investigations that contributed data on these SNPs used stringent QC standards (as described in detail in their original publications) and the two SNPs fulfilled these standards. Approval from local institutional review boards and steering committees was obtained, as deemed necessary for each study by its investigators. The contributed data were checked for completeness and with logical queries and any missing or unclear information was clarified through communication with the contributing investigators.

**Meta-analysis**

For each SNP, we performed meta-analyses including the data from all eligible cancer studies (regardless of the specific cancer phenotype addressed) and also subgroup meta-analyses, with each subgroup limited to studies addressing a specific cancer phenotype. A separate analysis compared the results of the association for prostate cancer versus the association for all other cancers combined. All analyses followed the per allele (log-additive) model of inheritance with effect sizes expressed in the odds ratio (OR) scale using both fixed and random effects models[45]. Heterogeneity testing used the Q statistic (considered statistically significant at \(p<0.10\)), and the \(I^2\) metric[46] and its 95% CIs [47]. Analyses excluding data from studies with pooled genotyping gave similar results (not shown).

Based on the accumulated total sample size and given the minor allele frequencies of these two polymorphisms in HapMap CEU (47% for rs4430796 A allele and 47% for rs7501939 T allele), the meta-analysis had 95% or higher power to detect an association of OR \(= 1.10\) at alpha \(= 0.05\) with each of the two polymorphisms for overall cancer risk, prostate cancer risk, or other cancer risk. Reported p-values are two-tailed. Analyses were performed in STATA 10.0 (College Station, Texas).

**Acknowledgments**


1Department of Health Sciences Research, College of Medicine, Mayo Clinic, Rochester, MN, USA. 2Laboratory of Translational Genomics, Research Institute, Cancer and Medical Genetics, National Institutes of Health, National Institute of Environmental Health and Human Sciences, Bethesda, MD, USA. 3Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Department of Health and Human Services, Bethesda, MD, USA. 4Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Department Health and Human Services, Bethesda, MD, USA. 5Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, MA, USA. 6Channing Laboratory, Department of Medicine, Brigham and Women’s Hospital, Harvard Medical School, Boston, MA, USA. 7Department of Epidemiology, Harvard School of Public Health, Boston, MA, USA. 8Department of Biostatistics, Harvard School of Public Health, Boston, MA, USA. 9Core Genotyping Facility, SAIC-Frederick Inc., NCI-Frederick, Frederick, MD, USA. 10Bioinformatic Consulting Services, Gaithersburg, MD, USA. 11Department of Obstetrics and Gynecology, New York University School of Medicine, New York, NY, USA. 12Department of Environmental Medicine, New York University School of Medicine, New York, NY, USA. 13New York University Cancer Institute, New York, NY, USA. 14National Institute for Public Health and the Environment (RIVM), Bilthoven, The Netherlands and Department of Gastroenterology and Hepatology, University Medical Centre Utrecht, Utrecht, The Netherlands. 15Samuel Lunenfeld Research Institute, Mount Sinai Hospital, Toronto, Ontario, Canada. 16Department of Laboratory Medicine/Pathology, School of Medicine, University of Minnesota, Minneapolis, MN, USA. 17Prevention and Research Center, Mercy Medical Center, Baltimore, MD, USA. 18Department of Epidemiology & Biostatistics, University of California San Francisco, San Francisco, CA, USA. 19Department of Epidemiology, American Cancer Society, Atlanta, GA, USA. 20Department of Oncology, the Johns Hopkins University School of Medicine, Baltimore, MD, USA. 21Department of Epidemiology, the Bloomberg School of Public Health, The Sol Goldman Pancreatic Cancer Research Center, The Johns Hopkins Medical Institutions, Baltimore, MD, USA. 22Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA, USA. 23Department of Gastrointestinal Medical Oncology, University of Texas M.D. Anderson Cancer Center, Houston, Texas, USA. 24Group Health Center for Health Studies, Seattle, WA, USA. 25Department of Epidemiology and Biostatistics, Memorial Sloan-Kettering Cancer Center, New York, NY, USA. 26Yale University School of Public Health, New Haven, CT, USA. 27Department of Medicine and Vanderbilt-Ingram Cancer Center, Vanderbilt University, Nashville, TN, USA. 28Inserm, Paris-Sud University, Institut Gustave-Roussey, Villejuif, France. 29Divisions of Preventive Medicine and Aging, Department of Medicine, Brigham and Women’s Hospital and Harvard Medical School, Boston, MA, USA.
References


