Int J Mol Epidemiol Genet 2013;4(1):49-60
www.ijmeg.org /ISSN:1948-1756/IJMEG1301002

Original Article

Pleiotropy and pathway analyses of genetic variants associated with both type 2 diabetes and prostate cancer

LA Raynor[1], James S Pankow[2], Laura J Rasmussen-Torvik[3], Weihong Tang[2], Anna Prizment[2], David J Couper[4]

[1]*Division of Academic General Pediatrics, Department of Pediatrics, University of Minnesota, Minneapolis, MN;* [2]*Division of Epidemiology and Community Health, School of Public Health, University of Minnesota, Minneapolis, MN;* [3]*Department of Preventive Medicine, Northwestern University, Feinberg School of Medicine, Chicago, IL;* [4]*Department of Biostatistics, University of North Carolina at Chapel Hill, Chapel Hill, NC*

Received January 28, 2013; Accepted February 23, 2013; Epub March 18, 2013; Published March 28, 2013

Abstract: Aims: Epidemiological evidence shows that diabetes is associated with a reduced risk of prostate cancer. The objective of this study was to identify genes that may contribute to both type 2 diabetes and prostate cancer outcomes and the biological pathways these diseases may share. Methods: The Atherosclerosis Risk in Communities (ARIC) Study is a population-based prospective cohort study in four U.S. communities that included a baseline examination in 1987-89 and three follow-up exams at three year intervals. Participants were 45-64 years old at baseline. We conducted a genomewide association (GWA) study of incident type 2 diabetes in males, summarized variation across genetic loci into a polygenic risk score, and determined if that diabetes risk score was also associated with incident prostate cancer in the same study population. Secondarily we conducted a separate GWA study of prostate cancer, performed a pathway analysis of both type 2 diabetes and prostate cancer, and qualitatively determined if any of the biochemical pathways identified were shared between the two outcomes. Results: We found that the polygenic risk score for type 2 diabetes was not statistically significantly associated with prostate cancer. The pathway analysis also found no overlap between pathways associated with type 2 diabetes and prostate cancer. However, it did find that the growth hormone signaling pathway was statistically significantly associated with type 2 diabetes (p=0.0001). Conclusion: The inability of this study to find an association between type 2 diabetes polygenic risk scores with prostate cancer or biological pathways in common suggests that shared genetic variants may not contribute significantly to explaining shared etiology.

Keywords: Type 2 diabetes, prostate cancer, polygenic risk score, pathway analysis

Introduction

Epidemiological evidence shows that diabetes is associated with a reduced risk of prostate cancer [1, 2]. Many of the mechanisms hypothesized to explain this inverse association posit that the effect of type 2 diabetes on prostate cancer risk is mediated by type 2 diabetes status. Specifically, having type 2 diabetes may decrease prostate cancer risk through (1) its influence on insulin levels; (2) its influence upon the bioavailability of insulin growth factor 1, leptin, and free testosterone; (3) type 2 diabetes drug treatments (i.e. metformin); and (4) changes in lifestyle and diet [3, 4]. However, the exact mechanism is unknown at this time.

The association between these two diseases could also be explained via pleiotropy, whereby specific genetic variants affect both type 2 diabetes and prostate cancer risk, independently [4]. Several genes recently identified in type 2 diabetes GWA studies have also been found to be associated with prostate cancer risk [3, 5-7]. Additionally, Pierce and Ahsan created a type 2 diabetes risk score using 18 common diabetes SNPs and found an inverse association with prostate cancer, indicating that individuals with increased genetic susceptibility to diabetes have decreased risk of prostate cancer [4].

To date, the majority of studies of type 2 diabetes genetic risk variants and prostate cancer

have largely been candidate gene analyses and no research study has been conducted to systematically identify the genes that overlap between diabetes and prostate cancer outcomes in the same study population. Furthermore, all studies to date have focused on individual SNPs and no analyses have been conducted to identify genetic pathways that may overlap between these two biologically related disease outcomes.

Therefore, to identify genes that may contribute to both diabetes and prostate cancer outcomes and the biological pathways these diseases may share, we conducted a GWA study of type 2 diabetes, summarized variation across genetic loci into a polygenic risk score, and determined if that diabetes risk score was also associated with prostate cancer. Secondarily, we performed a GWA study of prostate cancer and conducted separate pathway analyses for each outcome to determine if any of the biochemical pathways identified were shared between type 2 diabetes and prostate cancer.

Methods

Subjects

The ARIC study began in 1987-9 and recruited a population-based cohort from four U.S. communities including: Forsyth County, NC; Jackson, MS; the northwest suburbs of Minneapolis, MN; and Washington County, MD [8]. The Jackson, MS, site recruited exclusively self-reported African Americans. At the other sites, the racial composition of the cohort reflected that of the community. The baseline examinations (Visit 1) were conducted between 1987 and 1989; Visit 2 was held between 1990 and 1992; Visit 3 between 1993 and 1995; and Visit 4 was conducted between 1996 and 1998. A fifth clinic examination started in 2011. Of participants still alive at the time of the three follow-up visits to date, response rates for visits 2, 3, and 4 were 93, 86, and 81%, respectively. After the baseline exam, ARIC cohort members were contacted annually by telephone (even during the years in which they also had a clinical exam) to establish vital status and assess a history of cardiovascular disease, including hospitalizations.

Genotyping and QC description

Genotyping was performed at the Broad Institute of MIT and Harvard using the Affymetrix SNP Array 6.0. Genotyping, quality control, and imputation procedures for the ARIC genome-wide association study have previously been described in detail [9].

Statistical methods

Analyses were conducted using male self-reported Caucasian participants in the ARIC cohort with available GWA study data (N=4407). All participants were followed through 2006 for diabetes and prostate cancer, which is the most recent data available on incident cancer outcomes. Individuals with a history of prostate cancer or prevalent diabetes at the baseline examination were excluded from analysis. We also restricted analyses to individuals who had sufficient follow-up information to determine incidence of both prostate cancer and type 2 diabetes, leaving us with 3822 individuals for analysis.

Incident type 2 diabetes was defined as a self-reported physician diagnosis obtained by interviewer-administered questionnaire. Interviews were conducted at each of the in-person visits (though 1996-1998), and thereafter annually by phone. Incident prostate cancer outcomes were ascertained by linkage to the following cancer registries: the Minnesota Cancer Surveillance System, the North Carolina Cancer Registry, the Washington County (Maryland) Cancer Registry, and the (statewide) Maryland Cancer Registry. Cohort identifiers were linked to each cancer registry's database to obtain data regarding cancer occurrence, primary site, and diagnosis date. In addition to a search of cancer registries, the ARIC study asked participants to report all hospitalizations, and hospital surveillance was carried out in each community and cancer-related hospital discharges not identified by cancer registries were retrieved in each community.

To analyze diabetes events for the GWA study, we used Cox proportional hazard models to calculate hazard ratios and corresponding 95% confidence intervals using ProbABEL and assuming an additive genetic model [10]. Cox models were adjusted for age at baseline and field site. For incident diabetes, time to event was defined as the date of the interview at which the participant first reported a diagnosis of diabetes. Participants who did not report diabetes during follow-up were censored at the date of the last interview.

To create polygenic risk scores for type 2 diabetes, we reduced the number of SNPs available for analysis by filtering on minor allele frequency (MAF), genotyping rate, and linkage disequilibrium independent of their association with type 2 diabetes. Specifically, we selected a sample of SNPs with a MAF of ≥5%, a genotyping rate threshold of ≥99%, and a pairwise r^2 threshold of <0.25 within a 200-SNP sliding window [11]. Focusing the analysis on a subset of SNPs in approximate linkage equilibrium ensured the score represents the aggregate effect of a large number of independent SNPs [11]. After pruning, there were 99,966 SNPs out of 2,438,031 SNPs available for analysis.

Next we obtained sets of alleles that were associated with type 2 diabetes at increasingly liberal thresholds (P_T<0.05, 0.25, and 0.5) in Cox regression. For each individual, we calculated the sum of the number of score alleles they had, weighted by the allele-specific log hazard ratio estimated from the GWA for diabetes in ARIC. There were 5,119 alleles at the 0.05 threshold, 25,358 at 0.25, and 50,058 at 0.50.

Proc Score in SAS was used to calculate the scores (SAS Institute Inc., Version 9.2, Cary, NC). We used Cox regression to assess whether the aggregate polygenic risk scores for type 2 diabetes were associated with prostate cancer risk. We modeled the sets of score alleles as both a continuous variable, to estimate its association with prostate cancer under a linear assumption, and in quintiles, to explore the dose-response relationship.

Finally, we created a weighted polygenic risk score by adding together the number of genotyped or imputed risk alleles of 58 genes or regions (**Table 1**). The selection of these 58 genetic variants was based on a recent large-scale association analysis of European-Americans that combined genome-wide association data from multiple studies to identify genetic variants associated with type 2 diabetes [12]. The score weights the dosage for each SNP by the odds ratio reported for the combined DIAGRAM consortium analysis [12]. This risk score was also assessed using Cox regression, modeling the score both continuously and in quintiles. The creation of polygenic risk scores should significantly increase the statistical power necessary to detect associations with prostate cancer, as many of the loci with weak individual effects are more likely to be significantly associated with an outcome when combined into a risk score [13].

Pathway analysis

We performed GWA analyses for incident type 2 diabetes and incident prostate cancer and conducted a pathway analysis of the top signals using the Meta-Analysis Gene-set Enrichment of Variant Associations (MAGENTA) program, which queries for gene set enrichments in pathways found in the Gene Ontology (GO), Panther, Kyoto Encyclopedia of Genes and Genomes (KEGG), Reactome, and BioCarta pathway databases [14]. Because there are multiple comparisons of pathways within a database, we applied a Bonferroni correction for the number of pathways considered within each database and used that p-value as the significance threshold. MAGENTA produces a list of pathways for each disease outcome and their associated GSEA p-value. To determine if there was overlap in pathways between disease outcomes we qualitatively compared the top five most significant pathways associated with type 2 diabetes with the top five pathways for prostate cancer.

Results

There were 774 incident cases of self-reported, physician diagnosed type 2 diabetes and 373 incident cases of prostate cancer in the Caucasian ARIC males with GWA data (**Table 2**). The median follow-up time was 17.9 for type 2 diabetes and 17.8 years for prostate cancer. There were 80 individuals who had both events. The rate of type 2 diabetes was 13 per 1,000 person-years and the rate of prostate cancer was 6 per 1,000 person-years. Men with and without incident diabetes did not significantly differ in age at baseline; however, men with prostate cancer were statistically significantly older at baseline compared to those without prostate cancer (**Table 2**).

Diabetes polygenic risk scores were not statistically significantly associated with incident prostate cancer when modeled linearly (**Table 3**). Likewise, the score quintiles were largely unassociated with prostate cancer across all of the significance thresholds (p for trend=0.16; 0.07; 0.11; 0.19 for P_t <0.05, 0.25, and 0.5 and the 58 SNP risk score, respectively) (**Table 4**).

Table 1. Type 2 diabetes susceptibility loci used to construct the 58 SNP risk score

SNP	Chromosome	Locus	Position (Build 36 bp)	Risk Alleles	Other
rs10923931	1	*NOTCH2*	120,319,482	T	G
rs2075423	1	*PROX1*	212,221,342	G	T
rs780094	2	*GCKR*	27,594,741	C	T
rs10203174	2	*THADA*	43,543,534	C	T
rs243088	2	*BCL11A*	60,422,249	T	A
rs7569522	2	*RBMS1*	161,054,693	A	G
rs13389219	2	*GRB14/COBLL1*	165,237,122	C	T
rs2943640	2	*IRS1*	226,801,829	C	A
rs1801282	3	*PPARG*	12,368,125	C	G
rs1496653	3	*UBE2E2*	23,429,794	A	G
rs6795735	3	*ADAMTS9*	64,680,405	C	T
rs11717195	3	*ADCY5*	124,565,088	T	C
rs4402960	3	*IGF2BP2*	186,994,381	T	G
rs17301514	3	*ST64GAL1*	188,096,103	A	G
rs4458523	4	*WFS1*	6,340,887	G	T
rs459193	5	*MAP3K1/ANKRD55*	55,842,508	G	A
rs6878122	5	*ZBED3*	76,463,067	G	A
rs7756992	6	*CDKAL1*	20,787,688	G	A
rs17168486	7	*DGKB*	14,864,807	T	C
rs849135	7	*JAZF1*	28,162,938	G	A
rs10278336	7	*GCK*	44,211,888	A	G
rs13233731	7	*KLF14*	130,088,229	G	A
rs516946	8	*GOLGA7/ANK1*	41,638,405	C	T
rs7845219	8	*TP53INP1*	96,006,678	T	C
rs3802177	8	*SLC30A8*	118,254,206	G	A
rs16927668	9	*PTPRD*	8,359,533	T	C
rs10811661	9	*CDKN2A/B*	22,124,094	T	C
rs17791513	9	*TLE4*	81,095,410	A	G
rs2796441	9	*TLE1*	83,498,768	G	A
rs11257655	10	*CDC123/CAMK1D*	12,347,900	T	C
rs12242953	10	*VPS26A*	70,535,348	G	A
rs12571751	10	*ZMIZ1*	80,612,637	A	G
rs1111875	10	*HHEX/IDE*	94,452,862	C	T
rs7903146	10	*TCF7L2*	114,748,339	T	C
rs2334499	11	*HCCA2*	1,653,425	T	C
rs163184	11	*KCNQ1*	2,803,645	G	T
rs5215	11	*KCNJ11*	17,365,206	C	T
rs1552224	11	*ARAP1 (CENTD2)*	72,110,746	A	C
rs10830963	11	*MTNR1B*	92,348,358	G	C
rs11063069	12	*CCND2*	4,244,634	G	A
rs10842994	12	*PPFIBP1/KLHDC5*	27,856,417	C	T
rs2261181	12	*HMGA2*	64,498,585	T	C
rs7955901	12	*TSPAN8/LGR5*	69,719,560	C	T
rs12427353	12	*HNF1A (TCF1)*	119,911,284	G	C
rs1359790	13	*SPRY2*	79,615,157	G	A
rs4502156	15	*C2CD4A*	60,170,447	T	C
rs7177055	15	*HMG20A*	75,619,817	A	G
rs11634397	15	*ZFAND6*	78,219,277	G	A
rs2007084	15	*AP3S2*	88,146,339	G	A

rs12899811	15	*PRC1*	89,345,080	G	A
rs9936385	16	*FTO*	52,376,670	C	T
rs7202877	16	*WDR59/CTRB1*	73,804,746	T	G
rs2447090	17	*SRR*	2,245,724	A	G
rs4430796	17	*HNF1B (TCF2)*	33,172,153	G	A
rs12970134	18	*MC4R*	56,035,730	A	G
rs10401969	19	*ARMC6/SF4*	19,268,718	C	T
rs8108269	19	*GIPR/CD3EAP*	50,850,353	G	T
rs4812829	20	*HNF4A*	42,422,681	A	G

Table 2. Participant characteristics by incident diabetes and prostate cancer

	Males	Diabetes (N=774)	No Diabetes (N=3048)	*p*-value
Baseline	Age (years)	54.1	54.5	0.08
	Center: Forsyth County, NC	30.3	28.9	0.0008
	Minneapolis, MN	31.9	38.9	
	Washington County, MD	37.9	32.2	
		Prostate Cancer (N=373)	No Prostate Cancer (N=3449)	*p*-value
Baseline	Age (years)	56.5	54.2	<0.0001
	Center: Forsyth County, NC	29.2	29.2	0.85
	Minneapolis, MN	38.6	37.3	
	Washington County, MD	32.2	33.5	

Table 3. Adjusted associations between the type 2 diabetes polygenic risk scores and incident prostate cancer

Significance threshold	HR	95% Confidence Intervals	*P*-value
0.05	1.03	0.94-1.14	0.40
0.25	1.03	0.94-1.14	0.51
0.50	1.03	0.93-1.13	0.57
58 SNP	0.96	0.87-1.07	0.46

*Hazard ratios are per 1 standard deviation for the genetic score (SD=18.1, 44.8, and 56.8 risk for P_t 0.05, 0.25, and 0.50 and 4.5 for the 58 SNP risk score , respectively).

Table 5 shows the top five pathways identified by GSEA from the six pathway databases using the results of the prostate cancer GWA analysis. None of the biological pathways were statistically significantly associated with prostate cancer. **Table 6** shows the top five pathways identified by GSEA from six pathway databases using the results of the type 2 diabetes GWA analysis. There was only one biological pathway that was statistically significantly associated with type 2 diabetes incidence after the Bonferroni correction for multiple testing, the growth hormone signaling pathway from the BioCarta database (p=0.0001). However, this pathway was not significantly associated with prostate cancer (p=0.43). When type 2 diabetes gene set enrichment analysis p-values generated by MAGENTA were compared to prostate cancer, for each of the top type 2 diabetes pathways found in **Table 6**, none of the pathways were associated with prostate cancer (p>0.05).

Table 7 compares GSEA results for this study to the top five pathways for type 2 diabetes identified by Perry et al. in a previous analysis [15]. None of the top five pathways reported by Perry et al. (p<=0.005) was found to be associated with type 2 diabetes in our analysis (p>0.22) (13). **Table 8** shows that of the top five pathways found in the Genetic Database of Diabetes Mellitus (DMBase) analysis, only the growth hormone signaling pathway (p=0.0001) was statistically significantly associated with type 2 diabetes in our analysis [16].

Discussion

This purpose of this study was to determine if a type 2 diabetes polygenic risk score was also

Table 4. Adjusted associations between the type 2 diabetes polygenic risk scores, modeled as quintiles, and incident prostate cancer

	Quintile	HR	95% CI	*p*-value	P for trend
0.05	1	1.00	ref		0.16
	2	1.00	(0.72-1.40)	0.99	
	3	1.45	(1.06-1.98)	0.02	
	4	1.21	(0.86-1.71)	0.27	
	5	1.19	(0.86-1.65)	0.28	
0.25	1	1.00	ref		0.07
	2	1.01	(0.73-1.42)	0.92	
	3	1.34	(0.97-1.86)	0.07	
	4	1.44	(1.03-2.01)	0.04	
	5	1.21	(0.87-1.67)	0.26	
0.50	1	1.00	ref		0.11
	2	1.14	(0.83-1.60)	0.41	
	3	1.32	(0.95-1.84)	0.10	
	4	1.41	(1.01-1.99)	0.04	
	5	1.23	(0.89-1.71)	0.21	
58 SNP	1	1.00	ref		0.19
	2	0.83	(0.61-1.15)	0.26	
	3	0.82	(0.60-1.12)	0.21	
	4	0.88	(0.64-1.20)	0.41	
	5	0.78	(0.57-1.08)	0.14	

associated with incident prostate cancer; in addition, to conduct pathway analyses of both disease outcomes and identify shared biochemical pathways. Type 2 diabetes polygenic risk scores were not significantly associated with prostate cancer incidence in men in the same study population. None of the top five pathways most significantly associated with type 2 diabetes in gene set enrichment analysis, from each of the six pathway databases queried, were significantly associated with prostate cancer. Nor were the top five type 2 diabetes pathways associated with prostate cancer. Only one pathway was statistically significantly associated with type 2 diabetes after the Bonferroni correction, the growth hormone signaling pathway in the BioCarta database.

To date, the identification of variants associated with both outcomes has been limited and inconsistent. *HNF1B* is the only gene associated with both type 2 diabetes and prostate cancer that has been replicated across studies [6]. Our own study failed to detect genetic pathways shared between the two diseases. Wu et al. found no association between type 2 diabetes and prostate cancer risk, concluding that the inverse association between the two outcomes is the result of detection bias whereby, individuals with type 2 diabetes have attenuated prostate specific antigen, which results in a less frequent diagnosis of prostate cancer [17]. While this may not fully explain the association between the two outcomes, the number of genetic variants associated with both outcomes is limited and at this point in time does not contribute meaningfully to explaining shared etiology. Future analyses should consider alternative explanations for the association between these two diseases such as unmeasured confounding, metabolic and hormonal changes, or the effects of diabetes treatment.

The growth hormone signaling pathway is a biologically plausible candidate pathway for type 2 diabetes. There are 28 genes identified in the growth hormone signaling pathway including *HRAS, HNF1A, GRB2, STAT5A, STAT5B, SRF, SLC2A4, INS, SOS1, PIK3CA, SHC1, INSR, PIK3R1, GHR, PRKCA, PIK3CG, PTPN6, MAP2K1, SOCS1, RAF1, IRS1, PRKCB, MAPK1, GH1, RPS6KA1, PLCG1, MAPK3,* and *JAK2.* There were 27 out of these 28 genes represented in our GWA study and 2423 single nucleotide polymorphisms (SNPs) from these 27 genes were analyzed by MAGENTA. None of the individual SNPs had p-values that approached a Bonferroni corrected statistical threshold, lending support to the idea that when one combines nominally significant variants into biological pathways one may have greater statistical power to detect sets of variants associated with type 2 diabetes [18].

A number of these genes in the growth hormone signaling pathway have variants that have previously been found to be associated with type 2 diabetes. Specifically, *HNF1A* has both rare mutations resulting in monogenic forms of diabetes, in addition to common variants that predispose individuals to multifactorial diabetes [19]. *INS* has a variable number tandem repeat that has been proposed to exert pleiotropic effects on both birth weight and diabetes susceptibility [20]. A recent large-scale candidate gene association study found variants of *INS* and *SOS* were significantly associated with type 2 diabetes [18]. Heterozygous

Table 5. Top five most significant prostate cancer gene set enrichment analysis (GSEA) results for six pathway databases

Database	Number of pathways queried	Bonferroni corrected *p*-value	Pathway	*P*-value for association with incident type 2 diabetes from GSEA	*P*-value for association with incident prostate cancer from GSEA
GO	1778	0.00003	Endosome membrane	0.160	0.004
	1778	0.00003	Ubiquitin-specific protease activity	0.237	0.014
	1778	0.00003	Mitochondrial intermembrane space	0.923	0.015
	1778	0.00003	Erythrocyte differentiation	0.752	0.015
	1778	0.00003	NAD or NADH binding	0.450	0.015
Panther	527	0.0001	Protein complex assembly	0.180	0.003
	527	0.0001	Other cell adhesion molecule	0.710	0.007
	527	0.0001	Phospholipase	0.048	0.025
	527	0.0001	Neuronal activities	0.060	0.032
	527	0.0001	Zinc finger transcription factor	0.190	0.036
Ingenuity	81	0.0006	NRF2-mediated oxidative stress response	0.648	0.049
	81	0.0006	Axonal guidance signaling	0.993	0.105
	81	0.0006	Wnt beta-catenin signaling	0.992	0.249
	81	0.0006	Nitric oxide signaling in the cardiovascular system	0.078	0.257
	81	0.0006	Leukocyte extravasation signaling	0.918	0.362
KEGG	186	0.0003	Wnt signaling pathway	0.429	0.005
	186	0.0003	Vascular smooth muscle contraction	0.033	0.033
	186	0.0003	Oocyte meiosis	0.444	0.040
	186	0.0003	Glycerophospholipid metabolism	0.061	0.049
	186	0.0003	Alpha linolenic acid metabolism	0.190	0.053
BioCarta	214	0.0002	ALK pathway	0.377	0.004
	214	0.0002	BAD pathway	0.536	0.058
	214	0.0002	GCR pathway	0.272	0.058
	214	0.0002	CREB pathway	0.026	0.058
	214	0.0002	AGPCR pathway	0.603	0.062
Reactome	430	0.0001	Transmission across chemical synapses	0.014	0.056
	430	0.0001	Prefoldin mediated transfer of substrate to CCT TRIC	0.399	0.061
	430	0.0001	Membrane trafficking	0.357	0.070
	430	0.0001	Regulation of insulin secretion by glucagon like peptide 1	0.844	0.070
	430	0.0001	Neurotransmitter receptor binding and downstream transmission in postsynaptic cell	0.003	0.111

INSR mutations are the most common cause of monogenic insulin resistance and a recent study identified *INSR* haploinsufficiency is associated with severe insulin resistance and dysglycemia [19]. A polymorphism of *GHR* exon 3 has been found to be associated with type 2 diabetes and a recent GWA study identified a variant of *IRS1* associated with type 2 diabetes risk and this has been replicated by a second study by Yiannakouris et al. [22, 23]. Finally, the BioCarta growth hormone signaling pathway is also one of 19 pathways identified by the DMBase as being statistically significantly associated with type 2 diabetes (p=0.0000013). DMBase is an integrated web-based genetic information resource for diabetes mellitus

Table 6. Top five most significant type 2 diabetes gene set enrichment analysis (GSEA) results for six pathway databases

Database	Number of pathways queried	Bonferroni corrected *p*-value	Pathway	*P*-value for association with incident type 2 diabetes from GSEA	*P*-value for association with incident prostate cancer from GSEA
GO	1778	0.00003	Single-stranded DNA binding	0.0008	0.683
	1778	0.00003	Endocytic vesicle membrane	0.001	1
	1778	0.00003	Nuclear-transcribed mRNA catabolic process, nonsense-mediated decay	0.003	0.465
	1778	0.00003	Intracellular signaling cascade	0.003	0.722
	1778	0.00003	Arachidonic acid secretion	0.004	0.250
Panther	527	0.0001	Vision	0.003	0.924
	527	0.0001	Annexin	0.003	1
	527	0.0001	Protein targeting and localization	0.007	0.754
	527	0.0001	Calmodulin related protein	0.007	0.407
	527	0.0001	Chemokine	0.008	0.593
Ingenuity	81	0.0006	Role of BRCA1 in DNA damage response	0.002	1
	81	0.0006	JAK stat signaling	0.020	1
	81	0.0006	Chemokine signaling	0.028	1
	81	0.0006	14-3-3 mediated signaling	0.057	0.576
	81	0.0006	FGF signaling	0.059	1
KEGG	186	0.0003	VEGF signaling pathway	0.002	0.608
	186	0.0003	Acute myeloid leukemia	0.006	1
	186	0.0003	Chemokine signaling pathway	0.006	0.425
	186	0.0003	FC Gamma R mediated phagocytosis	0.007	0.840
	186	0.0003	FC epsilon RI signaling pathway	0.010	0.273
BioCarta	214	0.0002	GH Pathway	0.0001*	0.439
	214	0.0002	Calcineurin pathway	0.001	1
	214	0.0002	CXCR4 pathway	0.004	0.250
	214	0.0002	BCR pathway	0.006	1
	214	0.0002	CCR3 pathway	0.007	0.246
Reactome	430	0.0001	Insulin synthesis and secretion	0.0003	0.923
	430	0.0001	Regulation of gene expression in Beta cells	0.002	0.873
	430	0.0001	Activation of the pre-replicative complex	0.002	0.569
	430	0.0001	Botulinum neurotoxicity	0.002	1
	430	0.0001	Neurotransmitter receptor binding and downstream transmission in the postsynaptic cell	0.003	0.111

*Statistically significantly associated pathway after Bonferroni correction.

designed to provide genomic variants, genes, and secondary information derived for researchers [16].

There are several limitations to our study. Type 2 diabetes was self-reported. The inclusion of increasing numbers of score alleles with the

Table 7. GSEA p-values for top 5 pathways found in Perry et al. study of biological pathways associated with type 2 diabetes (Perry et al. 2009)

Database	Pathway	Perry et al. *p*-value for association with type 2 diabetes	GSEA *p*-value for association with type 2 diabetes (this study)
KEGG	Wnt signaling pathway	0.0007	0.429
KEGG	Olfactory transduction	0.0009	0.864
GO	Organic acid biosynthetic process	0.004	0.481
GO	Regulation of Wnt receptor signaling pathway	0.005	0.221
GO	Odontogenesis	0.005	0.758

Table 8. GSEA p-values for top five pathways found in DMBase of biological pathways associated with type 2 diabetes (Lee et al. 2011)

Database	Pathway	DMBase p-value for association with type 2 diabetes	GSEA *p*-value for association with type 2 diabetes (this study)
KEGG	Adiopocytokine signaling pathway	0.0000000000092	0.518
KEGG	Type II diabetes mellitus	0.000000011	0.039
KEGG	Insulin signaling pathway	0.000000026	0.028
KEGG	Maturity onset diabetes of the young	0.00000051	0.150
BioCarta	Growth hormone signaling	0.0000013	0.0001*

*Statistically significantly associated pathway after Bonferroni correction.

use of liberal thresholds could be introducing false positives that make it more difficult to discern the signal from the noise. A further limitation is the representation of each gene locus with a single SNP in the pathway analysis, when a disease-associated gene may have multiple functional variants [15]. In the pathway analysis, we could have also have failed to detect more pathways significantly associated with type 2 diabetes or overlap between type 2 diabetes and prostate cancer because (1) the relevant pathways or sets of functionally related genes were not tested; (2) the given distance around the gene may not capture potential signals from more distant transcriptional regulatory elements, such as enhancers or epigenetic marks; (3) rare variants were not tested (4) causal variants are spread across a large number of biological processes making it hard to detect clustering of associations into pathways and/or; (5) the fraction of causal genes in the given gene set may not be significantly higher than the total fraction of causal genes in the genome [14]. Also, the failure to identify shared pathways could be an issue of statistical power, due to the lower number of prostate cancer cases available for analysis compared to diabetes cases (10% versus 20% of the sample).

Finally, comparisons with DMBase must be considered with caution. Lee et al. extracted diabetes genes from the literature and consequently publication bias may exist, whereby non-significant findings remain unpublished, resulting in an artificially inflated magnitude of the effect for well-studied pathways [24]. Simulations have shown that in meta-analyses the use of published studies may over-estimate the effect sizes by as much as 30%, which threatens the validity of literature-based investigations of pathways [25]. Also, if publications from the literature are largely based on candidate gene studies, then pathways known to be biologically relevant from previous studies will be disproportionately represented in the databases used by DMBase.

The strengths of our analysis include the availability of a large group of men in which both prostate cancer and type 2 diabetes outcomes were ascertained in a well-characterized cohort study. Also, we had access to up to 19 years of follow-up data allowing us to prospectively evaluate the association between genes and incident type 2 diabetes and prostate cancer outcomes. Finally, to our knowledge this is the first pathway analysis that looks for common genet-

ic pathways shared between type 2 diabetes and prostate cancer.

In conclusion, polygenic risks scores derived from a GWA of type 2 diabetes in men were not statistically significantly associated with incident prostate cancer in the same study population. In addition, separate pathway analyses of type 2 diabetes and prostate cancer failed to identify pathways significantly associated with both diseases, which could be explained by the different genetic architecture of the diseases, the power of our analyses, or the strength and completeness of the algorithms and pathway databases used. However while we were unable to find pathways shared between type 2 diabetes and prostate cancer, our GWA analysis of type 2 diabetes did identify a pathway statistically significantly associated with type 2 diabetes, the growth hormone signaling pathway, confirming an association with this pathway reported earlier. Additional studies are needed to confirm the association between type 2 diabetes and the growth hormone signaling pathway; in addition, studies are needed to explore the genetic variants that comprise the pathway and how they may influence diabetes risk in isolation or in conjunction with other genes in the pathway.

Acknowledgments

The Atherosclerosis Risk in Communities Study is carried out as a collaborative study supported by National Heart, Lung, and Blood Institute contracts (HHSN268201100005C, HHSN268201100006C, HHSN268201100007C, HHSN268201100008C, HHSN268201100009C, HHSN268201100010C, HHSN268201100011C, and HHSN268201100012C), R01HL087641, R01HL59367 and R01HL086694; National Human Genome Research Institute contract U01HG004402; and National Institutes of Health contract HHSN268200625226C. The authors thank the staff and participants of the ARIC study for their important contributions. Infrastructure was partly supported by Grant Number UL1RR025005, a component of the National Institutes of Health and NIH Roadmap for Medical Research. Cancer incidence data were provided by the Maryland Cancer Registry, Center for Cancer Surveillance and Control, Department of Health and Mental Hygiene, 201 W. Preston Street, Room 400, Baltimore, MD 21201, http://fha.maryland.gov/cancer/mcr.home.cfm, 410-767-4055. We acknowledge the State of Maryland, the Maryland Cigarette Restitution Fund, and the National Program of Cancer Registries of the Centers for Disease Control and Prevention for the funds that support the collection and availability of the cancer registry data.

Conflict of interest statement

Nothing to declare.

Address correspondence to: LA Raynor, Division of Academic General Pediatrics, Department of Pediatrics, University of Minnesota, 717 Delaware St. SE Suite 353, 3rd Floor West, Minneapolis, MN 55414. Phone: 612-598-1148; Fax: 612-626-2134; E-mail: rayno007@umn.edu

References

[1] Bonova S, Filioussi K and Tsantes A. Diabetes mellitus and risk of prostate cancer: a meta-analysis. Diabetologia 2004; 47: 1071-1078.

[2] Kasper JS and Giovannucci E. A meta-analysis of diabetes mellitus and risk of prostate cancer. Cancer Epidemiol Biomarkers Prev 2006; 15: 2056-2062.

[3] Frayling TM, Colhoun H and Florez JC. A genetic link between type 2 between prostate cancer. Diabetologia 2008; 51: 1757-1760.

[4] Pierce BL and Ahsan H. Genetic susceptibility to type 2 diabetes is associated with reduced prostate cancer risk. Hum Hered 2010; 69: 193-201.

[5] Agalliu I, Suuriniemi M, Prokunina-Olsson L, Johanneson B, Collins FS, Standord JL and 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.

[6] Stevens VL, Ahn J, Sun J, Jacobs EJ, Moore SC, Patel AV, Berndt SI, Albanes D and Hayes RB. HNF1B and JAZF1 genes, diabetes, and prostate cancer risk. Prostate 2010; 70: 601-607.

[7] Meyer TE, Boerwinkle E, Morrison AC, Volcik KA, Sanderson M, Coker AL, Pankow JS and Folsom AR. Diabetes genes and prostate cancer in the Atherosclerosis Risk in Communities Study. Cancer Epidemiol Biomarkers Prev 2010; 19: 558-565.

[8] The ARIC Investigators. The Atherosclerosis Risk in Communities (ARIC) Study: design and objectives. Am J Epidemiol 1989; 129: 687-702.

[9] Rasmussen-Torvik LJ, Alonso A, Li M, Kao W, Kottgen A, Yan Y, Couper D, Boerwinkle E, Bielinski SJ and Pankow JS. Impact of repeated

measures and sample selection on genome-wide association studies of fasting glucose. Genet Epidemiol 2010; 34: 665-673.

[10] Aulechenko YS, Struchalin MV and van Duijn CM. ProbABEL package for genome-wide association analysis of imputed data. BMC Bioinformatics 2010; 11: 134.

[11] International Schizophrenia Consortium, Purcell SM, Wray NR, Stone JL, Visscher PM, O'Donovan MC, Sullivan PF and Sklar P. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. Nature 2009; 460: 748-752.

[12] The Diabetes Genetics Replication And Meta-analysis (DIAGRAM) Consortium, Morris AP, Voight BF, Teslovich TM, Ferreira T, Segrè AV, Steinthorsdottir V, Strawbridge RJ, Khan H, Grallert H, Mahajan A, Prokopenko I, Kang HM, Dina C, Esko T, Fraser RM, Kanoni S, Kumar A, Lagou V, Langenberg C, Luan J, Lindgren CM, Müller-Nurasyid M, Pechlivanis S, Rayner NW, Scott LJ, Wiltshire S, Yengo L, Kinnunen L, Rossin EJ, Raychaudhuri S, Johnson AD, Dimas AS, Loos RJ,Vedantam S, Chen H, Florez JC, Fox C, Liu CT, Rybin D, Couper DJ, Kao WH, Li M, Cornelis MC, Kraft P, Sun Q, van Dam RM, Stringham HM, Chines PS, Fischer K,Fontanillas P, Holmen OL, Hunt SE, Jackson AU, Kong A, Lawrence R, Meyer J, Perry JR, Platou CG, Potter S, Rehnberg E, Robertson N, Sivapalaratnam S, Stančáková A, Stirrups K, Thorleifsson G, Tikkanen E, Wood AR, Almgren P, Atalay M, Benediktsson R, Bonnycastle LL, Burtt N, Carey J, Charpentier G, Crenshaw AT, Doney AS, Dorkhan M, Edkins S, Emilsson V, Eury E, Forsen T, Gertow K, Gigante B, Grant GB, Groves CJ, Guiducci C, Herder C, Hreidarsson AB, Hui J, James A, Jonsson A, Rathmann W, Klopp N, Kravic J, Krjutškov K, Langford C, Leander K, Lindholm E, Lobbens S, Männistö S, Mirza G, Mühleisen TW, Musk B, Parkin M, Rallidis L, Saramies J, Sennblad B, Shah S, Sigurðsson G, Silveira A, Steinbach G, Thorand B, Trakalo J, Veglia F, Wennauer R, Winckler W, Zabaneh D, Campbell H, van Duijn C, Uitterlinden AG, Hofman A, Sijbrands E, Abecasis GR, Owen KR, Zeggini E, Trip MD, Forouhi NG, Syvänen AC, Eriksson JG, Peltonen L, Nöthen MM, Balkau B, Palmer CN, Lyssenko V, Tuomi T, Isomaa B, Hunter DJ, Qi L; Wellcome Trust Case Control Consortium; Meta-Analyses of Glucose and Insulin-related traits Consortium (MAGIC) Investigators; Genetic Investigation of ANthropometric Traits (GIANT) Consortium; Asian Genetic Epidemiology Network–Type 2 Diabetes (AGEN-T2D) Consortium; South Asian Type 2 Diabetes (SAT2D) Consortium, Shuldiner AR, Roden M, Barroso I, Wilsgaard T, Beilby J, Hovingh K, Price JF, Wilson JF, Rauramaa R, Lakka TA, Lind L, Dedoussis G, Njølstad I, Pedersen NL, Khaw KT, Wareham NJ, Keinanen-Kiukaanniemi SM, Saaristo TE, Korpi-Hyövälti E, Saltevo J, Laakso M, Kuusisto J, Metspalu A, Collins FS, Mohlke KL, Bergman RN, Tuomilehto J, Boehm BO, Gieger C, Hveem K, Cauchi S, Froguel P, Baldassarre D, Tremoli E, Humphries SE, Saleheen D, Danesh J, Ingelsson E, Ripatti S, Salomaa V, Erbel R, Jöckel KH, Moebus S, Peters A, Illig T, de Faire U, Hamsten A, Morris AD, Donnelly PJ, Frayling TM, Hattersley AT, Boerwinkle E, Melander O, Kathiresan S, Nilsson PM, Deloukas P, Thorsteinsdottir U, Groop LC, Stefansson K, Hu F, Pankow JS, Dupuis J, Meigs JB, Altshuler D, Boehnke M and McCarthy MI. Large-scale association analysis provides insight into the genetic architecture and pathophysiology of type 2 diabetes. Nat Genet 2012; 44: 981-990.

[13] Fontaine-Bisson B, Renstrom F, Rolandsson O, MAGIC, Payne F, Hallmans G, Barroso I and Franks PW. Evaluating the discriminative power of multi-trait genetic risk scores for type 2 diabetes in a northern Swedish population. Diabetologia 2010; 53: 2155-2162.

[14] Segre AV, DIAGRAM Consortium, MAGIC Investigators, Groop L, Mootha VK, Daly MJ and Altshuler D. Common inherited variation in mitochondrial genes is not enriched for associations with type 2 diabetes or related glycemic traits. PLoS Genetics 2010; 6: e1001058.

[15] Perry JR, McCarthy MI, Hattersley AT, Zeggini E, Wellcome Trust Case Control Consortium, Weedon MN and Frayling TM. Interrogating type 2 diabetes genome-wide association data using a biological pathway-based approach. Diabetes 2009; 58: 1463-1467.

[16] Lee SY, Park Y, Kim J and Kim YJ. DMBase: an integrated genetic information. Interdisciplinary Bio Central 2011; 3: 1-4.

[17] Wu C, Moreira DM, Gerber L, Rittmaster RS, Andriole GL and Freedland SJ. Diabetes and prostate cancer risk in the REDUCE trial. Prostate Cancer Prostatic Dis 2011; 14: 326-331.

[18] Baranzini SE, Galwey NW, Wang J, Khankhanian P, Lindberg R, Pelletier D, Wu W, Uitdehaag BM, Kappos L, GeneMSA Consortium, Polman CH, Matthews PM, Hauser SL, Gibson RA, Oksenberg JR and Barnes MR. Pathway and network-based analysis of genome-wide association studies in multiple sclerosis. Hum Mol Genet 2009; 18: 2078-2090.

[19] Voight BF, Scott LJ, Steinthorsdottir V, Morris AP, Dina C, Welch RP, Zeggini E, Huth C, Aulchenko YS, Thorleifsson G, McCulloch LJ, Ferreira T, Grallert H, Amin N, Wu G, Willer CJ, Raychaudhuri S, McCarroll SA, Langenberg C, Hofmann OM, Dupuis J, Qi L, Segrè AV, van Hoek M, Navarro P, Ardlie K, Balkau B, Bene-

diktsson R, Bennett AJ, Blagieva R, Boerwinkle E, Bonnycastle LL, Bengtsson Boström K, Bravenboer B, Bumpstead S, Burtt NP, Charpentier G, Chines PS, Cornelis M, Couper DJ, Crawford G, Doney AS, Elliott KS, Elliott AL, Erdos MR, Fox CS, Franklin CS, Ganser M, Gieger C, Grarup N, Green T, Griffin S, Groves CJ, Guiducci C, Hadjadj S, Hassanali N, Herder C, Isomaa B, Jackson AU, Johnson PR, Jørgensen T, Kao WH, Klopp N, Kong A, Kraft P, Kuusisto J, Lauritzen T, Li M, Lieverse A, Lindgren CM, Lyssenko V, Marre M, Meitinger T, Midthjell K, Morken MA, Narisu N, Nilsson P, Owen KR, Payne F, Perry JR, Petersen AK, Platou C, Proença C, Prokopenko I, Rathmann W, Rayner NW, Robertson NR, Rocheleau G, Roden M, Sampson MJ, Saxena R, Shields BM, Shrader P, Sigurdsson G, Sparsø T, Strassburger K, Stringham HM, Sun Q, Swift AJ, Thorand B, Tichet J, Tuomi T, van Dam RM, van Haeften TW, van Herpt T, van Vliet-Ostaptchouk JV, Walters GB, Weedon MN, Wijmenga C, Witteman J, Bergman RN, Cauchi S, Collins FS, Gloyn AL, Gyllensten U, Hansen T, Hide WA, Hitman GA, Hofman A, Hunter DJ, Hveem K, Laakso M, Mohlke KL, Morris AD, Palmer CN, Pramstaller PP, Rudan I, Sijbrands E, Stein LD, Tuomilehto J, Uitterlinden A, Walker M, Wareham NJ, Watanabe RM, Abecasis GR, Boehm BO, Campbell H, Daly MJ, Hattersley AT, Hu FB, Meigs JB, Pankow JS, Pedersen O, Wichmann HE, Barroso I, Florez JC, Frayling TM, Groop L, Sladek R, Thorsteinsdottir U, Wilson JF, Illig T, Froguel P, van Duijn CM, Stefansson K, Altshuler D, Boehnke M, McCarthy MI, MAGIC investigators and GIANT Consortium. Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. Nat Genet 2010; 42: 579-589.

[20] Barroso I, Luan J, Middleberg RP, Harding AH, Franks PW, Jakes RW, Clayton D, Schafer AJ, O'Rahilly S and Wareham NJ. Candidate gene association study in type 2 diabetes indicates a role for genes involved in beta-cell function as well as insulin action. PLoS Biol 2003; 1: E20.

[21] Sulliman SG, Stanik J, McCulloch LJ, Wilson N, Edghill EL, Misovicova N, Gasperikova D, Sandrikova V, Elliott KS, Barak L, Ellard S, Volpi EV, Klimes I and Gloyn AL. Severe insulin resistance and intrauterine growth deficiency associated with haploinsufficiency for INSR and CHN2: new insights into synergistic pathways involved in growth and metabolism. Diabetes 2009; 58: 2954-2961.

[22] Strawbridge RJ, Karvestedt L, Li C, Efendic S, Ostenson CG, Gu HF and Brismar K. GHR exon 3 polymorphism: association with type 2 diabetes mellitus and metabolic disorder. Growth Horm IGF Res 2007; 17: 392-398.

[23] Yiannakouris N, Cooper JA, Shah S, Drenos F, Ireland HA, Stephens JW, Li KW, Elkeles R, Godsland IF, Kivimaki M, Hingorani AD, Kumari M, Talmud PJ and Humphries SE. IRS1 gene variants, dysglycaemic metabolic changes and type-2 diabetes risk. Nutr Metab Cardiovasc Dis 2011; 22: 1024-1030.

[24] Munafo MR, Clark TG and Flint J. Assessing publication bias in genetic association studies: evidence from a recent meta-analysis. Psychiatry Res 2004; 129: 39-44.

[25] Soeken KL and Sripusanapan A. Assessing publication bias in meta-analysis. Nurs Res 2003; 52: 57-60.