

Article

Noncoding Variations in the Gene Encoding Ceramide Synthase 6 are Associated with Type 2 Diabetes in a Large Indigenous Australian Pedigree

David A. Good¹, David L. Duffy², Manuela Good¹, Cheng Xia Guo¹, Frances Busfield¹, Anthony Shaw¹ and Joanne T. E. Shaw^{1,2,*}

¹Diabetes Research Laboratories, School of Clinical Medicine Prince Charles Hospital Northside Clinical Unit, Faculty of Medicine, University of Queensland, Brisbane, Queensland, Australia and ²QIMR Berghofer Medical Research Institute, Brisbane, Queensland, Australia

Abstract

Type 2 diabetes (T2D) is a chronic disease that disproportionately affects Indigenous Australians. We have previously reported the localization of a novel T2D locus by linkage analysis to chromosome 2q24 in a large admixed Indigenous Australian pedigree (Busfield et al. (2002). *American Journal of Human Genetics*, 70, 349–357). Here we describe fine mapping of this region in this pedigree, with the identification of SNPs showing strong association with T2D: rs3845724 (diabetes $p = 7 \times 10^{-4}$), rs4668106 (diabetes $p = 9 \times 10^{-4}$) and rs529002 (plasma glucose $p = 3 \times 10^{-4}$). These associations were successfully replicated in an independent collection of Indigenous Australian T2D cases and controls. These SNPs all lie within the gene encoding ceramide synthase 6 (*CERS6*) and thus may regulate ceramide synthesis.

Keywords: chronic disease; diabetes; genetic association; indigenous health

(Received 20 January 2019; accepted 11 March 2019)

Type 2 diabetes (T2D) is characterized by hyperglycemia due to impaired insulin secretion, insulin resistance in muscle and nonsuppressible hepatic glucose production (DeFronzo, 1988). Genetic factors are highly influential in the development of diabetes. The monogenic forms of diabetes are responsible for approximately 5% of cases and are caused by mutations in genes encoding insulin; the insulin receptor glucokinase; and the transcription factors hepatocyte nuclear factor-1 alpha (HNF-1alpha), HNF-1beta, HNF-4alpha, insulin promoter factor-1 and NeuroD1/BETA2 (Bell et al., 1996; Horikawa et al., 1997; Malecki et al., 1999; Steiner et al., 1995; Stoffers et al., 1997; Taylor, 1995; Yamagata et al., 1996; Yamagata et al., 1996). Mutations in mitochondrial genes can cause diabetes, often in correlation with hearing loss (Maassen & Kadowaki, 1996). Large genomewide association study (GWAS) meta-analyses have now identified over 240 less strongly penetrant risk loci (Mahajan et al., 2018; Suzuki et al., 2019), so that 18% of the total T2D heritability can be attributed to known loci. This still leaves a substantial role for other loci, particularly for Indigenous Australian populations, where there is a dearth of genetic data.

While the overall T2D prevalence among Australians over 25 years of age is considerable at 7.2% (Zimmet & Welborn, 2001), the rates are six times higher in Indigenous Australian communities (Bastian, 1979; Gault et al., 1996). The significant prevalence

of T2D and associated metabolic disturbances in Indigenous Australians are of substantial importance for study as they are major factors contributing to high morbidity and mortality in this group (O'Dea, 1992). Complications of T2D, including nephropathy, retinopathy and foot ulcers, that lead to amputation are common, and the clustering of hyperglycemia with insulin resistance, elevated plasma triglyceride, low high-density lipoprotein cholesterol, hypertension and central obesity also predisposes to macrovascular disease (Guest et al., 1993; Hayman, 1997; O'Dea, 1991; O'Dea et al., 1990).

Epidemiological indications strongly implicate key genetic factors in the predisposition of groups of Indigenous Australians to T2D. A major piece of evidence for the genetic influences that underlie ethnic differences in diabetes prevalence is the inverse connection of diabetes and insulin resistance to European admixture. Among Indigenous Australian groups residing in New South Wales, diabetes was twice as common in inland towns as in coastal towns where European admixture was often present (Williams et al., 1987), and in central Australia, glucose intolerance and 2-hour insulin levels were inversely connected with European admixture (O'Dea et al., 1993). This T2D predisposition is transmitted in a family-specific manner, rather than uniformly present in all Indigenous Australians. However, these genetic effects interact with environmental and social factors such as weight gain and low physical activity, as T2D has remained infrequent among Indigenous Australians who maintain hunter-gatherer lifestyles (O'Dea et al., 1993).

In other populations there have been breakthroughs in this field. In the Greenlandic Inuit population, a common *TBC1D4*

Author for correspondence: Joanne T. E. Shaw, Email: jteshaw@gmail.com

Cite this article: Good DA, Duffy DL, Good M, Xia Guo C, Busfield F, Shaw A, and Shaw JTE. (2019) Noncoding Variations in the Gene Encoding Ceramide Synthase 6 are Associated with Type 2 Diabetes in a Large Indigenous Australian Pedigree. *Twin Research and Human Genetics* 22: 79–87, <https://doi.org/10.1017/thg.2019.13>

variant that confers muscle resistance and T2D (Moltke *et al.*, 2014) was recently reported. Independently within the same population, a separate signal was located in a cluster of fatty acid desaturases that determine omega-3 polyunsaturated fatty acid levels. The associated alleles modulate fatty acid composition, as well as multiple metabolic and anthropometric phenotypes, including weight and height (Fumagalli *et al.*, 2015; Garup *et al.*, 2018). In the Samoan population, a GWAS discovered a variant of *CREBRF* associated with obesity, which when overexpressed selectively decreased energy use and increased fat storage in an adipocyte cell model (Minster *et al.*, 2016).

There are complexities involved in the mapping of diabetes susceptibility genes. Different ethnic groups have distinct population histories, and in each population, different genes are likely to be responsible for disease susceptibility. The genetic analysis of the Indigenous Australian population is likely to make a unique contribution to the understanding of diabetes susceptibility and resistance.

We previously recruited a large multigenerational pedigree of admixed Indigenous Australian descent with a high prevalence of T2D (Busfield *et al.*, 2002). The pedigree consists of 232 living members, 56 of whom are affected with T2D. We conducted a genomewide scan at an average resolution of 10 cM, for T2D susceptibility genes, and undertook linkage analysis and found strong evidence for linkage to chromosome 2q24.3 with a multipoint LOD score of $3.9 < 1$ cM centromeric from marker D2S2345, with a 3-LOD support interval of 18 cM. This interval does contain the known adiposity/T2D locus *COBLL1* (Manning *et al.*, 2012) and fasting glucose locus *G6PC2* (Bouatia-Naji *et al.*, 2008). Secondary peak two-point LOD scores on the genomewide scan were noted on chromosome 8p (LOD score 1.77 at D8S549) and chromosome 3q (LOD score 1.8 at D3S1311; Busfield *et al.*, 2002). In the present article, we describe fine mapping of the 2q24.3 region, with the identification of four SNPs showing strong association with T2D, which all lie within a gene encoding ceramide synthase 6 (*CerS6*).

Limitations

This study is limited by its focused scope on a population of Indigenous Australians, and therefore its findings may not be applicable to broader populations with T2D. However, this research is still highly relevant for endocrine study as it provides insight into the tangible influences that genetics can have on the onset of T2D.

Methods

Subjects

We ascertained clinical details from a community of Indigenous Australian descent in North Stradbroke Island, Queensland. The pedigree consists of 232 living members, 56 of whom are affected with T2D. We have previously reported the clinical details and phenotypic characterization of this pedigree (Busfield *et al.*, 2002). While it was necessary to divide the pedigree for linkage analysis in the previous article, in the present association analyses, we used the complete pedigree, usually incorporating an empirical genetic relationship matrix.

We also recruited a replication sample of 234 unrelated individuals from a number of mainland Indigenous communities in our region. These participants were all of Indigenous descent and well matched in terms of age (52.0 ± 17 pedigree vs. 51.3 ± 16 controls)

and sex; 84 (36%) were individuals with T2D, further confirming the high rates of this disease in Indigenous Australians.

The study protocol was approved by the Princess Alexandra Hospital, the Prince Charles Hospital and University of Queensland Ethics Committee, and all participants provided informed consent before the study. These studies were performed after full consultation and with the approval of the Indigenous Australian communities.

SNP Mapping

Initially, 136 SNPs spanning both chromosome 2 and 3 were selected from the linkage-based minimal candidate region to have a minor allele frequency greater than 0.3 in both the African American and Caucasian populations. In the follow-up genotyping, an additional 61 SNPs for the *CerS6* gene were tagged SNPs selected from the HapMap database (International HapMap Consortium (IHMC), 2005) for comprehensive coverage of the entire *CerS6* locus, and included an additional 70 kbp both up- and downstream of exon 1 and exon 10, respectively. SNP genotyping was carried out at the Australian Genome Research Facility (AGRF), Brisbane, Australia using the homogenous MassExtend (hME — single base extension) assay analyzed on the Sequenom Autoflex Mass spectrometer and the Samsung 24-pin nanodispenser.

Genotype Imputation

The BEAGLE 5.0 package (Browning *et al.*, 2018) was used to impute ungenotyped SNPs in the 2q24 region. As this does not specifically model the pedigree relationships, the 'best-guess' imputed genotypes gave rise to a large number of Mendelian inconsistencies even for SNPs with a $DR^2 = 1$, a fact that must be borne in mind when assessing results from imputed SNPs. To guard against problems, we examined genotyped and imputed genotype association results separately (see Figure 1).

Mutation Screening

We sequenced all 10 exons as well as intron/exon boundaries and 600 base pairs into the putative promoter region of *CERS6* to investigate a causal mutation. In brief, PCR products were initially generated from salted-out lymphocyte DNA (Miller *et al.*, 1988) of 12 individuals to represent each of the three SNP genotypes (T/T, G/T, G/G). PCR reactions contained the following reagents; 10 pmol each primer, 2.5 mM $MgCl_2$, 0.25 mM dNTPs, 0.5 U Red Hot Taq (ABgene) and $1 \times$ reaction buffer IV (ABgene).

PCR products were generated on a Corbett Research Palm-Cycler using the following program: 94°C for 2 mins, and 30 cycles of 94°C for 30 s, 55–72°C for 30 s, 72°C for 40 s; a final 5-min incubation took place at 72°C. PCR products were purified using centrifugation and washing through High Pure PCR gel matrix columns (Roche) and sequenced using ABI dye terminator chemistry using the standard protocol. Washed and precipitated sequence reactions were separated by the AGRF. Sequence data were analyzed using the Chromas software package. All polymorphisms identified were screened in the family members using the aforementioned method.

Statistical Analyses

The SPSS 7.0 and R statistical packages (R Core Team, 2019) were used to carry out standard descriptive and statistical analyses. We estimated the genetic ancestry of study participants using the PC-AIR (which adjusts for relatedness) and PCA packages in R (Conomos *et al.*, 2019) to analyze a set of microsatellite markers

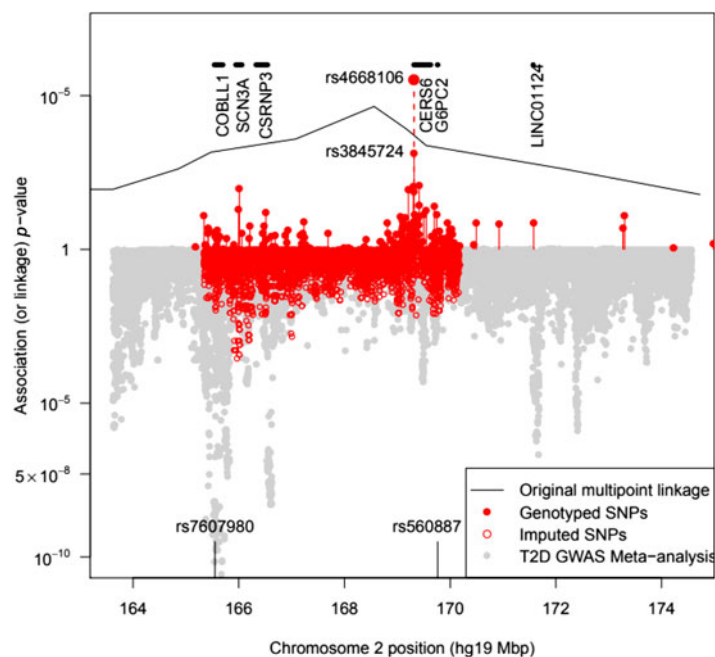


Fig. 1. Plot of the $-\log_{10} p$ value for association of SNPs with T2D in pedigree members. Upper panel shows results for genotyped SNPs and the original multipoint linkage results converted into p values (recall linkage analysis uses a less severe multiple testing threshold). The result for rs4668106 combining the pedigree and replicate samples is labeled. The lower panel shows results for imputed SNPs, along with Fisher-combined p values from two large published GWAS meta-analyses. The position of the published T2D and FPG SNPs rs7607980 in *COLBB1* and rs560887 in *G6PC2* are indicated at the bottom (meta-analysis $p = 10^{-18}$ and 10^{-128} , respectively).

from our earlier study that overlapped with those used in the Human Genome Diversity Project. Allelic association analysis was carried out as a logistic mixed-model regression in the GENESIS package (Chen et al., 2016; Conomos et al., 2019), including an empirical genomic relationship matrix based on the microsatellite and SNP genotype data estimated using the Sib-pair program (Duffy, 1997), and with sex, age and the first principal component from the PC-AIR analysis as fixed effects. We also used MENDEL 16.0 (Lange et al., 2013) to perform individual SNP and haplotype logistic regression-based association analysis, including birth year and sex as covariates. For the replication Indigenous sample, we compared allele frequencies of the control and affected groups using a chi-square test, then combined these data with the pedigree data and carried out a mixed-model analysis, including the sample as a stratifying variable. To identify the most strongly associated observed SNP, we carried out Bayesian model averaging (Raftery et al., 2006).

Association testing of our panel of 197 SNPs in our two intervals on chromosomes 2 and 3 is equivalent to 139 independent statistical tests after adjustment for linkage disequilibrium (Moskvina & Schmidt, 2008). A 5% experiment wide, type-1 error rate is therefore obtained by using a critical p -value threshold of .0004.

We also compared our results to those from two recently published large meta-analyses of T2D in European (Mahajan et al., 2018) and Asian (Suzuki et al., 2019) populations that have made individual SNP association results available. We combined the p values from each study for each overlapping SNP using the Fisher method.

We consulted a number of bioinformatics databases to assess the loci exhibiting association in our results, including the GTEx portal (GTEx Consortium, 2013), and the Pancreatic Islet Regulome Browser (Miguel-Escalada et al., 2018; Mularoni et al., 2017).

Results

The genetic principal components analysis using the previously collected microsatellite markers showed no significant differences between the ancestries of diabetic and unaffected individuals within the pedigree. Cases varied markedly from each other in their scores on the first four principal components.

We genotyped pedigree members for a total of 136 SNPs spanning both chromosomes 2 and 3. Of the 136 SNPs selected, 5 were dropped from the analysis either because of low minor allele frequency or because of Hardy-Weinberg disequilibrium suggesting additional assay problems. This analysis identified a single SNP (rs4668106) that demonstrated strong association with T2D, $p = 4.3 \times 10^{-5}$ (Table 1, Figure 1).

We then proceeded to screen for rs4668106 in a random unrelated cohort of 234 Indigenous Australians, 84 of whom suffered from T2D: rs4668106 showed no evidence of Hardy-Weinberg disequilibrium, while the rs4668106*T allele was shown to be associated with T2D in this cohort ($p = .001$; Table 2).

The SNP rs4668106 is located within intron 8 of the ceramide synthase 6 gene (*CERS6*). An additional 61 SNPs were selected from the HapMap database (IHMC, 2005) suitable for haplotype coverage of the entire *CERS6* locus, as well as approximately 70 kbp of flanking sequence. These additional SNPs were screened for association with T2D. The strongest association was found for the T allele of rs3845724 with a p value of 1.1×10^{-5} , but two other SNPs rs4668102, and rs4668089 exhibited similar strength of association ($p = 4.6 \times 10^{-5}$, $p = 1.3 \times 10^{-5}$; Tables 1 and 3). There is almost perfect disequilibrium between rs4668106 and rs3845724 (Table 4).

In a Bayesian model averaging analysis testing 20 SNPs (larger windows lead to numerical problems) in an interval containing the four best SNPs, the posterior probability of a nonzero log odds ratio for rs4668089 was 100%, while it was 13.6% for the next best SNP,

Table 1. SNPs and their association with type 2 diabetes in members of a large indigenous Australian pedigree

Marker	N all ^a	All ^b	All2 ^c	MAF ^d	N typ ^e	Likelihood ratio		DF ^f	p value	
						test statistic			Mendel	Solar
D2S142	6				121	3.73	0.00	5	.59	
D2S141	8				25	15.95	0.00	7	.03	
D2S156	12				101	0.00	0.00	11	1.00	
D2S2157	10				131	11.69	0.00	9	.23	
D2S124	8				133	5.64	0.00	7	.58	
D2S2330	9				141	8.82	0.00	8	.36	
D2S1395	6				98	10.19	0.00	5	.07	
D2S2345	7				133	6.79	0.00	6	.34	
rs4503959	2	G	T	0.337	166	6.36	5.59	1	.01	.02
rs12233000	2	C	T	0.355	165	3.34	3.55	1	.07	.06
rs7589268	2	G	A	0.476	166	4.29	3.92	1	.04	.05
rs2724157	2	C	T	0.187	163	0.31	1.38	1	.57	.24
rs10497340	2	G	A	0.137	164	0.20	0.40	1	.65	.53
rs1377502	2	G	A	0.369	164	0.09	0.26	1	.77	.61
rs17806793	2	C	T	0.169	154	0.02	0.00	1	.88	.96
rs1124893	2	A	G	0.288	163	0.93	0.74	1	.33	.39
rs1453405	2	G	A	0.460	163	0.39	0.67	1	.53	.41
rs1453406	2	G	A	0.186	161	0.01	0.01	1	.94	.92
rs13007801	2	A	G	0.449	156	2.26	4.56	1	.13	.03
rs964987	2	T	C	0.381	151	2.48	1.33	1	.12	.25
rs17808106	2	T	C	0.259	162	1.23	1.25	1	.27	.26
rs1993224	2	G	A	0.352	162	3.86	4.31	1	.05	.04
rs10930327	2	C	T	0.363	160	0.73	0.66	1	.39	.42
rs1020603	2	T	C	0.409	153	1.66	3.55	1	.20	.06
rs17808500	2	C	T	0.171	152	2.66	2.82	1	.10	.09
rs13398999	2	C	T	0.147	167	0.05	0.03	1	.82	.86
rs13006127	2	T	C	0.476	144	5.61	7.10	1	.02	.01
rs13031884	2	T	A	0.353	163	6.01	3.34	1	.01	.07
rs11695350	2	T	C	0.183	161	3.90	5.32	1	.05	.02
rs17203177	2	G	C	0.285	156	4.95	4.25	1	.03	.04
rs6757628	2	A	G	0.341	160	5.08	4.41	1	.02	.04
rs13398133	2	A	G	0.119	160	0.18	0.01	1	.68	.92
rs4668079	2	G	A	0.374	159	0.01	0.30	1	.94	.58
rs1574437	2	C	T	0.468	155	0.05	0.27	1	.82	.60
rs4668086	2	G	A	0.142	165	2.91	5.21	1	.09	.02
rs4668089	2	G	T	0.206	165	19.02	20.56	1	1x10⁻⁵	6x10⁻⁶
rs1806328	2	C	T	0.308	154	0.55	0.48	1	.46	.49
rs6707065	2	G	C	0.393	164	2.31	4.36	1	.13	.04
rs13391857	2	G	A	0.181	166	0.04	0.25	1	.85	.62
rs1863146	2	G	C	0.271	157	2.75	2.15	1	.10	.14
rs4668097	2	T	G	0.288	163	0.00	0.02	1	.97	.89
rs17249675	2	G	A	0.269	158	0.00	0.07	1	.95	.80
rs7601307	2	T	C	0.469	162	2.45	3.03	1	.12	.08

(Continued)

Table 1. (Continued)

Marker	N all ^a	All ^b	All2 ^c	MAF ^d	N typ ^e	Likelihood ratio			p value	
						test statistic		DF ^f	Mendel	Solar
					Mendel	Solar				
rs11900907	2	T	C	0.419	161	7.33	7.02	1	.01	.01
rs1424937	2	T	C	0.332	164	0.64	1.47	1	.42	.23
rs11681255	2	T	C	0.352	166	0.11	0.22	1	.74	.64
rs17251018	2	G	A	0.240	154	0.15	1.01	1	.70	.31
rs4668102	2	T	C	0.319	157	16.63	11.00	1	4.6x10⁻⁵	9.1x10⁻⁵
rs4668103	2	A	G	0.174	164	2.94	2.68	1	.09	.1
rs4668106	2	G	T	0.355	165	16.74	16.00	1	4.3x10⁻⁵	6.3x10⁻⁵
rs3845724	2	C	T	0.357	164	19.28	17.01	1	1.1x10⁻⁵	3.7x10⁻⁵
rs9287901	2	G	A	0.473	164	4.62	3.85	1	.03	.05
rs8304	2	G	T	0.174	161	2.27	2.05	1	.13	.15
rs727068	2	T	G	0.327	162	8.14	8.97	1	.00	.00
rs13017628	2	A	T	0.127	165	1.21	1.87	1	.27	.17
rs3914920	2	G	T	0.264	161	1.18	2.42	1	.28	.12
rs829950	2	C	T	0.327	159	0.01	0.64	1	.92	.43
rs829958	2	C	T	0.224	161	0.34	0.01	1	.56	.92
rs6433093	2	A	C	0.265	164	1.07	0.06	1	.30	.81
rs830028	2	A	G	0.401	167	0.63	0.30	1	.43	.58
rs558922	2	C	T	0.439	164	0.78	0.07	1	.38	.79
rs10198693	2	A	G	0.148	162	0.02	0.15	1	.88	.70
rs830021	2	T	G	0.221	154	0.16	1.68	1	.69	.19
rs487756	2	T	C	0.466	163	0.31	0.63	1	.57	.43
rs7585565	2	C	T	0.213	150	0.22	0.62	1	.64	.43
rs12386224	2	C	G	0.379	165	0.12	0.17	1	.72	.68
rs531328	2	T	G	0.285	165	0.44	0.15	1	.51	.70
rs532779	2	A	T	0.208	161	0.30	0.30	1	.58	.58
rs529002	2	A	C	0.354	158	7.12	7.05	1	.008	.008
rs11676084	2	A	G	0.108	158	0.49	0.08	1	.48	.78
D2S2177	6				114	15.92	0.00	5	.007	
D2S335	11				137	9.58	0.00	10	.48	

^aNumber of alleles. ^bAllele 1. ^cAllele 2. ^dMinor allele frequency. ^eNumber typed. ^fDegrees of freedom. The SNPs demonstrating a statistically significant relationship are shown in bold type.

rs4668106. The examination of the haplotypes segregating in the family found that the most common haplotype, involving the four best SNPs (TCTT), was increased in diabetics (75%) compared to nondiabetics (50%). However, haplotype was not a significantly better predictor of diabetes than rs4668089.

We tested these SNPs for association to diabetes-associated continuous phenotypes (body mass index and waist hip ratio; Figure 2). Testing the SNPs for association with fasting plasma glucose (FPG) levels indicates that the T/T allele of rs4668089 was associated with affected individuals with very elevated FPG levels. When the SNPs were tested for association with waist circumference, the best evidence for association was with rs4668102 ($p = .003$) (Figure 3).

On examination of the GTEx eQTL database, the best hits for rs4668106 and rs3845724 were for expression levels of *G6PC2* in three tissue types: omentum, skin and tibial nerve. In each case, the common allele was associated with a 0.3 SD increase in

Table 2. Genotype frequencies for rs4668106 in original pedigree and indigenous replication sample and combined association analysis

Group		G/G	G/T	T/T	MAF (T)	Association p value
Pedigree	Unaffecteds N = 112	14	66	32	0.583	4×10^{-5}
	T2D cases N = 52	1	19	32	0.242	
Replication sample	Unaffecteds N = 147	23	83	41	0.439	9×10^{-4}
	T2D cases N = 83	5	37	41	0.198	
Combined association test (MFT mixed model, replicate as covariate)						3×10^{-6}
Test for homogeneity of samples						.66

Table 3. Frequency of top diabetes-associated SNPs

	Effect (other) allele	Effect allele frequency (ASE)				
		T2D cases in pedigree	Unaffected pedigree members	MGRB ^a N = 2,840	gnomAD NFE	GnomAD East Asia
rs4668089	T (G)	0.851 (0.053)	0.781 (0.040)	0.755	0.752	0.732
rs4668102	C (T)	0.729 (0.068)	0.625 (0.057)	0.474	0.447	0.608
rs4668106	T (G)	0.717 (0.067)	0.561 (0.047)	0.434	0.413	0.553
rs3845724	T (C)	0.735 (0.066)	0.581 (0.057)	0.446	0.424	0.554

Note: P-values are in previous table.

^aMGRB is the Sydney Genomics Collaborative Medical Genomics Reference Bank.

G6PC2 transcript levels (best p value 3×10^{-7} for rs3845724 in nerve). Known regulatory variants within *G6PC2* are strongly associated with FPG level (Bouatia-Naji et al., 2008). When we look up results for either of these SNPs in the DIAGRAM and Japan BioBank T2D GWAS meta-analyses (Mahajan et al., 2018; Suzuki et al., 2019), neither is significantly associated in those populations (all p values $> .3$). However, this is also the case for rs560887 in *G6PC2*, a SNP associated with FPG at $p = 10^{-218}$ in a recent meta-analysis (Depuis et al., 2010).

As the identified SNPs are intronic, we sequenced all coding regions of this gene to rule out coding polymorphisms in linkage disequilibrium with the SNPs. Sequence analysis of all 10 exons, 800 bp upstream of exon 1 and all intron/exon boundaries in four affected and four unaffected family members identified a number of genetic variations. These were screened in all pedigree members; however, none were found to be associated with T2D.

Discussion

We have detected association between common SNP alleles intronic to *CERS6* and risk of T2D at a statistically significant level (rs4668106 unadjusted $p = 3 \times 10^{-6}$, adjusted for 139 effective tests $p = 4.2 \times 10^{-4}$). Ceramide synthase 6 (*CERS6*, previously known as *LASS6*) is one of the six members of the ceramide synthase gene family in humans. This enzyme generates C14 ceramide and C16 ceramide. *CerS6*-knockout mice demonstrate a significant decrease in C16-containing sphingolipids, accompanied by behavioral abnormalities, protection from high-fat-diet-induced obesity and glucose intolerance, as well as protection from colitis and neutrophil activation. Activation of *CerS6* and subsequent elevation of C16- ceramide were demonstrated in response to folate stress, serum starvation, deficiency of cytochrome c oxidase and methotrexate. *CerS6* is an integral membrane protein and has three domains located within the lumen of the endoplasmic reticulum, five transmembrane regions and three cytoplasmic regions. Tissue expression of human *CerS6* transcripts demonstrated a wide distribution profile, with highest levels observed in transformed lymphocytes and fibroblasts, skin, small intestine, cerebellar hemisphere and frontal cortex, heart ventricle, ovary, stomach and uterus (Holmes et al., 2018). Variants in *CERS6* in other populations have been associated with diabetic nephropathy, insulin resistance, lipid levels, Parkinson's disease, TNF levels and adenocarcinoma of the lung (National Center for Biotechnology Information (NCBI), 2019). The potential role of ceramides in diabetes and its complications are of interest (Galadari et al., 2013). Ceramide serves as a hub in sphingolipid metabolism. Evidence from a plethora of studies would suggest that ceramide could play a significant role in diabetes by several different mechanisms, including inducing

Table 4. Proportions of rs4668089-rs4668102-rs4668106-rs3845724 haplotypes in diabetic patients and unaffected individuals from this study compared with U.S. population controls

Haplotype	Nondiabetics	Diabetic patients	U.S. population controls
T C T T	0.50381	0.75679	0.33288
T T G C	0.20283	0.10384	0.33243
G T G C	0.15876	0.06475	0.13726
G C T T	0.05053	0.03329	0.05379
T C G C	0.02505	0.03040	0.06453
G C G C	0.03371	-	0.03220
T T T T	0.01427	0.01093	0.04691
G T T T	0.00654	-	-
T C G T	0.00451	-	-

pancreatic beta cell apoptosis, increasing insulin resistance and reducing insulin gene expression.

Two major hypotheses have been proposed to explain how increased adipose stores affect overall insulin resistance leading to T2D. First, when adipocytes exceed their storage capacity, fat begins to accumulate in other tissues leading to the production of metabolites that inhibit insulin signal transduction. Second, obesity triggers a chronic inflammatory state, and cytokines released from either adipocytes or macrophages infiltrate tissue and antagonize insulin action. The sphingolipid ceramide is a putative intermediate linking both excess nutrients (i.e. saturated fatty acids) and inflammatory cytokines (e.g. tumor necrosis factor alpha) to the induction of insulin resistance. Moreover, ceramide has been shown to be toxic in a variety of different cell types, including pancreatic beta cell, suggesting a possible pathway leading to the beta cell dysfunction seen in T2D. Ceramides have been shown to function as a second messenger in a variety of cellular events, including apoptosis and differentiation. Ceramides have the ability to regulate Akt/protein kinase B (PKB), which has been shown to promote the uptake of glucose, lipid and amino acids in various storage tissues. In both skeletal muscle and adipose tissue, Akt/PKB stimulates the translocation of an insulin-stimulated glucose uptake transporter (GLUT4) from the plasma to the cell membrane, thereby facilitating the increased rate of glucose into the tissue. Studies have reported that ceramide inhibits insulin-stimulated glucose uptake, GLUT4 translocation and glycogen synthesis; however, the precise mechanisms remain unclear. Increased *CerS6* expression has been associated with the glucose intolerance and hepatic insulin resistance associated with

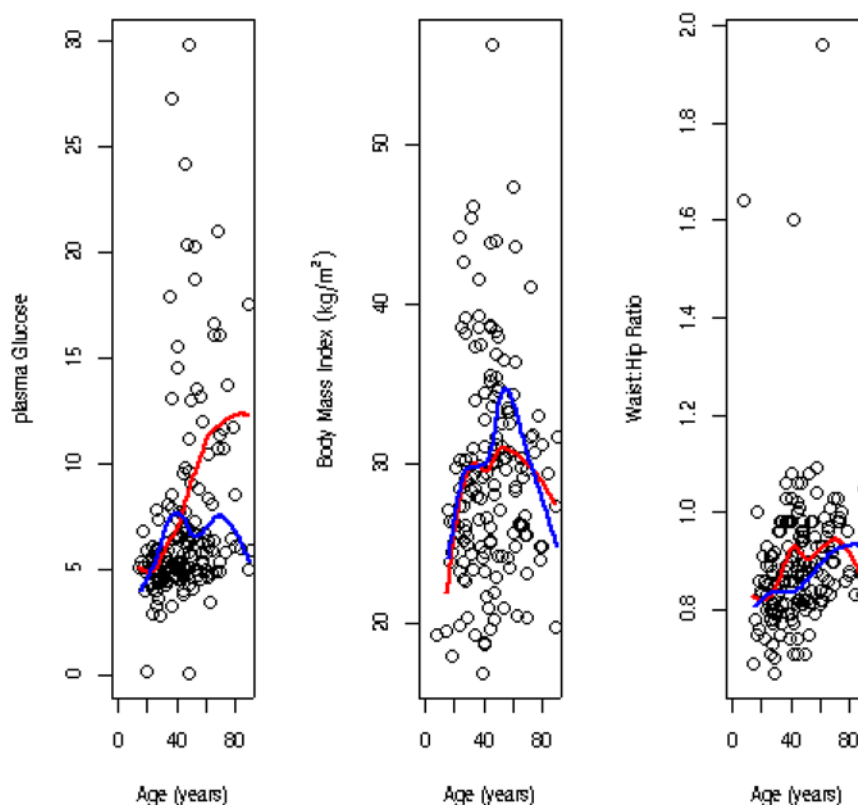


Fig. 2. (Colour online) Relationship between *rs4668089* genotype and blood glucose, body mass index and waist-to-hip ratio adjusting for age. The red curve is locally weighted regression curve for the T/T genotype, and the blue curve is for the G/T genotype.

hypercortisolism (NCBI, 2019). Short, as well as long, incubation with palmitate resulted in increased accumulation of ceramide in rat hepatocytes, which caused insulin resistance (Konstantynowicz-Nowicka et al., 2015). We conclude that further research is indicated to functionally analyze the potential role of *CerS6* in the pathogenesis and pathophysiology of T2D.

After highlighting *CERS6*, we also note that *rs4668106* and *rs3845724* genotypes are both suggestively related to expression of *G6PC2*. The glucose-6-phosphatase catalytic subunit 2 (also known as IGRP) is a pancreatic islet specific component of a complex that hydrolyses glucose-6-phosphate. In type 1 diabetes, the protein is an important autoantigen, but more critically in the current setting, variants in this gene are strongly associated with FPG levels, with the C allele of the known regulatory SNP *rs560887* increasing FPG by 0.075 mmol/l (Baerenwald et al., 2013). *G6PC2* lies 100 kbp telomeric to *CERS6*, and we are increasingly aware that regulatory SNPs can be at such distances from the target; for example, *FTO* SNPs and *IRX3* expression. The Islet Regulome Browser (Mularoni et al., 2017; Miguel-Escalada et al., 2018) Hi-C experiment shows 'high-confidence' evidence of a chromatin interaction between the region of *CERS6* close to our association peak with a strong enhancer region of *G6PC2* (Figure 4).

In conclusion, this study demonstrates a strong association between noncoding variants in the *CERS6* gene on chromosome 2q24.3 and T2D in a large Indigenous Australian pedigree and in a population of unrelated Indigenous Australian subjects. This finding highlights a potentially significant role of this gene in the development of T2D in this population.

Acknowledgements. We are grateful to the people who participated in the research program, which would not have been possible without the

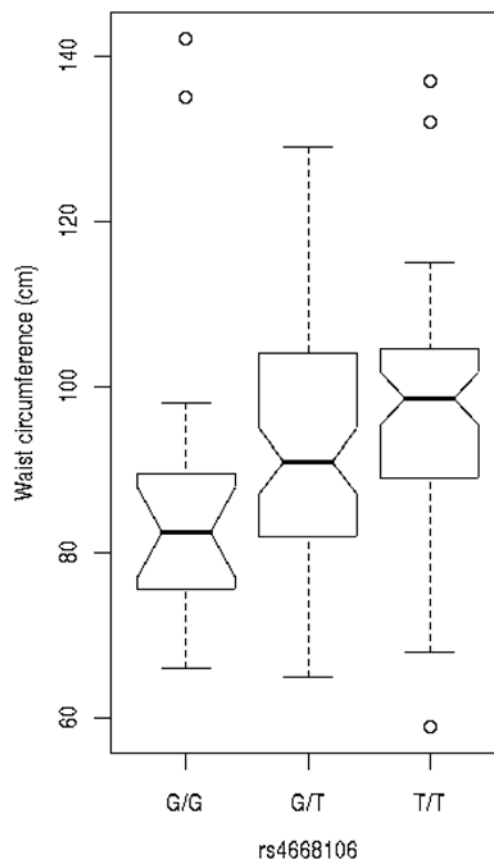


Fig. 3. SNP genotype association with waist circumference for members of the Indigenous Australian pedigree.

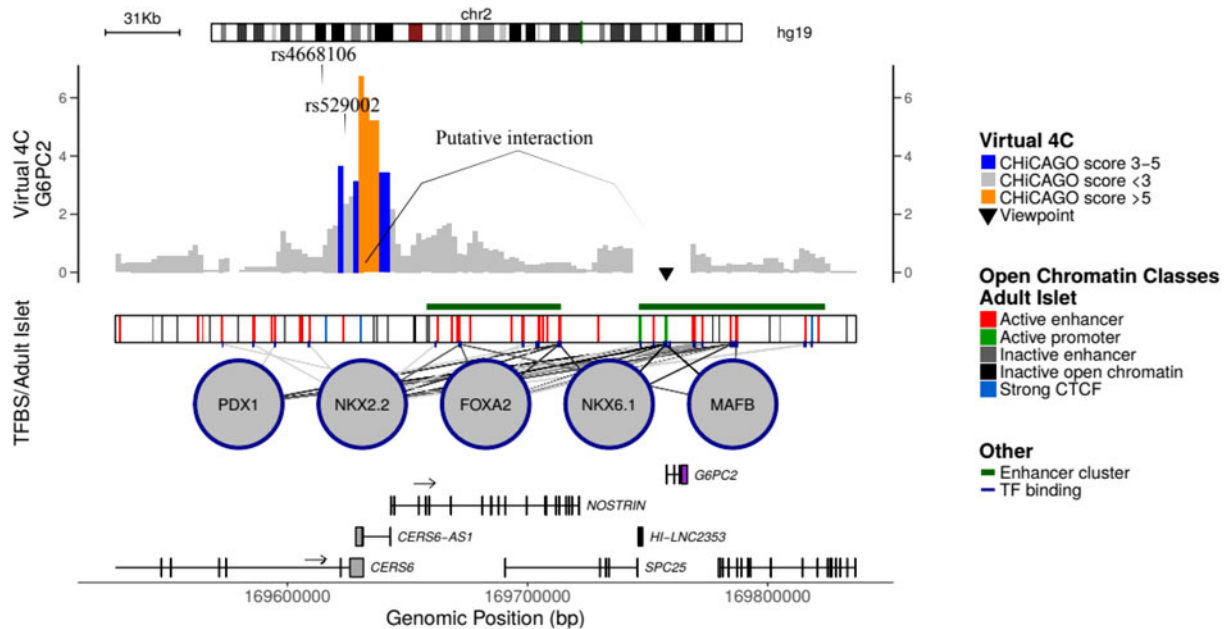


Fig. 4. The plot from Islet Regulome Browser (<http://www.isletregulome.org>) showing Virtual 4C estimates for interaction between 'bait' region on G6PC2 and distal CERS6/CERS6-AS1. The chromatin interaction map is constructed from promoter capture HiC data (four human islet samples) centered on the queried bait region, hence, virtual 4C. A CHiCAGO score >5 is a high-confidence interaction (Miguel-Escalada et al., 2018).

commitment and enthusiasm of the friends and staff of the Yulu-Burri-Ba health centre, North Stradbroke Island Queensland, Cherbourg Community health Centre and the Beaudesert community. This work was funded by an NHMRC Project Grant, by the Princess Alexandra Hospital Research and Development Foundation, by the Lions Medical Research Foundation, by the Novo Nordisk Regional Diabetes Support Scheme, by the Australian Government's Cooperative Research Centre for the Discovery of Genes for Common Human Diseases, by National Health and Medical Research Council of Australia (NHMRC) grants, the National Institute of Health (NIH) and the Prince Charles Hospital Foundation. Special thanks are extended to Mrs. Chrissa Aroney for her generous donation. In recognition of the support of the Sylvia and Charles Viertel Charitable Foundation. This paper is dedicated to the memory of Professor Robert Turner.

References

- Baerenwald, D. A., Bonnefond, A., Bouatia-Naji, N., Flemming, B. P., Umunakwe, O. C., Oeser, J. K., ... O'Brien, R. M. (2013). Multiple functional polymorphisms in the G6PC2 gene contribute to the association with higher fasting plasma glucose levels. *Diabetologia*, 56, 1306–1316.
- Bastian, P. (1979). Coronary heart disease in tribal Aborigines — The West Kimberley survey. *Australian and New Zealand Journal of Medicine*, 9, 284–292.
- Bell, G. I., Pilkis, S. J., Weber, I. T., & Polonsky, K. S. (1996). Glucokinase mutations, insulin secretion, and diabetes mellitus. *Annual Review of Physiology*, 58, 171–186.
- Bouatia-Naji, N., Rocheleau, G., Van Lommel, L., Lemaire, K., Schuit, F., Cavalcanti-Proenca, C., ... Froguel, P. (2008). A polymorphism within the G6PC2 gene is associated with fasting plasma glucose levels. *Science*, 320, 1085–1088.
- Browning, B. L., Zhou, Y., & Browning, S. R. (2018). A one-penny imputed genome from next generation reference panels. *American Journal of Human Genetics*, 103, 338–348.
- Busfield, F., Duffy, D. L., Kesting, J. B., Walker, S. M., Lovelock, P. K., Good, D., ... Shaw, J. T. E. (2002). A genome-wide search for type diabetes susceptibility in Indigenous Australians: Linkage to a novel region on human chromosome 2q. *American Journal of Human Genetics*, 70, 349–357.
- Chen, H., Wang, C., Conomos, M. P., Stilp, A. M., Li, Z., Sofer, T., ... Lin, X. (2016). Control for population structure and relatedness for binary traits in

genetic association studies using logistic mixed models. *American Journal of Human Genetics*, 98, 653–666.

- Conomos, M. P., Gogarten, S. M., Brown, L., Chen, H., Rice, K., Sofer, T., ... Yu, C. (2019). GENESIS: GENetic ESTimation and Inference in Structured samples (GENESIS): Statistical methods for analyzing genetic data from samples with population structure and/or relatedness. R package version 2.12.4. Retrieved from <https://github.com/UW-GAC/GENESIS>.
- DeFronzo, R. A. (1988). Lilly Lecture 1987. The triumvirate: Beta-cell, muscle, liver. A collusion responsible for NIDDM. *Diabetes*, 37, 667–687.
- Duffy, D. L. (1997). Sib-pair: A program for non-parametric linkage/association analysis. *American Journal of Human Genetics*, 61, 197.
- Dupuis, J., Langenberg, C., Prokopenko, I., Saxena, R., Soranzo, N., Jackson, A. U., Barroso, I. (2010). New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. *Nature Genetics*, 42(2), 105–116.
- Fumagalli, M., Moltke, I., Garup, N., Racimo, F., Bjerregaard, P., Jorgensen, M. E., ... Neilsen, R. (2015). Greenlandic Inuit show genetic signatures of diet and climate adaptation. *Science*, 349, 1343–1347.
- Galadari, S., Rahman, A., Pallichankandy, S., Galadari, A., & Thayyullathil, F. (2013). Role of ceramide in diabetes mellitus: Evidence and mechanisms. *Lipids in Health and Disease*, 12, 98.
- Garup, N., Moltke, I., Andersen, M. K., Dalby, M., Vitting-Seerup, K., Kern, T., ... Hansen, T. (2018). Loss-of-function variants in ADCY3 increases risk of obesity and type 2 diabetes. *Nature Genetics*, 50, 172–174.
- Gault, A., O'Dea, K., Rowley, K. G., McLeay, T., & Traianedes, K. (1996). Abnormal glucose tolerance and other coronary heart disease risk factors in an isolated Aboriginal community in central Australia. *Diabetes Care*, 19, 1269–1273.
- GTEx Consortium. (2013). The Genotype-Tissue Expression (GTEx) project. *Nature Genetics*, 45, 580–585.
- Guest, C. S., O'Dea, K., Hopper, J. L., & Larkins, R. G. (1993). Hyperinsulinaemia and obesity in Aborigines of south-eastern Australia, with comparisons from rural and urban European populations. *Diabetes Research and Clinical Practice*, 20, 155–164.
- Hayman, N. (1997). Cardiovascular disease amongst Indigenous Australians. *Aboriginal and Islander Health Worker Journal*, 21, 15–17.
- Holmes, R. S., Barron, K. A., & Krupenko, N. I. (2018). Ceramide synthase 6: Comparative analysis, phylogeny and evolution. *Biomolecules*, 8, 111.

- Horikawa, Y., Iwasaki, N., Hara, M., Furuta, H., Hinokio, Y., Cockburn, B. N., ... Bell, G. I. (1997). Mutation in hepatocyte nuclear factor-1 beta gene (TCF2) associated with MODY. *Nature Genetics*, *17*, 384–385.
- International HapMap Consortium (IHM). (2005). A haplotype map of the human genome. *Nature*, *437*, 1299–1320.
- Konstantynowicz-Nowicka, K., Harasim, E., Baranowski, M., & Chabowski, A. (2015). New evidence for role of ceramide in the development of hepatic insulin resistance. *PLoS One*, *10*, e0116858.
- Lange, K., Papp, J. C., Sinsheimer, J. S., Sripracha, R., Zhou, H., & Sobel, E. M. (2013). Mendel: The Swiss army knife of genetic analysis programs. *Bioinformatics*, *29*, 1568–1570.
- Maassen, J. A., & Kadowaki, T. (1996). Maternally inherited diabetes and deafness: A new diabetes subtype. *Diabetologia*, *39*, 375–382.
- Mahajan, A., Taliun, D., Thurner, M., Robertson, N. R., Torres, J. M., Rayner, N. W., ... McCarthy, M. I. (2018). Fine-mapping type 2 diabetes loci to single-variant resolution using high-density imputation and islet-specific epigenome maps. *Nature Genetics*, *50*, 1505–1513.
- Malecki, M. T., Jhala, U. S., Antonellis, A., Fields, L., Doria, A., Orban, T., ... Krolewski, A. S. (1999). Mutations in NEUROD1 are associated with the development of type 2 diabetes mellitus. *Nature Genetics*, *23*, 323–328.
- Manning, A. K., Hivert, M. F., Scott, R. A., Grimsby, J. L., Bouatia-Naji, N., Chen, H., ... Langenberg, C. (2012). A genome-wide approach accounting for body mass index identifies genetic variants influencing fasting glycemic traits and insulin resistance. *Nature Genetics*, *44*, 659–669.
- Miguel-Escalada, I., Bonàs-Guarch, S., Cebola, I., Ponsa-Cobas, J., Mendieta-Esteban, J., Rolando, D. M. Y., ... Ferrer, J. (2018). Human pancreatic islet 3D chromatin architecture provides insights into the genetics of type 2 diabetes. *bioRxiv* 400291.
- Miller, S. A., Dykes, D. D., & Polesky, H. F. (1988). A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Research*, *16*, 1215.
- Minster, R. L., Hawley, N. L., Su, C. T., Sun, G., Kershaw, E. E., Cheng, H., ... McGarvey, S. T. (2016). A thrifty variant in CREBRF strongly influences body mass index in Samoans. *Nature Genetics*, *48*, 1049–1054.
- Moltke, I., Grarup, N., Jorgensen, M. E., Bjerregaard, P., Treebak, J. T., Fumagalli, M., ... Hansen, T. (2014). A common Greenlandic TBC1D4 variant confers muscle insulin resistance and type 2 diabetes. *Nature*, *512*, 190–193.
- Moskvina, V., & Schmidt, K. M. (2008). On multiple-testing correction in genome-wide association studies. *Genetic Epidemiology*, *32*, 567–573.
- Mularoni, L., Ramos, M., & Pasquali, L. (2017). The pancreatic islet regulome browser. *Frontiers in Genetics*, *8*, 13.
- National Center for Biotechnology Information (NCBI). (2019). *Phenotype-genotype integrator*. Retrieved from <https://www.ncbi.nlm.nih.gov/gap/phenogeni?tab=1&gene=253782>
- O'Dea, K. (1991). Westernisation, insulin resistance and diabetes in Australian Aborigines. *Medical Journal of Australia*, *155*, 258–264.
- O'Dea, K. (1992). Diabetes in Australian Aborigines: Impact of the western diet and life style. *Journal of Internal Medicine*, *232*, 103–117.
- O'Dea, K., Lion, R. J., Lee, A., Traianedes, K., Hopper, J. L., & Rae, C. (1990). Diabetes, hyperinsulinemia, and hyperlipidemia in small Aboriginal community in northern Australia. *Diabetes Care*, *13*, 830–835.
- O'Dea, K., Patel, M., Kubisch, D., Hopper, J., & Traianedes, K. (1993). Obesity, diabetes, and hyperlipidemia in a central Australian Aboriginal community with a long history of acculturation. *Diabetes Care*, *16*, 1004–1010.
- Raftery, A., Hoeting, J., Volinsky, P., Painter, I., & Yeung, K. Y. (2006). *BMA: Bayesian model averaging R package version 3.03*. Retrieved from <http://www.r-project.org>, <http://www.research.att.com/~volinsky/bma.html>
- R Core Team. (2019). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <https://www.R-project.org/>.
- Steiner, D. F., Tager, H. S., Nanjo, K., Chan, S. J., & Rubenstein, A. H. (1995). Familial syndromes of hyperproinsulinaemia and hyperinsulinaemia with mild diabetes. In C. R. Scriver, A. L. Beaudet, W. S. Sly, & D. Valle (Eds.), *The metabolic and molecular bases of inherited disease* (pp. 897–904). New York, NY: McGraw-Hill.
- Stoffers, D. A., Zinkin, N. T., Stanojevic, V., Clarke, W. L., & Habener, J. F. (1997). Pancreatic agenesis attributable to a single nucleotide deletion in the human IPF1 gene coding sequence. *Nature Genetics*, *15*, 106–110.
- Suzuki, K., Akiyama, M., Ishigaki, K., Kanai, M., Hosoe, J., Shojima, N., ... Kadowaki, T. (2019). Identification of 28 new susceptibility loci for type 2 diabetes in the Japanese population. *Nature Genetics*, *51*, 379–386.
- Taylor, S. I. (1995). Diabetes mellitus. In C. R. Scriver, A. L. Beaudet, W. S. Sly, & D. Valle (Eds.), *The metabolic and molecular bases of inherited disease* (pp. 843–896). New York, NY: McGraw-Hill.
- Williams, D. R., Moffitt, P. S., Fisher, J. S., & Bashir, H. V. (1987). Diabetes and glucose tolerance in New South Wales coastal Aborigines: possible effects of non-Aboriginal genetic admixture. *Diabetologia*, *30*, 72–77.
- Yamagata, K., Furuta, H., Oda, N., Kaisaki, P. J., Menzel, S., Cox, N. J., ... Bell, G. I. (1996). Mutations in the hepatocyte nuclear factor-4alpha gene in maturity-onset diabetes of the young (MODY1). *Nature*, *384*, 458–460.
- Yamagata, K., Oda, N., Kaisaki, P. J., Menzel, S., Furuta, H., Vaxillaire, M., ... Bell, G. I. (1996). Mutations in the hepatocyte nuclear factor-1alpha gene in maturity-onset diabetes of the young (MODY3). *Nature*, *384*, 455–458.
- Zimmet, P., & Welborn, T. (2001). *The final report of the Australian Diabetes and Obesity Study (AusDiab)*. Canberra, Australia: Commonwealth Department of Health and Aged Care.