Original Paper

Human Heredity

Hum Hered 2010;69:193–201
DOI: 10.1159/000289594

Received: August 25, 2009
Accepted after revision: December 18, 2009
Published online: March 5, 2010

Genetic Susceptibility to Type 2 Diabetes Is Associated with Reduced Prostate Cancer Risk

Brandon L. Pierce[a] Habibul Ahsan[a, b]

Department of [a]Health Studies and [b]Medicine and Human Genetics and Cancer Research Center, The University of Chicago, Chicago, Ill., USA

Key Words

Genetic susceptibility · Type 1 diabetes · Type 2 diabetes · Prostate cancer · Single nucleotide polymorphism · Mendelian randomization

Abstract

Objective: To examine the collective effects of type 1 (T1D) and type 2 diabetes (T2D) risk alleles on prostate cancer (PCa) risk. ***Methods:*** Using data on 14 and 18 single nucleotide polymorphisms (SNPs) that effect T1D and T2D risk, respectively, we generated risk scores (a 'risk allele count' and a 'genetic relative risk') for both T1D and T2D for 1,171 non-Hispanic white, PSA-screened PCa cases and 1,101 matched controls from the Cancer Genetic Markers of Susceptibility study. Logistic regression was used to estimate odds ratios (OR) and 95% confidence intervals (CI) for associations between the diabetes risk scores and PCa risk. ***Results:*** Both T2D risk scores, but neither T1D score, showed an inverse association with PCa ($p < 0.01$). These associations remained significant after excluding HNF1B SNP rs4430796 (a known PCa risk factor) from the analysis. The highest quartile of the T2D allele count (>20 risk alleles) was associated with reduced PCa risk (OR = 0.77; CI: 0.60–0.99) compared to the lowest category (<17 risk alleles). ***Conclusions:*** These results suggest that individuals with increased genetic susceptibility to T2D have decreased risk for PCa. This association is consistent with the observation that individuals with T2D are at decreased risk for PCa; however, data on T2D status was not available for this analysis.

Introduction

Epidemiologic research suggests that diabetes mellitus (DM) is associated with reduced prostate cancer (PCa) risk. Two meta-analyses report PCa relative risks of 0.91 [1] and 0.84 [2] for diabetics compared to non-diabetics. Several authors have hypothesized that this inverse association is attributable to metabolic and hormonal changes associated with type 2 diabetes (T2D, accounting for ~90% of all DM cases) such as decreased testosterone or insulin, which lead to a less carcinogenic environment (reviewed in [1, 2]). Alternative explanations for this association include unmeasured confounding, decreased PSA levels in diabetics [3–5], effects of DM treatment on PCa risk, and/or shared genetic factors for DM and PCa. There is also evidence suggesting that type 1 diabetes (T1D) may influence PCa risk, but very few studies have addressed this hypothesis [6].

Recent genome-wide association (GWA) studies and subsequent meta-analyses have identified numerous common susceptibility variants for both T1D [7–11] and

KARGER

Fax +41 61 306 12 34
E-Mail karger@karger.ch
www.karger.com

Accessible online at:
www.karger.com/hhe

Dr. Brandon L. Pierce
Center for Cancer Epidemiology and Prevention
The University of Chicago
5841 South Maryland Avenue, Suite N101D, Chicago, IL 60637 (USA)
Tel. +1 773 834 3708, Fax +1 773 834 0139, E-Mail brandonpierce@uchicago.edu

T2D [12–16], bringing the number of T1D and T2D susceptibility variants to ~15 and ~20, respectively. GWA studies have also provided further support for a link between DM and PCa risk, as variants in the HNF1B and JAZF1 genes have been shown to influence both PCa and T2D risk, confirming the existence of shared genetic factors and suggesting related disease mechanisms [17]. In light of these findings, a comprehensive analysis of the roles of T1D and T2D risk variants in relation to PCa risk is warranted.

In this analysis, we use information on T1D and T2D susceptibility variants to test associations between T1D and T2D genetic risk scores and PCa risk in a large case-control study of PCa. An observed association between DM (i.e., T2D or T1D) genetic susceptibility and PCa risk would support one (or more) of the following hypotheses: (1) DM status or its consequences (e.g. DM treatment) affect PCa risk or PSA levels; (2) genetic susceptibility to DM (or an associated metabolic phenotype) affects PCa risk, independent of the DM phenotype, or (3) a specific DM susceptibility variant(s) affects PCa risk, independent of the DM phenotype (i.e., pleiotropy). We discuss the plausibility of these hypotheses in light of our results and highlight the strengths and limitations of using data on DM genetic susceptibility to investigate the relationship between DM and PCa.

Materials and Methods

Study Participants

Single nucleotide polymorphism (SNP) data from the Cancer Genetic Markers of Susceptibility (CGEMS) GWA study of PCa were obtained from the National Cancer Institute (NCI) (http://cgems.cancer.gov/data/). These data consist of genotypes for approximately 550,000 SNPs on 1,176 PCa cases and 1,101 matched controls selected from the screening arm of the Prostate, Lung, Colon, and Ovarian (PLCO) Cancer Screening Trial (enrollment from 1993–2003). Eligible PLCO participants (n = 25,521) were non-Hispanic white males who had (1) no history of PCa prior to randomization; (2) ≥1 PSA screening test prior to October 1, 2003; (3) completed a baseline questionnaire; (4) provided consent, and (5) provided sufficient blood/DNA.

Of the 1,361 eligible PCa cases (diagnosed prior to 2004), all 737 aggressive PCa cases (Gleasons Score ≥7 or Stage ≥III) were included. In order to enrich this study for aggressive PCa, only 493 (of 624) non-aggressive cases (Gleason Score <7 and Stage <III) were included, for a total of 1,230 cases. CGEMS controls (n = 1,230) were selected using incidence-density sampling and frequency matched to cases on year of entry into the cohort, age at entry, and number of years under observation. However, the CGEMS data is only publically available in a single-selection format, i.e., each control individual is included in the dataset only once, allowing for the use of standard case-control analysis methods (details in [18]). In other words, controls who were selected more than once were only included once, and controls selected prior to becoming a case were excluded, resulting in a total of 1,160 eligible controls.

DNA quality and quantity were assessed at the Core Genotyping Facility of the NCI; 1,188 cases and 1,110 controls were suitable for genotyping. Genotyping was carried out under contract by the Illumina Corporation, using both the Sentrix® HumanHap300 and HumanHap240 chips (over 560,000 total SNPs). Samples with completion rates <90% and SNP will call rates <90% were excluded. Detailed quality control procedures have been described previously [18]. Of the 1,177 cases and 1,105 controls with data suitable for association analyses, three individuals who were population outliers and two individuals from study centers with insufficient participation were removed [19], resulting in 1,176 cases and 1,101 controls whose data are available from CGEMS.

T1D and T2D SNP Data

We conducted a literature review to identify all SNPs that have been confirmed as T1D or T2D risk variants. Using published GWA studies and review articles, we identified 15 and 19 regions harboring a common SNP(s) associated with T1D [7–11, 20] and T2D [12–16, 21], respectively. Many of these SNPs were present on the genotyping platform used in this study, while others were not. For T1D and T2D SNPs not present in the CGEMS dataset, surrogate tagSNPs ($r^2 > 0.8$) were identified using the Tagger program [22] and data from HapMap CEU samples (release 21 and 22; http://www.hapmap.org/). No suitable tagSNP for T1D SNP rs3087243 (CTLA4) or T2D SNP rs10830963 (MTNR1B) was available in the CGEMS dataset, so these SNPs were not included in the analysis. In total, data on 14 T1D SNPs and 18 T2D SNPs (or a suitable tagSNP) were present in the CGEMS dataset and used in this analysis (table 1).

T1D SNP rs264704 (HLA) had substantial missing data (n = 56) and was replaced by rs9275184 (the second most significant of the >180 significant HLA SNPs reported [7]). There were small amounts of missing data (<1%) for T1D SNPs rs2476601 (n = 1), rs1990760 (n = 2), rs4505848 (n = 1), rs9275184 (n = 4), rs12251307 (n = 3), rs1004446 (n = 1), rs1701704 (n = 1), rs229527 (n = 1); and for T2D SNPs rs2641348 (n = 18), rs13414140 (n = 15), rs4411878 (n = 15), rs4402960 (n = 19), rs13266634 (n = 1), rs2383208 (n = 2), rs11257655 (n = 17), rs1111875 (n = 1), rs2237892 (n = 1), rs5215 (n = 1), rs1353362 (n = 1), and rs8050136 (n = 1). There was substantial missing data for rs4430796 (n = 107; 4.7%); however, no adequate tagSNP was available, so this SNP was included in the analysis. Missingness was not associated with case-control status for any T1D or T2D SNP.

To enable the calculation of 'genetic risk scores' for T1D and T2D (see statistical analysis section below), we obtained effect size estimates for each T1D and T2D risk allele (on T1D and T2D risk, respectively) from the existing literature. For all T1D SNPs, we obtained per-allele odds ratios for T1D risk and allele frequencies from published GWA [7, 8, 10, 11] and follow-up studies [20]. Similar information for all confirmed T2D risk variants were obtained from published GWA studies [12–16] and a review paper [21]. Per-allele odds ratios are referred to as generic 'relative risks' (RRs) from this point forward, because our calculations treat these odds ratios as estimates of risk ratios. For RRs previously reported using the risk allele as the reference allele, the reciprocal RR is reported, using the non-risk allele as the reference (table 1).

Table 1. Characteristics of T1D[a] and type 2 diabetes T2D[b] risk variants

Gene	Locus	CGEMS SNP	Previous reports			CGEMS controls	
			risk/non-risk allele	RR for T1D or T2D[c]	risk allele frequency	risk allele frequency	HWE p value
T1D							
PTPN22	1p13	rs2476601	T/C	2.03	0.11	0.10	0.22
IFIH1	2q24	rs1990760	A/G	1.18	0.62	0.61	0.09
IL2/IL21	4q27	rs4505848[d]	A/G	1.20	0.66	0.66	0.69
HLA	6p21	rs9275184[e]	C/T	5.18	0.09	0.10	0.87
BACH2	6q15	rs3757247[f]	A/G	1.13	0.46	0.46	0.36
IL2RA	10p15	rs12251307	C/T	1.33	0.88	0.88	0.49
PRKCQ	10p15	rs947474	A/G	1.16	0.81	0.81	0.28
INS	11p15	rs1004446	C/T	1.61	0.65	0.63	0.20
ERBB3	12q13	rs1701704	C/A	1.25	0.35	0.34	0.25
C12orf30	12q24	rs17696736	G/A	1.20	0.42	0.44	0.16
CTSH	15q24	rs2870085[g]	C/T	1.16	0.68	0.67	0.73
CLEC16A	16p13	rs2903692	G/A	1.50	0.62	0.65	0.11
PTPN2	18p11	rs1893217	G/A	1.20	0.17	0.16	0.65
C1QTNF6	22p13	rs229527[h]	A/C	1.11	0.40	0.43	0.85
T2D							
NOTCH2	1p13	rs2641348	G/A	1.13	0.10	0.11	0.17
THADA	2p21	rs13414140[i]	C/T	1.15	0.90	0.87	0.20
PPARG	3p25	rs6802898[j]	C/T	1.14	0.87	0.87	0.25
ADAMTS9	3p14	rs4411878[k]	C/T	1.09	0.76	0.75	0.37
IGF2BP2	3q27	rs4402960	T/G	1.14	0.32	0.32	0.42
WFS1	4p16	rs10012946[l]	C/T	1.12	0.60	0.59	0.48
CDKAL1	6p22	rs7756992	G/A	1.14	0.32	0.28	0.40
JAZF1	7p15	rs1635852[m]	T/C	1.10	0.50	0.49	0.50
SLC30A8	8q24	rs13266634	C/T	1.15	0.69	0.71	0.42
CDKN2A/B	9p2	rs2383208[n]	A/G	1.20	0.83	0.81	0.32
CDC123	10p14	rs11257655[o]	T/C	1.11	0.18	0.20	0.33
HHEX/IDE	10q23	rs1111875	C/T	1.15	0.65	0.59	0.49
TCF7L2	10q25	rs7903146	T/C	1.37	0.31	0.30	0.41
KCNQ1	11p15	rs2237892	C/T	1.29	0.93	0.94	0.13
KCNJ11	11p15	rs5215	C/T	1.14	0.35	0.37	0.45
TSPAN8	12q21	rs1353362[p]	C/T	1.09	0.27	0.27	0.40
FTO	16q12	rs8050136	A/C	1.17	0.40	0.42	0.48
HNF1B	17q12	rs4430796	G/A	1.10	0.47	0.50	0.50

RR = Relative risk.

[a] No tagSNP available for rs3087243 (CTLA4); [b] no tagSNP available for rs10830963 (MTNR1B); [c] RRs are previously reported per-allele odds ratios for the risk allele; [d] r^2 = 0.826 with rs17388568; [e] rs2647044 (from Hakonarson) was out of HWE; [f] r^2 = 0.904 with rs11755527; [g] r^2 = 0.926 with rs3825932; [h] r^2 = 1.0 with rs229541; [i] r^2 = 1.0 with rs7578597; [j] r^2 = 1.0 with rs1801282; [k] r^2 = 0.948 with rs4607103; [l] r^2 = 1.0 with rs100010131; [m] r^2 = 0.97 with rs864745; [n] r^2 = 1.0 with rs10811661 (from Scott et al.); [o] r^2 = 0.80 with rs12779790; [p] r^2 = 0.955 with rs79615.

Statistical Analysis

To identify and exclude participants showing evidence of intercontinental admixture, data on all CGEMS participants were combined with data from all three HapMap populations, and a set of ~12,000 genome-wide unlinked SNPs (described in Yu et al. [19]) was used to conduct principal components analysis (PCA) (using smartpca [23], http://genepath.med.harvard.edu/~reich/EIGENSTRAT.htm). The same marker set was then used to estimate genetic relatedness among study participants (PLINK [24], http://pngu.mgh.harvard.edu/purcell/plink/) and to derive principal components (PCs) representing the two major axes of European ancestry (using smartpca). We tested Hardy-Weinberg equilibrium for each SNP (using PLINK) and compared the observed allele frequency to previously reported estimates to further ensure genotyping quality.

For each individual, we generated two susceptibility measures for T1D and T2D. The first measure was a 'risk allele count', the total number of T1D (or T2D) risk alleles carried by an individu-

Table 2. Associations between T1D and T2D risk scores (continuous) and PCa, by aggressiveness, using 1,101 control subjects

Diabetes risk measures	All PCa (n = 1,171)			Less aggressive PCa (n = 487)			More aggressive PCa (n = 684)		
	OR[a]	95% CI	p value	OR[a]	95% CI	p value	OR[a]	95% CI	p value
T1D genetic risk measures									
Risk allele count	0.99	0.96–1.03	0.74	0.99	0.95–1.04	0.78	0.99	0.95–1.04	0.74
Relative risk	0.99	0.95–1.03	0.57	0.96	0.90–1.03	0.21	1.00	0.96–1.05	0.91
Log-relative risk	0.95	0.87–1.04	0.27	0.94	0.84–1.05	0.27	0.96	0.87–1.06	0.46
T2D genetic risk measures									
Risk allele count	0.95	0.92–0.99	0.004	0.96	0.92–1.00	0.05	0.95	0.92–0.99	0.007
Relative risk	0.72	0.58–0.89	0.002	0.76	0.58–1.01	0.06	0.68	0.53–0.88	0.003
Log-relative risk	0.73	0.59–0.91	0.005	0.78	0.59–1.04	0.09	0.69	0.54–0.89	0.005
T2D genetic risk measures (excluding HNF1B SNP)									
Risk allele count	0.96	0.92–0.99	0.02	0.97	0.93–1.01	0.15	0.96	0.92–0.99	0.02
Relative risk	0.74	0.60–0.93	0.008	0.80	0.61–1.06	0.12	0.71	0.55–0.91	0.008
Log-relative risk	0.76	0.61–0.95	0.01	0.82	0.62–1.10	0.13	0.71	0.55–0.92	0.01

[a] ORs and 95% CI are adjusted for categorical age and genetic ancestry.

al. The second susceptibility measure was a 'genetic RR' for T1D (or T2D), which takes the previously-reported effect sizes for each SNPs into account. This measure is the product of the previously reported per-allele RRs for each T1D (or T2D) SNP, with each per-allele RR raised to the power of the number of risk alleles carried by an individual at that SNP. In other words, if an individual carried 25 T2D risk alleles, we took the product of the 25 corresponding per-allele RRs (the non-risk alleles contribute an RR of 1 to this calculation and can be ignored). Individuals with a missing SNP genotype were assigned average allele count in controls, prior to the calculation of the RR measure. Because the reference group (denominator) for this RR is an individual carrying zero risk alleles, we divided this measure by the mean RR in the sample, to generate in a more interpretable RR measure that represents an individual's RR for T1D (or T2D) compared to the average population risk. The distributions of these RR measure have a mean of one and are skewed to the right, as they have a lower bound of zero and no defined upper bound. A third score, a log-transformed genetic RR with an approximate normal distributions and mean zero, was also generated.

All tests of association were two-sided and performed using logistic regression adjusted for categorical age and two axes of ancestry (i.e. principal components from smartpca program). We first tested associations with PCa for three continuous measures of T1D and T2D risk: (1) risk allele count; (2) genetic RR, and (3) log-transformed genetic RR. We then tested quartiles of these T1D and T2D risk scores for association with PCa risk. Analyses of T2D risk were conducted again after excluding HNF1B SNP rs4430796 (a SNP known to associate with both PCa and T2D) from the risk score calculations. This was done to test the hypothesis that cumulative genetic risk for T2D is associated with PCa independent of this known PCa risk factor. In addition, each individual T1D and T2D SNP was tested for association with PCa risk using a log-additive model with genotypes coded as 0, 1, or 2 risk alleles. For single-SNP analyses, only individuals with non-missing SNP data were analyzed. All regressions were performed using SAS software, Version 9.1 (SAS Institute Inc., Cary NC, USA).

Results

According to our analysis of ancestry using PCA, no CGEMS participants showed any evidence of substantial intercontinental admixture, as all subjects clustered tightly with CEU HapMap samples according to the first two principal components (online suppl. fig. 1, www.karger.com/doi/10.1159/000289594). According to identity-by-decent estimates, 5 pairs of subjects were first-degree relatives (one from each pair was removed) and all participants were confirmed to be genetically male. The distribution of the first two principal components derived from PCA of CGEMS data only (i.e., PCs representing European-specific population structure) is shown in online supplementary figure 2. The first principal component, which had a highly skewed distribution, was associated with decreased PCa risk (p = 0.02, from logistic regression).

The eligible 1,171 CGEMS cases and 1,101 controls used in this analysis were closely matched on age (using the following age categories: 55–59, 60–64, 65–69, 70–74). Cases were more likely to have a family history of PCa (11%) than controls (6%). A total of 684 cases were classified as having aggressive PCa, while 487 cases were clas-

Table 3. Associations between T1D and T2D risk scores (categorical) and PCa

	Cases (n = 1,176)	Controls (n = 1,101)	OR[a]	95% CI
T1D risk allele count (%)				
7–12	27.41	27.43	1.00	Ref
13–14	32.02	32.24	0.97	0.78–1.20
15	17.25	16.17	1.07	0.83–1.37
16–22	23.31	23.16	1.01	0.80–1.27
			p-trend = 0.78	
T1D relative risk (%)				
0.07–0.28	25.28	24.70	1.00	Ref
0.29–0.46	25.79	24.34	1.05	0.83–1.33
0.47–0.89	25.45	24.34	1.02	0.81–1.29
0.90–42.9	23.48	26.61	0.86	0.68–1.09
			p-trend = 0.21	
T2D risk allele count (%)				
10–16	21.18	19.35	1.00	Ref
17–18	28.69	26.70	0.99	0.77–1.25
19–20	28.78	28.52	0.93	0.73–1.18
21–27	21.35	25.49	0.77	0.60–0.99
			p-trend = 0.03	
T2D relative risk (%)				
0.26–0.76	25.70	24.25	1.00	Ref
0.77–1.01	25.96	23.98	1.03	0.82–1.30
1.02–1.36	24.85	25.16	0.94	0.74–1.18
1.37–5.44	23.48	26.61	0.85	0.70–1.07
			p-trend = 0.11	
T2D risk allele count (excluding HNF1B; %)				
10–15	19.04	18.53	1.00	Ref
16–17	29.97	27.52	1.06	0.83–1.35
18–19	28.95	29.79	0.94	0.74–1.21
20–26	22.08	24.16	0.88	0.69–1.14
			p-trend = 0.20	
T2D relative risk (excluding HNF1B; %)				
0.28–0.77	25.53	24.43	1.00	Ref
0.78–1.01	25.79	24.16	1.03	0.82–1.30
1.02–1.37	24.59	25.43	0.93	0.73–1.17
1.38–5.45	24.08	25.98	0.90	0.71–1.14
			p-trend = 0.26	

[a] ORs and 95% CI are adjusted for categorical age and genetic ancestry.

sified as having less aggressive PCa. Characteristics of this sample have been described previously [18].

The T1D and T2D risk variants analyzed in this study are described in table 1. The observed allele frequencies in controls were similar to those previously reported in GWA studies of individuals of European ancestry. No markers showed statistically significant deviations from Hardy-Weinberg equilibrium in controls, with the exception of T1D SNP rs264704 (HLA), which was replaced by rs9275184 due to missing data issues (see Methods). For tagSNPs identified using Tagger, pair-wise correlations with the previously reported SNPs are shown in the footnotes to table 1.

All three continuous measures of T2D risk showed a significant inverse association with PCa risk (e.g. T2D risk allele count: p = 0.004) (table 2). Furthermore, after excluding the one SNP known to associate with both T2D and PCa risk (HNF1B, which has opposite effects on T2D and PCa), these associations remained significant, but were slightly attenuated (e.g. T2D risk allele count: p = 0.02). Associations for T2D risk were similar for less aggressive and more aggressive PCa, although slightly stronger for more aggressive PCa. When analyzed as quartiles, the highest category of T2D allele count (21–27 risk alleles) was associated with reduced PCa risk (OR = 0.77; 95% CI: 0.60–0.99; p-trend = 0.03) (table 3). However, no quartile of T2D allele count (after excluding HNF1B) or T2D RR showed significant association with PCa risk. In addition, none of the continuous measures of T1D risk (or their quartiles) was significantly associated with PCa risk.

The only T1D SNP to show a nominally significant association with PCa was rs3757247 (BACH2; OR = 1.12, 95% CI = 1.00–1.26) (table 4). The only T2D SNPs to show nominally significant associations with PCa were rs8050136 (FTO; OR = 0.87, 95% CI: 0.77–0.98) and rs4430796 (HNF1B; OR = 0.87, 95% CI: 0.77–0.98). Adjustment for European axes of ancestry had very little effect on the SNP-specific ORs and the ORs for T1D and T2D genetic risk measures.

Discussion

In this large case-control study of PCa, we show that increasing T2D genetic susceptibility, measured using 18 T2D-related SNPs, is associated with reduced PCa risk. This association was slightly stronger for more aggressive PCa. We did not observe a significant association between T1D genetic susceptibility, measured using 14 SNPs, and PCa risk.

There are several possible causal explanations for the observed association between T2D susceptibility and PCa risk, as outlined in the Introduction. First, it is possible that the effect of T2D susceptibility on PCa risk is mediated by T2D status. However, this is unlikely to be the sole explanation, considering the modest reported associations between DM and PCa [1, 2] and the small amount of variance in T2D susceptibility that is ex-

Table 4. Associations between T1D and T2D risk variants and PCa risk

Gene	SNP	Risk allele	Risk allele frequency		Age- and ancestry-adjusted (sorted by OR)		
			cases	controls	OR[a]	95% CI	p value[b]
T1D							
PTPN2	rs1893217	G	0.146	0.159	0.90	0.76–1.05	0.17
CLEC16A	rs2903692	G	0.628	0.650	0.91	0.80–1.02	0.12
ERBB3	rs1701704	C	0.328	0.341	0.93	0.82–1.05	0.28
HLA	rs9275184	C	0.098	0.104	0.93	0.77–1.13	0.53
IFIH1	rs1990760	A	0.601	0.613	0.94	0.83–1.06	0.25
PTPN22	rs2476601	T	0.092	0.096	0.95	0.78–1.15	0.51
INS	rs1004446	C	0.621	0.630	0.97	0.86–1.09	0.61
IL2/IL21	rs4505848	A	0.656	0.662	0.99	0.87–1.11	0.81
C1QTNF6	rs229527	A	0.422	0.427	0.99	0.88–1.12	0.87
PRKCQ	rs947474	A	0.812	0.810	1.01	0.87–1.18	0.80
CTSH	rs2870085	C	0.682	0.670	1.06	0.93–1.20	0.37
C12orf30	rs17696736	G	0.454	0.438	1.07	0.95–1.21	0.25
BACH2	rs3757247	A	0.489	0.461	1.12	1.00–1.26	0.06
IL2RA	rs12251307	C	0.893	0.877	1.17	0.97–1.40	0.10
T2D							
FTO	rs8050136	A	0.391	0.423	0.87	0.77–0.98	0.02
HNF1B	rs4430796	G	0.459	0.498	0.87	0.77–0.97	0.02
NOTCH2[c]	rs2641348	G	0.096	0.109	0.87	0.71–1.05	0.14
PPARG	rs6802898	C	0.852	0.868	0.88	0.74–1.04	0.14
KCNQ1	rs2237892	C	0.931	0.940	0.88	0.69–1.12	0.30
KCNJ11	rs5215	C	0.347	0.373	0.89	0.78–1.00	0.07
IGF2BP2	rs4402960	T	0.300	0.321	0.91	0.81–1.04	0.16
CDKN2A/B	rs2383208	A	0.802	0.810	0.95	0.82–1.10	0.47
SLC30A8	rs13266634	C	0.699	0.705	0.97	0.86–1.11	0.70
TCF7L2	rs7903146	T	0.287	0.298	0.97	0.85–1.10	0.66
HHEX	rs1111875	C	0.586	0.592	0.98	0.87–1.10	0.79
JAZF1[c]	rs1635852	T	0.490	0.495	0.98	0.87–1.10	0.69
ADAMTS9[c]	rs4411878	C	0.746	0.748	0.98	0.85–1.12	0.74
CDKAL1	rs7756992	G	0.280	0.282	1.00	0.88–1.14	0.99
WFS	rs10012946	C	0.597	0.594	1.02	0.91–1.15	0.81
CDC123[c]	rs11257655	T	0.208	0.202	1.03	0.89–1.19	0.67
TSPAN8[c]	rs1353362	C	0.274	0.264	1.04	0.92–1.19	0.52
THADA[c]	rs13414140	C	0.885	0.874	1.10	0.92–1.32	0.30

[a] Per-allele ORs are adjusted for categorical age and ancestry and calculated using the non-risk allele as the reference group; [b] p value is derived from an Armitage test of trend; [c] SNPs identified in a previous meta-analysis.

plained by these variants (<10% [21]). If we make liberal assumptions regarding the total phenotypic variance in T2D explained by these SNPs (5%), the effect of T2D on PCa (OR = 0.85), and the prevalence of T2D in controls (0.10), simulation-based power analyses suggest that a study of this size would have very low power to detect the effect of T2D susceptibility (<5%), if mediated entirely by T2D. Therefore, it is not likely that the effect of T2D susceptibility on PCa is dependent on a T2D diagnosis. Because we do not have T2D phenotype data, we were unable to determine the degree to which this effect is mediated by T2D status or T2D-related phenotypes. Such data should be integrated into future studies, where instrument variable analyses (i.e. Mendelian randomization) can be used to generate estimates of the causal effect of T2D-related phenotypes on PCa [25].

A second potential explanation for this association is pleiotropy, i.e., a specific variant affects both T2D and PCa risk, independently. However, the T2D risk variants appear to affect PCa risk in a collective manner, suggesting that a pleiotropy is not the sole explanation. We conclude this based upon the many weak, inverse associa-

tions with PCa we observe for the individual T2D risk alleles (i.e., 14 out of 18 T2D risk alleles showed ORs <1, which is more than is expected by chance; two-sided binomial test p = 0.03). Furthermore, after removing the T2D SNP previously known to influence PCa risk (HNF1B rs4430796) we still observe a significant, although slightly attenuated, association between the T2D risk score and PCa, indicating that the association is not solely due to the pleiotropic effects of this specific variant. The collective nature of this association suggests that effects of these variants on PCa may be driven, at least in part, by biological processes related to T2D. In addition to possible mediation through T2D status itself, this effect could also be mediated through pre-diabetes phenotypes and/or undiagnosed T2D.

Because we are studying genetic determinants of T2D, rather than T2D status itself, we can conclude that the observed association is not due to reverse causation or confounding (by any factor other than ancestry, for which we have adjusted). This result lends credence to the hypothesis that T2D (or T2D-related phenotypes) have a protective effect on PCa, although we cannot formally test this hypothesis without T2D phenotype data.

Evidence from previous studies shows that diabetics' risk of PCa decreases as time since DM diagnoses increases [26–29], suggesting a causal duration-response relationship. Such a causal relationship between DM and PCa could be attributed to several different factors. For example, long-term T2D is associated with decreases in circulating concentrations of several potential tumor promoting hormones such as insulin [30], insulin-like growth factor 1 [31], testosterone [32], and leptin [33]. DM is also associated with decreased PSA levels [3–5], potentially leading to decreased PCa risk; however, evidence is not consistent regarding whether the DM-PCa association was stronger in the pre-PSA era [27] or in the PSA era [2]. Low PSA levels among diabetics have been associated with specific DM treatments (e.g. insulin) and high hemoglobin A1c levels [3]. Also unclear is the degree to which screening habits differ by DM status. In the Multiethnic Cohort Study, diabetics were less likely to receive PSA screening, and screening factors (including PSA levels and screening habits) were shown to account for a modest proportion (~20%) of the DM-PCa association [4]. In this analysis, all CGEMS participants had at least one PCa screen prior to enrollment and were screened for the duration of the study, so it is possible that our results could be due to differences in PSA levels between those with low and high T2D susceptibility.

We observe four (out of 18) T2D risk alleles with an OR >1.00 for PCa risk, associations that are inconsistent with our hypothesis. However, three of these four SNPs were identified in a T2D GWA study meta-analysis [16]. SNPs identified in meta-analyses are likely to make lesser overall contributions to T2D genetic susceptibility than SNPs identified in primary analyses, as their effects are more difficult to detect (i.e. more power is needed). It is also possible that a subset of the SNPs identified in this meta-analysis are false positives, as they have not been formally replicated in independent samples. Considering these possibilities, it seems reasonable that a small number of SNPs show associations that are not consistent with our hypothesis, due to a combination of sampling variation, weak effects, and false positives.

This study demonstrates how disease susceptibility variants with weak individual effects can be aggregated into meaningful risk measures and used to investigate relationships between diseases. Our 'genetic RR' calculation (based on methods described in [34]) incorporates information on previously reported RRs; however, because all of the T2D risk variants have similar and modest reported RRs (all <1.4), it may not be critical to account for different effect sizes when constructing a T2D susceptibility measure. The T2D allele count (which does not account for differences) and the T2D 'genetic RR' measure show similar associations with PCa, supporting the notion that effect sizes can be ignored, as if often done in studies assessing the predictive value of T2D variants [35, 36]. Because T1D risk variants show larger differences in their effect sizes, and it may be more critical to account for effect sizes when generating measures of T1D risk.

This study may have limited power to detect an association between T1D susceptibility, if the association is primarily mediated through the T1D phenotype. In other words, even if complete information on T1D status was available for this cohort, its prevalence would be quite low (<0.5%), and its association with PCa would have to be very large in order to be detected (and even larger if using genetic proxies for T1D). Alternative study designs would likely be required to study the effect of T1D on PCa risk. Interestingly, the hypothesized mechanisms by which T2D is related to PCa are also plausible explanations for a similar relationship between T1D on PCa [6]. However, T1D and T2D are clearly quite distinct in terms of etiology, genetic susceptibility, pathology, duration, and treatment.

The low prevalence of T1D (and T2D) does, however, emphasize the potential benefit of using common genetic risk factors to investigate relationships between diseases. Unlike T1D status, T1D risk variants are common and

easily analyzable in an unselected sample. Information on T1D susceptibility alone may be valuable if it represents important sub-clinical biological processes related to T1D that could influence PCa risk regardless of whether a T1D diagnosis is made. Nevertheless, in this study, we observe only a modest, non-significant association with PCa risk that is driven by individuals in the highest quintile of T1D risk (OR = 0.86; 95% CI: 0.68–1.09). This association needs to be tested in larger studies.

Several T2D SNPs (in addition to HNF1B rs4430796) have been previously examined in relation to PCa, and our results are consistent with these studies. Several small case-control studies of Caucasian men have shown modest (non-statistically significant) associations between the PPARG T2D risk allele and reduced PCa risk [37, 38]. Candidate gene studies of TCF7L2 have shown either a modest (non-statistically significant) association between the T2D risk allele and decreased PCa risk [39] or associations close to the null [40]. Variation in JAZF has been implicated in both PCa and T2D [16]; however, these associations do not appear to be due to a common causal SNP (i.e., the T2D- and PCa-associated SNPs are uncorrelated), suggesting that variation in JAZF may influence PCa and T2D through independent genetic mechanisms [17].

In conclusion, this is the first study to use comprehensive SNP data to examine the association between both T1D and T2D genetic susceptibility and PCa risk. These results are consistent with the hypothesis that DM has a protective effect on PCa risk. Future research should combine genetic susceptibility data with T2D phenotype data to determine to what degree the association between T2D genetic risk and PCa is mediated by T2D or related phenotypes (such as fasting glucose, fasting insulin, or glucose tolerance).

Acknowledgements

This work was supported by grants from the National Institutes of Health [CA122171 and CA102484 to H.A.]. The data used in this analysis was provided by the Cancer Genetic Markers of Susceptibility (CGEMS) project, a National Cancer Institute initiative. We would like to thank Dr. Stephen J. Chanock and Dr. Kai Yu of the National Cancer Institute, who provided technical advice and information regarding the CGEMS dataset.

References

1 Bonovas S, Filioussi K, Tsantes A: Diabetes mellitus and risk of prostate cancer: A meta-analysis. Diabetologia 2004;47:1071–1078.
2 Kasper JS, Giovannucci E: A meta-analysis of diabetes mellitus and the risk of prostate cancer. Cancer Epidemiol Biomarkers Prev 2006;15:2056–2062.
3 Muller H, Raum E, Rothenbacher D, Stegmaier C, Brenner H: Association of diabetes and body mass index with levels of prostate-specific antigen: Implications for correction of prostate-specific antigen cutoff values? Cancer Epidemiol Biomarkers Prev 2009.
4 Waters KM, Henderson BE, Stram DO, Wan P, Kolonel LN, Haiman CA: Association of diabetes with prostate cancer risk in the multiethnic cohort. Am J Epidemiol 2009;169:937–945.
5 Werny DM, Saraiya M, Gregg EW: Prostate-specific antigen values in diabetic and non-diabetic US men, 2001–2002. Am J Epidemiol 2006;164:978–983.
6 Pierce BL, Plymate S, Ostrander EA, Stanford JL: Diabetes mellitus and prostate cancer risk. Prostate 2008;68:1126–1132.
7 Hakonarson H, Grant SF, Bradfield JP, Marchand L, Kim CE, Glessner JT, Grabs R, Casalunovo T, Taback SP, Frackelton EC, Lawson ML, Robinson LJ, Skraban R, Lu Y, Chiavacci RM, Stanley CA, Kirsch SE, Rappaport EF, Orange JS, Monos DS, Devoto M, Qu HQ, Polychronakos C: A genome-wide association study identifies kiaa0350 as a type 1 diabetes gene. Nature 2007;448:591–594.
8 Hakonarson H, Qu HQ, Bradfield JP, Marchand L, Kim CE, Glessner JT, Grabs R, Casalunovo T, Taback SP, Frackelton EC, Eckert AW, Annaiah K, Lawson ML, Otieno FG, Santa E, Shaner JL, Smith RM, Onyiah CC, Skraban R, Chiavacci RM, Robinson LJ, Stanley CA, Kirsch SE, Devoto M, Monos DS, Grant SF, Polychronakos C: A novel susceptibility locus for type 1 diabetes on chr12q13 identified by a genome-wide association study. Diabetes 2008;57:1143–1146.
9 Smyth DJ, Cooper JD, Bailey R, Field S, Burren O, Smink LJ, Guja C, Ionescu-Tirgoviste C, Widmer B, Dunger DB, Savage DA, Walker NM, Clayton DG, Todd JA: A genome-wide association study of nonsynonymous snps identifies a type 1 diabetes locus in the interferon-induced helicase (ifih1) region. Nat Genet 2006;38:617–619.
10 Todd JA, Walker NM, Cooper JD, Smyth DJ, Downes K, Plagnol V, Bailey R, Nejentsev S, Field SF, Payne F, Lowe CE, Szeszko JS, Hafler JP, Zeitels L, Yang JH, Vella A, Nutland S, Stevens HE, Schuilenburg H, Coleman G, Maisuria M, Meadows W, Smink LJ, Healy B, Burren OS, Lam AA, Ovington NR, Allen J, Adlem E, Leung HT, Wallace C, Howson JM, Guja C, Ionescu-Tirgoviste C, Simmonds MJ, Heward JM, Gough SC, Dunger DB, Wicker LS, Clayton DG: Robust associations of four new chromosome regions from genome-wide analyses of type 1 diabetes. Nat Genet 2007;39:857–864.
11 Cooper JD, Smyth DJ, Smiles AM, Plagnol V, Walker NM, Allen JE, Downes K, Barrett JC, Healy BC, Mychaleckyj JC, Warram JH, Todd JA: Meta-analysis of genome-wide association study data identifies additional type 1 diabetes risk loci. Nat Genet 2008;40:1399–1401.
12 Gudmundsson J, Sulem P, Steinthorsdottir V, Bergthorsson JT, Thorleifsson G, Manolescu A, Rafnar T, Gudbjartsson D, Agnarsson BA, Baker A, Sigurdsson A, Benediktsdottir KR, Jakobsdottir M, Blondal T, Stacey SN, Helgason A, Gunnarsdottir S, Olafsdottir A, Kristinsson KT, Birgisdottir B, Ghosh S, Thorlacius S, Magnusdottir D, Stefansdottir G, Kristjansson K, Bagger Y, Wilensky RL, Reilly MP, Morris AD, Kimber CH, Adeyemo A, Chen Y, Zhou J, So WY, Tong PC, Ng MC, Hansen T, Andersen G, Borch-Johnsen K, Jorgensen T, Tres A, Fuertes F, Ruiz-Echarri M, Asin L, Saez B, van Boven E, Klaver S, Swinkels DW, Aben KK, Graif T, Cashy J, Suarez BK, van Vierssen Trip O, Frigge ML, Ober C, Hofker MH, Wijmenga C, Christiansen C, Rader DJ, Palmer CN, Rotimi C, Chan JC, Pedersen O, Sigurdsson G, Benediktsson R, Jonsson E, Einarsson GV, Mayordomo JI, Catalona WJ, Kiemeney LA,

Barkardottir RB, Gulcher JR, Thorsteinsdottir U, Kong A, Stefansson K: Two variants on chromosome 17 confer prostate cancer risk, and the one in tcf2 protects against type 2 diabetes. Nat Genet 2007;39:977–983.

13 Scott LJ, Mohlke KL, Bonnycastle LL, Willer CJ, Li Y, Duren WL, Erdos MR, Stringham HM, Chines PS, Jackson AU, Prokunina-Olsson L, Ding CJ, Swift AJ, Narisu N, Hu T, Pruim R, Xiao R, Li XY, Conneely KN, Riebow NL, Sprau AG, Tong M, White PP, Hetrick KN, Barnhart MW, Bark CW, Goldstein JL, Watkins L, Xiang F, Saramies J, Buchanan TA, Watanabe RM, Valle TT, Kinnunen L, Abecasis GR, Pugh EW, Doheny KF, Bergman RN, Tuomilehto J, Collins FS, Boehnke M: A genome-wide association study of type 2 diabetes in finns detects multiple susceptibility variants. Science 2007; 316:1341–1345.

14 Sladek R, Rocheleau G, Rung J, Dina C, Shen L, Serre D, Boutin P, Vincent D, Belisle A, Hadjadj S, Balkau B, Heude B, Charpentier G, Hudson TJ, Montpetit A, Pshezhetsky AV, Prentki M, Posner BI, Balding DJ, Meyre D, Polychronakos C, Froguel P: A genome-wide association study identifies novel risk loci for type 2 diabetes. Nature 2007;445:881–885.

15 Steinthorsdottir V, Thorleifsson G, Reynisdottir I, Benediktsson R, Jonsdottir T, Walters GB, Styrkarsdottir U, Gretarsdottir S, Emilsson V, Ghosh S, Baker A, Snorradottir S, Bjarnason H, Ng MC, Hansen T, Bagger Y, Wilensky RL, Reilly MP, Adeyemo A, Chen Y, Zhou J, Gudnason V, Chen G, Huang H, Lashley K, Doumatey A, So WY, Ma RC, Andersen G, Borch-Johnsen K, Jorgensen T, van Vliet-Ostaptchouk JV, Hofker MH, Wijmenga C, Christiansen C, Rader DJ, Rotimi C, Gurney M, Chan JC, Pedersen O, Sigurdsson G, Gulcher JR, Thorsteinsdottir U, Kong A, Stefansson K: A variant in cdkal1 influences insulin response and risk of type 2 diabetes. Nat Genet 2007;39:770–775.

16 Zeggini E, Scott LJ, Saxena R, Voight BF, Marchini JL, Hu T, de Bakker PI, Abecasis GR, Almgren P, Andersen G, Ardlie K, Bostrom KB, Bergman RN, Bonnycastle LL, Borch-Johnsen K, Burtt NP, Chen H, Chines PS, Daly MJ, Deodhar P, Ding CJ, Doney AS, Duren WL, Elliott KS, Erdos MR, Frayling TM, Freathy RM, Gianniny L, Grallert H, Grarup N, Groves CJ, Guiducci C, Hansen T, Herder C, Hitman GA, Hughes TE, Isomaa B, Jackson AU, Jorgensen T, Kong A, Kubalanza K, Kuruvilla FG, Kuusisto J, Langenberg C, Lango H, Lauritzen T, Li Y, Lindgren CM, Lyssenko V, Marvelle AF, Meisinger C, Midthjell K, Mohlke KL, Morken MA, Morris AD, Narisu N, Nilsson P, Owen KR, Palmer CN, Payne F, Perry JR, Pettersen E, Platou C, Prokopenko I, Qi L, Qin L, Rayner NW, Rees M, Roix JJ, Sandbaek A, Shields B, Sjogren M, Steinthorsdottir V, Stringham HM, Swift AJ, Thorleifsson G, Thorsteinsdottir U, Timpson NJ, Tuomi T, Tuomilehto J, Walker M, Watanabe RM, Weedon MN, Willer CJ, Illig T, Hveem K, Hu FB, Laakso M, Stefansson K, Pedersen O, Wareham NJ, Barroso I, Hattersley AT, Collins FS, Groop L, McCarthy MI, Boehnke M, Altshuler D: Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. Nat Genet 2008;40:638–645.

17 Frayling TM, Colhoun H, Florez JC: A genetic link between type 2 diabetes and prostate cancer. Diabetologia 2008;51:1757–1760.

18 Yeager M, Orr N, Hayes RB, Jacobs KB, Kraft P, Wacholder S, Minichiello MJ, Fearnhead P, Yu K, Chatterjee N, Wang Z, Welch R, Staats BJ, Calle EE, Feigelson HS, Thun MJ, Rodriguez C, Albanes D, Virtamo J, Weinstein S, Schumacher FR, Giovannucci E, Willett WC, Cancel-Tassin G, Cussenot O, Valeri A, Andriole GL, Gelmann EP, Tucker M, Gerhard DS, Fraumeni JF Jr., Hoover R, Hunter DJ, Chanock SJ, Thomas G: Genome-wide association study of prostate cancer identifies a second risk locus at 8q24. Nat Genet 2007;39:645–649.

19 Yu K, Wang Z, Li Q, Wacholder S, Hunter DJ, Hoover RN, Chanock S, Thomas G: Population substructure and control selection in genome-wide association studies. PLoS ONE 2008;3:e2551.

20 Smyth DJ, Cooper JD, Howson JM, Walker NM, Plagnol V, Stevens H, Clayton DG, Todd JA: Ptpn22 trp620 explains the association of chromosome 1p13 with type 1 diabetes and shows a statistical interaction with hla class ii genotypes. Diabetes 2008;57:1730–1737.

21 Prokopenko I, McCarthy MI, Lindgren CM: Type 2 diabetes: New genes, new understanding. Trends Genet 2008;24:613–621.

22 de Bakker PI, Yelensky R, Pe'er I, Gabriel SB, Daly MJ, Altshuler D: Efficiency and power in genetic association studies. Nat Genet 2005;37:1217–1223.

23 Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D: Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet 2006;38:904–909.

24 Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC: PLINK: A tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 2007;81:559–575.

25 Lawlor DA, Harbord RM, Sterne JA, Timpson N, Davey Smith G: Mendelian randomization: Using genes as instruments for making causal inferences in epidemiology. Stat Med 2008;27:1133–1163.

26 Giovannucci E, Rimm EB, Stampfer MJ, Colditz GA, Willett WC: Diabetes mellitus and risk of prostate cancer (United States). Cancer Causes Control 1998;9:3–9.

27 Kasper JS, Liu Y, Giovannucci E: Diabetes mellitus and risk of prostate cancer in the health professionals follow-up study. Int J Cancer 2009;124:1398–1403.

28 Rodriguez C, Patel AV, Mondul AM, Jacobs EJ, Thun MJ, Calle EE: Diabetes and risk of prostate cancer in a prospective cohort of us men. Am J Epidemiol 2005;161:147–152.

29 Zhu K, Lee IM, Sesso HD, Buring JE, Levine RS, Gaziano JM: History of diabetes mellitus and risk of prostate cancer in physicians. Am J Epidemiol 2004;159:978–982.

30 Giovannucci E, Michaud D: The role of obesity and related metabolic disturbances in cancers of the colon, prostate, and pancreas. Gastroenterology 2007;132:2208–2225.

31 Monzavi R, Cohen P: Igfs and igfbps: Role in health and disease. Best Pract Res Clin Endocrinol Metab 2002;16:433–447.

32 Laaksonen DE, Niskanen L, Punnonen K, Nyyssonen K, Tuomainen TP, Valkonen VP, Salonen R, Salonen JT: Testosterone and sex hormone-binding globulin predict the metabolic syndrome and diabetes in middle-aged men. Diabetes Care 2004;27:1036–1041.

33 Saglam K, Aydur E, Yilmaz M, Goktas S: Leptin influences cellular differentiation and progression in prostate cancer. J Urol 2003;169:1308–1311.

34 Pharoah PD, Antoniou AC, Easton DF, Ponder BA: Polygenes, risk prediction, and targeted prevention of breast cancer. N Engl J Med 2008;358:2796–2803.

35 Lango H, Palmer CN, Morris AD, Zeggini E, Hattersley AT, McCarthy MI, Frayling TM, Weedon MN: Assessing the combined impact of 18 common genetic variants of modest effect sizes on type 2 diabetes risk. Diabetes 2008;57:3129–3135.

36 Meigs JB, Shrader P, Sullivan LM, McAteer JB, Fox CS, Dupuis J, Manning AK, Florez JC, Wilson PW, D'Agostino RB, Sr., Cupples LA: Genotype score in addition to common risk factors for prediction of type 2 diabetes. N Engl J Med 2008;359:2208–2219.

37 Paltoo D, Woodson K, Taylor P, Albanes D, Virtamo J, Tangrea J: Pro12ala polymorphism in the peroxisome proliferator-activated receptor-gamma (ppar-gamma) gene and risk of prostate cancer among men in a large cancer prevention study. Cancer Lett 2003;191:67–74.

38 Zmuda JM, Modugno F, Weissfeld JL, Cauley JA, Trump DL, Moffett SP, Ferrell RE: Peroxisome proliferator-activated receptor-gamma polymorphism, body mass and prostate cancer risk: Evidence for gene-environment interaction. Oncology 2006;70:185–189.

39 Agalliu I, Suuriniemi M, Prokunina-Olsson L, Johanneson B, Collins FS, Stanford JL, Ostrander EA: Evaluation of a variant in the transcription factor 7-like 2 (tcf7l2) gene and prostate cancer risk in a population-based study. Prostate 2008;68:740–747.

40 Folsom AR, Pankow JS, Peacock JM, Bielinski SJ, Heiss G, Boerwinkle E: Variation in tcf7l2 and increased risk of colon cancer: The atherosclerosis risk in communities (aric) study. Diabetes Care 2008;31:905–909.