Identification of Novel Candidate Genes for Type 2 Diabetes From a Genome-Wide Association Scan in the Old Order Amish

Evidence for Replication From Diabetes-Related Quantitative Traits and From Independent Populations

Evadnie Rampersaud,[1] Coleen M. Damcott,[1] Mao Fu,[1] Haiqing Shen,[1] Patrick McArdle,[1] Xiaolian Shi,[1] John Shelton,[1] Jing Yin,[1] Yen-Pei C. Chang,[1] Sandra H. Ott,[1] Li Zhang,[1] Yiju Zhao,[1] Braxton D. Mitchell,[1] Jeffery O'Connell,[1] and Alan R. Shuldiner[1,2]

OBJECTIVE—We sought to identify type 2 diabetes susceptibility genes through a genome-wide association scan (GWAS) in the Amish.

RESEARCH DESIGN AND METHODS—DNA from 124 type 2 diabetic case subjects and 295 control subjects with normal glucose tolerance were genotyped on the Affymetrix 100K single nucleotide polymorphism (SNP) array. A total of 82,485 SNPs were tested for association with type 2 diabetes. Type 2 diabetes–associated SNPs were further prioritized by the following: *1*) associations with 5 oral glucose tolerance test (OGTT) traits in 427 nondiabetic Amish subjects, and *2*) in silico replication from three independent 100L SNP GWASs (Framingham Heart Study Caucasians, Pima Indians, and Mexican Americans) and a 500K GWAS in Scandinavians.

RESULTS— The strongest association ($P = 1.07 \times 10^{-5}$) was for rs2237457, which is located in growth factor receptor–bound protein 10 (*Grb10*), an adaptor protein that regulate insulin receptor signaling. rs2237457 was also strongly associated with OGTT glucose area under the curve in nondiabetic subjects ($P = 0.001$). Of the 1,093 SNPs associated with type 2 diabetes at $P < 0.01$, 67 SNPs demonstrated associations with at least one OGTT trait in nondiabetic individuals; 80 SNPs were nominally associated with type 2 diabetes in one of the three independent 100K GWASs, 3 SNPs (rs2540317 in *MFSD9*, rs10515353 on chromosome 5, and rs2242400 in *BCAT1* were associated with type 2 diabetes in more than one population), and 11 SNPs were nominally associated with type 2 diabetes in Scandinavians. One type 2 diabetes–associated SNP (rs3845971, located in *FHIT*) showed replication with OGTT traits and also in another population.

CONCLUSIONS—Our GWAS of type 2 diabetes identified several gene variants associated with type 2 diabetes, some of which are worthy of further study. ***Diabetes*** **56:3053–3062, 2007**

Type 2 diabetes, a complex disease that is characterized by insulin resistance and impaired β-cell function, represents a serious global public health problem, with more than 100 million people affected worldwide. While the primary molecular defects in type 2 diabetes remain largely unknown, it is clear that both genetic and environmental risk factors (including diet and physical inactivity) play critical roles. More than 20 genome-wide linkage scans of type 2 diabetes have been published, with evidence for linkage reported to a number of loci, including regions on chromosomes 1, 3, 8, 10, 12, 14, and 20 (1–8). Of the numerous candidate genes studied for their functional role in pancreatic β-cell function, insulin action, or energy metabolism, as well as positional candidate genes identified under linkage peaks, very few have variants that are consistently associated with type 2 diabetes. Indeed, common variants in only a few genes (*PPARγ*, *KCNJ11*, *CALPN10*, *TCF7L2*, and *HNF4A*) have been replicated in multiple populations (9).

The Old Order Amish are a closed founder population who emigrated from Switzerland in the early 1700s. They are a well-suited population for carrying out genetic studies since they live a relatively homogeneous lifestyle and maintain extensive family history records. The Amish Family Diabetes Study (AFDS) was initiated in 1995 with the goal of identifying the genetic determinants of type 2 diabetes (10). The sibling relative risk (λ_s) of type 2 diabetes in the Amish is 3.28 (95% CI 1.58–6.80), similar to that observed in other Caucasian populations. Genome-wide linkage analysis of type 2 diabetes and impaired glucose tolerance conducted in AFDS pedigrees (6) revealed regions on chromosomes 1q and 14q, both of which have been implicated in linkage scans from other populations (1–5,7). Specific variants in several well-replicated type 2 diabetes susceptibility genes are associated with type 2 diabetes in the Amish, including *TCF7L2* rs7903146

From the [1]Division of Endocrinology, Diabetes and Nutrition, University of Maryland, Baltimore, Maryland; and the [2]Geriatric Research and Education Clinical Center, Baltimore Veterans Administration Medical Center, Baltimore, Maryland.

Address correspondence and reprint requests to Alan R. Shuldiner, MD, Division of Endocrinology, Diabetes and Nutrition, University of Maryland School of Medicine, 660 West Redwood St., Room 494, Baltimore, MD 21201. E-mail: ashuldin@medicine.umaryland.edu.

Received for publication 3 April 2007 and accepted in revised form 5 September 2007.

Published ahead of print at http://diabetes.diabetesjournals.org on 10 September 2007. DOI: 10.2337/db07-0457.

Additional information for this article can be found in an online appendix at http://dx.doi.org/10.2337/db07-0457.

AFDS, Amish Family Diabetes Study; DGI, Broad-Lund-Novartis Diabetes Genome Initiative; GAUC, glucose area under the curve; GWAS, genome-wide association scan; HOMA-IR, homeostasis model assessment of insulin resistance; IAUC, insulin area under the curve; ISI, insulinogenic index; LD, linkage disequilibrium; NGT, normal glucose tolerant; OGTT, oral glucose tolerance test; SNP, single nucleotide polymorphism.

© 2007 by the American Diabetes Association.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

TABLE 1
Description of sample characteristics for type 2 diabetes GWAS in the Amish

Characteristics	Type 2 diabetic case and NGT control subject dataset	
	Type 2 diabetic case subjects	NGT control subjects
n	124	295
Male subjects (%)	33	52
Age (years)	51.3 ± 10.5	64.4 ± 12.9
BMI (kg/m^2)	29.3 ± 5.8	27.4 ± 4.7

Characteristics	OGTT-derived quantitative trait dataset		
	All	Men	Women
n	427	200	227
Age (years)	51.9 ± 11.9	52.2 ± 11.9	51.7 ± 11.9
BMI (kg/m^2)	27.7 ± 5.0	26.4 ± 4.0	28.9 ± 5.5
Fasting glucose (mmol/l)	5.09 ± 0.45	5.12 ± 0.47	5.07 ± 0.43
GAUC (mmol · l^{-1} · h^{-1})	379.2 ± 68.0	362.6 ± 65.3	393.9 ± 67.1
IAUC (mU· l^{-1} · h^{-1})	137.9 ± 88.9	108.5 ± 60.6	164.5 ± 101.3
HOMA-IR (mU per mmol/l^2)	2.6 ± 1.7	2.6 ± 2.2	2.6 ± 1.1
ISI (units/g)	0.9 ± 1.5	0.85 ± 1.9	0.93 ± 1.0

Data are means ± SD. All were nondiabetic subjects.

(odds ratio 1.60, $P = 0.008$) (11) and *HNF4A* rs2425640 (1.60, $P = 0.03$) (12). These findings suggest that the common type 2 diabetes gene variants in the Amish will likely be relevant to more outbred Caucasian populations.

Increased knowledge of common variation in the human genome learned as part of the HapMap initiative, coupled with advances in technologies, make possible the genotyping of thousands of single nucleotide polymorphisms (SNPs) in genome-wide association scans (GWAS). This is a powerful approach for identifying novel susceptibility genes for complex diseases (13,14). Recently, four GWAS studies of type 2 diabetes have identified variants at several novel loci, including *SLC30A8*, *IGF2BP2*, *CDKAL1*, *CDKN2A/CKDN2B*, and *HHEX/IDE*, that show strong replicated association with type 2 diabetes (15–18). In this article, we report results from a GWAS of type 2 diabetes in the Amish using the Affymetrix 100K SNP genotyping platform. We further characterize our findings using diabetes-related quantitative traits measured in nondiabetic Amish individuals. Lastly, we interpret the results of this scan in the context of three recently completed 100K GWAS studies for type 2 diabetes, as part of the Type 2 Diabetes 100K GWAS Consortium, along with a publicly available 500K GWAS of type 2 diabetes recently performed in a Scandinavian population.

RESEARCH DESIGN AND METHODS

Study population and phenotype assessment. Individuals with type 2 diabetes were identified from the AFDS. Details of the AFDS have been previously described (10). Phenotypic characterization of participants included medical and family history, anthropometry, and a 3-h 75-g oral glucose tolerance test (OGTT) with insulin levels. We based our primary analyses on 124 type 2 diabetic case and 295 normal glucose tolerant (NGT) control subjects. Type 2 diabetes was defined by fasting plasma glucose level (≥7 mmol/l), 2-h OGTT plasma glucose level (≥11.1 mmol/l), random plasma glucose level (≥11.1 mmol/l), the use of insulin or prescription oral glucose-lowering agents, or a diagnosis of diabetes documented by a physician. To minimize potentially misclassifying subjects with type 1 diabetes as having type 2 diabetes, case subjects with age of diagnosis <35 years were excluded. NGT control subjects were aged >38 years at the time of study and were selected based on fasting plasma glucose level (<6.1 mmol/l) and 2-h OGTT plasma glucose level (<7.8 mmol/l).

We performed secondary quantitative analyses of our mostly highly associated signals ($P < 0.01$) in a set of 427 nondiabetic Amish study participants, 132 of whom had impaired glucose tolerance and 295 of whom were part of the NGT control group used in our primary analysis. We estimated the mean levels of two OGTT-derived quantitative glucose traits (fasting glucose and glucose area under the curve [GAUC]) and three insulin traits (insulinogenic index [ISI], insulin area under the curve [IAUC], and homeostasis model assessment of insulin resistance [HOMA-IR]) according to the SNP genotypes in these individuals. Total GAUC and IAUC were calculated based on measurements at 0, 30, 60, 90, 120, 150, and 180 min using the trapezoidal method. The ISI was calculated as (insulin at 30 min − fasting insulin)/(glucose at 30 min − fasting glucose). HOMA-IR was calculated as fasting insulin (mU/l) × fasting glucose (mmol/l)/22.5. Table 1 describes the characteristics of this sample. The study protocol was approved by the institutional review board at the University of Maryland School of Medicine, and informed consent was obtained from each study participant.

Genotyping. Genomic DNA from leukocytes were genotyped using the Affymetrix GeneChip Mapping 100K array set, which consists of two microarray chips (*Xba*I and *Hind*III) (Affymetrix, Santa Clara, CA). Total genomic DNA (250 ng) was digested with *Xba*I or *Hind*III restriction enzymes and processed according to the Affymetrix protocol. The GeneChip Genotyping Analysis Software (GTYPE 4.0) was used to generate dynamic modeling algorithm–derived genotypes that were reanalyzed with the BRLMM (Bayesian RLMM) genotype calling algorithm (confidence threshold of 0.33) to improve the proportion of heterozygote calls (19). As an initial quality-control measure, BRLMM-generated chip files with call rates <90% for both enzymes across all SNPs were excluded. The resulting median call rate across all of the remaining 419 case-control samples was 97.5% (97.6% for *Xba*I and 97.4% for *Hind*III). We further removed individual SNPs with genotype call rates <90%, monomorphic SNPs and SNPs with minor allele frequency < 5%, and those deviating from Hardy-Weinberg equilibrium in control subjects ($P < 0.001$). The number of monomorphic and low-frequency SNPs ($n = 26{,}816$) in the Amish was not appreciably different from that observed in more outbred Caucasians of the HapMap CEU sample. For this report, we focused our analysis on the 82,485 autosomal SNPs that passed our quality-control standards.

The concordance rate for 11 quality-control samples that were run twice on the Affymetrix GeneChip mapping panel was 97.5%. We also calculated a cross-platform concordance rate of 98% for 419 samples in which 61 SNPs were genotyped using the Affymetrix GeneChip Mapping 100K panel and an independent Illumina 1536-plex GoldenGate assay. Supplementary Table 3 (available in an online appendix at http://dx.doi.org/10.2337/db07-0457) summarizes the quality checks and informativeness of the data.

Association testing and SNP prioritization scheme. Our GWAS analysis and SNP prioritization scheme is shown in Fig. 1. We selected the SNPs most highly associated with type 2 diabetes in our Amish case-control dataset based on *P* value rankings (*P* value cutoff <0.01) and then used two complementary approaches to further prioritize them. In one approach, we evaluated the most highly type 2 diabetes–associated SNPs for association with diabetes-related quantitative traits in an expanded set of 427 nondiabetic Amish subjects, 295 of whom were NGT control subjects from the primary type 2 diabetes association analysis (internal consistency). In a parallel approach, we as-

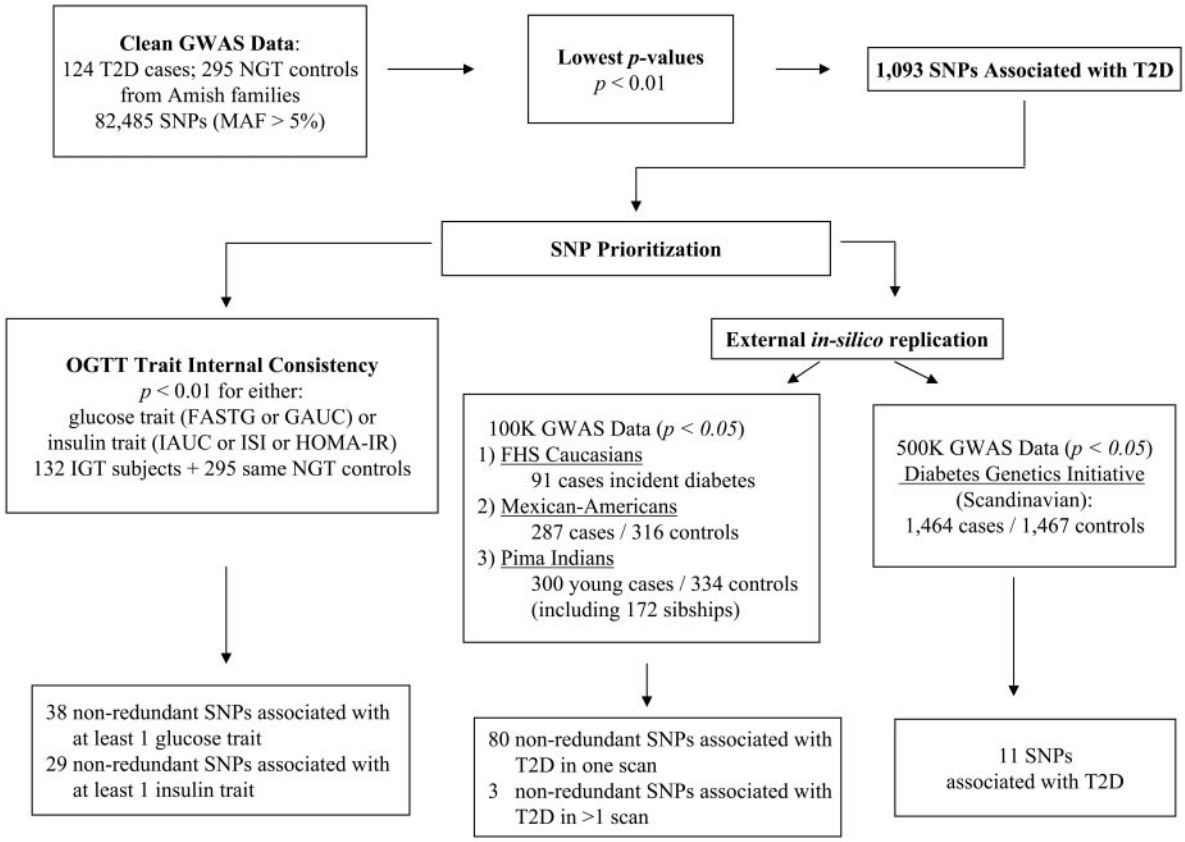

FIG. 1. Schematic diagram of analysis and SNP prioritization approach for a 100K type 2 diabetes GWAS in the Amish. FASG, fasting glucose during an OGTT; FHS, Framingham Heart Study.

sessed replication of the most highly associated type 2 diabetes–associated SNPs in the Amish in four independent GWASs from other populations (external replication).

Type 2 diabetes association analysis. We performed case-control association analysis using a variance component approach as implemented in SOLAR software (20). Using a liability threshold model, we modeled the probability that the individual was a case or control subject as a function of the individual's age, sex, and genotype, conditional on the correlations in phenotype among relative pairs. Statistical testing was performed using a likelihood ratio test, in which we compared the likelihood of the data under a model in which the genotype effect was estimated with the likelihood of a nested model in which the genotype effect was constrained to be zero. Odds ratios (ORs) were computed from variance components models. We chose to report the additive model as our primary analysis. Supplementary analyses using a dominant or recessive model did not yield any SNP showing genome-wide significance. Of 82,485 SNPs, 611 had $P < 0.01$ under a dominant model and 569 had $P < 0.01$ under the recessive model. Our complete dataset with results from all models is available online (available at http://www.medschool.umaryland.edu/amishstudies/index.asp). Pairwise linkage disequilibrium (LD) correlation statistics (r^2) were computed using the Helixtree software, version 5.0.2 (GoldenHelix, Bozeman, MT).

Quantitative trait analysis. For quantitative trait analyses performed in nondiabetic Amish subjects, we used the measured genotype approach, in which we estimated the likelihood of an additive genetic model given the pedigree structure (21). Before analysis, all insulin traits (IAUC, ISI, and HOMA-IR) were transformed by their natural logarithm to reduce skewness. Parameter estimates were obtained by maximum likelihood methods, and the significance of association was tested by the likelihood ratio test. Within each model, we simultaneously estimated the effects of age and sex. These analyses were performed using the SOLAR program (20).

Power calculations. Power calculations, based on the genetic power calculator of Purcell et al. (22), indicated that our sample would provide 80% power to detect a diabetes susceptibility allele having a genotype relative risk of 1.8 (for allele frequency of 30%, 124 case and 295 control subjects, 8% population prevalence of diabetes, assuming a multiplicative model) and 80% power to detect a quantitative trait loci accounting for 4% or higher of the trait variance for a continuously distributed phenotype (427 subjects).

In silico replication samples. We considered whether our best type 2 diabetes association signals (P value cutoff <0.01) replicated in at least one of three distinct populations (Framingham Caucasians, Mexican Americans, and Pima Indians), each with different study designs but performed using the same Affymetrix 100K genotyping platform. Descriptions of each of the Type 2 Diabetes 100K GWAS Consortium study populations are provided in accompanying articles (23–25) and in supplementary Table 4. We directly checked whether any of the 1,093 SNPs with the best type 2 diabetes association signals ($P < 0.01$) in the Amish were also associated with type 2 diabetes based on generalized estimating equations and family-based association tests in the Framingham Heart Study, Fisher's exact allelic association test in the Mexican-American study, and case-control and sib-based association tests in the Pima Indian study. We also utilized publicly available prereleased data (March 2007) from a type 2 diabetes GWAS carried out in a Scandinavian cohort of 1,464 type 2 diabetic case and 1,467 matched control subjects and genotyped using the Affymetrix 500K platform by the Broad-Lund-Novartis Diabetes Genetics Initiative (DGI) (available at http://www.broad.mit.edu/diabetes/) (18). We specifically checked replication of 295 of 1,093 of our most highly type 2 diabetes–associated SNPs that were present on both 100K and 500K Affymetrix genotyping arrays. Since LD structure may differ across populations, and to limit multiple comparisons, we defined replication only if the same SNP was associated with type 2 diabetes at $P < 0.05$ with an OR in the same direction (i.e., reflective of the same allelic risk).

RESULTS

Following quality-control and Hardy-Weinberg equilibrium checks, 82,485 informative SNPs were included in our analyses. The median physical inter-SNP distance was 11.3 kb, and the average distance between SNPs was 29 kb. Under the additive model, a total of 1,093 SNPs, some of which were in LD, were associated ($P < 0.01$) with type 2 diabetes (Fig. 2). The 50 most strongly type 2 diabetes–associated SNPs (i.e., lowest P values) are shown in Table 2. The complete dataset is available online (available at

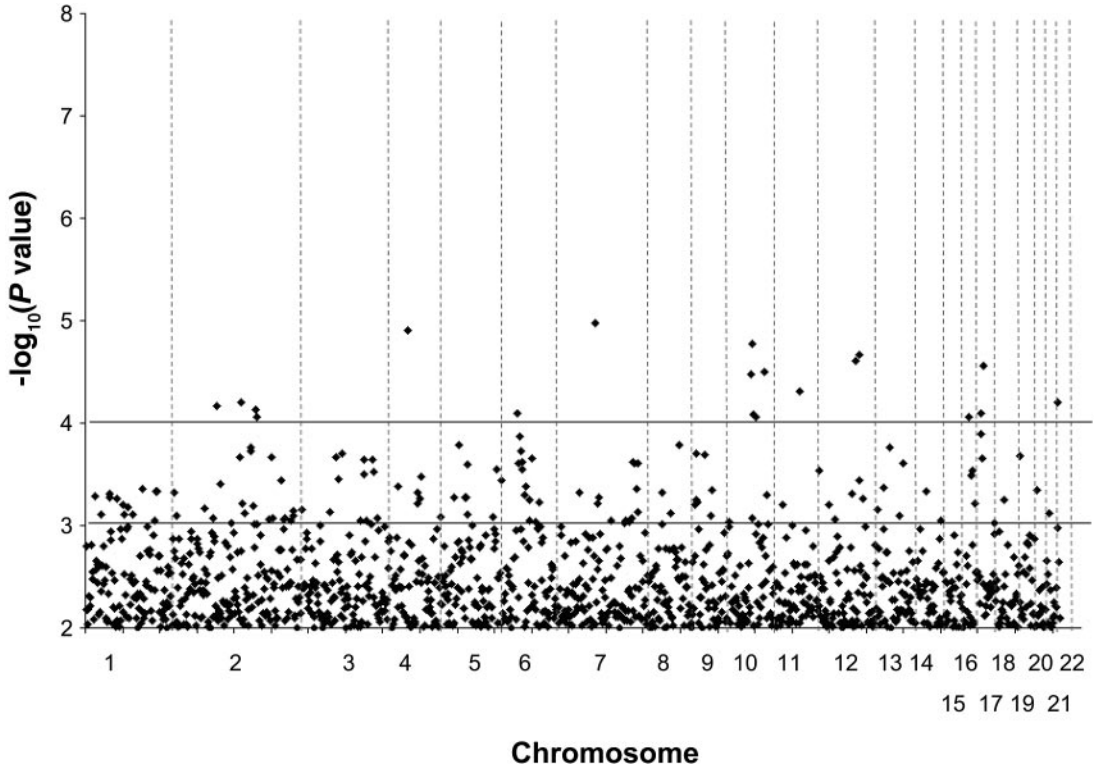

FIG. 2. SNP association *P* values (<0.01) across all autosomal chromosomes.

http://www.medschool.umaryland.edu/amishstudies/afds.asp). No SNP was associated with type 2 diabetes at a conservative Bonferroni-corrected level. The strongest association ($P = 1.07 \times 10^{-5}$) was for rs2237457 on chromosome 7, which is located in intron 4 of growth factor receptor–bound protein 10 (*Grb10*), an adaptor protein known to regulate signaling of insulin and IGF receptors (26–28). In addition to *Grb10*, 15 SNPs were associated with type 2 diabetes at $P < 1 \times 10^{-4}$ (Fig. 2 and Table 2). These SNPs are located in or near *MSH6* (chromosome 2), *PRKG2* (chromosome 4), *COL13A1* (chromosome 10), *MTHFSD* (chromosome 16), and *SPECC1* (chromosome 17), none of which are obvious candidate genes for type 2 diabetes. Adjustment for BMI did not have a large impact on the strength of the associations of these SNPs with type 2 diabetes (Table 2).

As a measure of internal consistency, we tested whether the 1,093 SNPS associated with type 2 diabetes ($P < 0.01$) were also associated with OGTT-derived quantitative traits in nondiabetic individuals. In these analyses, we considered two OGTT glucose traits (fasting glucose and GAUC) and three OGTT insulin traits (IAUC, HOMA-IR, and ISI), with $P < 0.01$ as our threshold for significance. Thirty-eight nonredundant ($r^2 < 0.80$) type 2 diabetes–associated SNPs were also associated with at least one glucose trait and showed the same allelic association as that for diabetes (i.e., the diabetes risk allele was also associated with higher glucose levels), while 29 nonredundant type 2 diabetes–associated SNPs were also associated with at least one insulin-related trait (Fig. 1; Table 3). Of the top 16 SNPs associated with type 2 diabetes at $P < 1 \times 10^{-4}$, rs2237457 in *Grb10* was the only one also associated with an OGTT trait ($P = 0.001$ for GAUC). Two perfectly correlated ($r^2 = 1$) type 2 diabetes–associated SNPs in *ADAMTS1* (chromosome 5) ($P = 0.004–0.005$) were associated with one glucose trait ($P = 0.006$ for GAUC) and one insulin trait ($P = 0.007$ for IAUC).

We next sought to determine which of our 1,093 most highly type 2 diabetes–associated SNPs were also associated with type 2 diabetes in any of three independent populations for which the same 100K Affymetrix platform was used or in the DGI Scandinavian population for which the 500K Affymetrix platform was used. We identified 80 nonredundant SNPs for which the same risk allele was also associated with type 2 diabetes in one of the three studies from the Type 2 Diabetes 100K GWAS Consortium ($P < 0.05$) and 11 nonredundant SNPs that showed consistent association in the DGI sample ($P < 0.05$) (Fig. 1; supplementary Table 3). In total, three SNPs demonstrated associations in the Amish as well as in two independent populations. The T-allele for rs2540317 in *MFSD9* on chromosome 2 was associated with decreased risk of type 2 diabetes in the Amish (OR 0.72, $P = 0.007$) and showed nominal association in the Pima Indian dataset (case-control OR 0.67, $P = 0.042$; sib-based OR 0.50, $P = 0.043$; and summary OR 0.63, $P = 0.016$) and also in the Mexican-American sample (case-control OR 0.75, $P = 0.047$). The G-allele in rs10515353 on chromosome 5 was associated with decreased risk of type 2 diabetes in the Amish (OR 0.61, $P = 0.005$) and also with decreased type 2 diabetes risk in Mexican-American (0.69, $P = 0.035$) and DGI (0.79, $P = 0.007$) samples. The T-allele in rs2242400 in *BCAT1* on chromosome 10 was associated with decreased risk of type 2 diabetes in the Amish (0.71, $P = 0.004$) and also in the Pima Indian dataset (sib-based OR 0.66, $P = 0.019$; summary OR 0.78, $P = 0.034$) and the Mexican-American dataset (OR 0.67, $P = 0.009$); borderline association was also seen in the DGI sample (0.86, $P = 0.051$). The direction of effect was the same for all studies.

TABLE 2
Fifty SNPs most highly associated with type 2 diabetes from Amish GWAS

SNP	Chromosome	Position*	Gene†	Alleles 1/2	Strand	Case subjects (*n*)	Control subjects (*n*)	Allele 2 case subjects	Allele 2 control subjects	Type 2 diabetes *P*‡	OR§	Type 2 diabetes *P* (BMI)‖
rs2237457	7	50693638	*Grb10*	A/G	−	124	293	0.53	0.68	1.07×10^{-5}	0.61	1.46×10^{-5}
rs980720	4	82272087	*PRKG2*	A/G	−	122	287	0.80	0.90	1.25×10^{-5}	0.52	4.87×10^{-5}
rs10509199	10	65295820		G/T	−	124	295	0.31	0.46	1.68×10^{-5}	0.62	4.46×10^{-5}
rs1373147	17	20147237	*SPECC1*	A/T	−	124	293	0.47	0.60	2.79×10^{-5}	0.63	5.26×10^{-6}
rs4082516	10	71374662	*COL13A1*	C/G	−	122	293	0.77	0.88	3.19×10^{-5}	0.53	1.27×10^{-5}
rs10509195	10	65193372		A/C	−	118	281	0.42	0.28	3.30×10^{-5}	1.62	7.1×10^{-5}
rs1395931	2	123535443		A/G	+	124	293	0.68	0.58	6.36×10^{-5}	1.57	2.90×10^{-5}
rs3136279	2	47871272	*MSH6*	G/T	−	124	294	0.20	0.11	6.81×10^{-5}	1.80	9.2×10^{-5}
rs1446732	2	134374210		G/T	−	121	295	0.51	0.37	7.38×10^{-5}	1.55	0.0001
rs10485249	6	70161908		C/G	+	124	295	0.82	0.92	7.96×10^{-5}	0.54	0.0004
rs2703813	17	20055907	*SPECC1*	A/G	−	123	285	0.51	0.64	8.12×10^{-5}	0.64	2.36×10^{-5}
rs1916412	10	65340801		C/T	−	124	293	0.62	0.49	8.35×10^{-5}	1.54	0.0002
rs1916411	10	65340839		C/G	−	123	292	0.38	0.51	$8.48 \times 10-5$	0.65	0.0002
rs3751797	16	85124993	*MTHFSD*	A/T	−	123	285	0.76	0.64	8.74×10^{-5}	1.60	0.0005
rs930621	2	134418548		C/T	+	122	290	0.54	0.41	8.74×10^{-5}	1.55	0.0001
rs10509201	10	65342248		A/G	−	123	293	0.62	0.49	8.80×10^{-5}	1.53	0.0002
rs2158473	17	20079682	*SPECC1*	C/T	+	119	292	0.51	0.39	0.0001	1.54	3.00×10^{-5}
rs2502497	6	75399083		G/T	+	116	279	0.05	0.10	0.0001	0.43	0.0002
rs430123	5	106109632		A/G	−	114	272	0.16	0.27	0.0002	0.61	0.0002
rs10504553	8	75038957	*TCEB1*	A/G	−	115	275	0.77	0.89	0.0002	0.56	0.0003
rs9287428	2	133912538	*FLJ34870*	C/T	+	124	287	0.83	0.71	0.0002	1.62	0.0002
rs10507601	13	55047379		A/G	+	123	294	0.04	0.12	0.0002	0.44	0.0006
rs994952	6	78319674		A/G	+	120	287	0.52	0.42	0.0002	1.52	0.0002
rs7604549	2	133923233	*FLJ34870*	A/G	−	123	294	0.26	0.40	0.0002	0.65	0.0002
rs2737245	8	116727757	*TRPS1*	A/C	−	113	276	0.58	0.71	0.0002	0.67	0.0003
rs1513287	3	114754898	*SIDT1*	A/G	+	120	289	0.58	0.44	0.0002	1.49	0.0003
rs7817780	8	119486147	*SAMD12*	C/T	−	120	294	0.91	0.80	0.0002	1.85	0.0002
rs297765	20	4435111		A/G	+	124	295	0.75	0.63	0.0002	1.55	0.0004
rs4410442	3	112434500		A/G	−	123	294	0.48	0.34	0.0002	1.52	0.0008
rs1351916	2	139479918		A/T	−	122	288	0.37	0.56	0.0002	0.67	0.0006
rs1507666	2	123532455		A/T	−	119	285	0.44	0.56	0.0002	0.66	0.0002
rs10498934	6	83516566		C/T	−	124	293	0.57	0.69	0.0002	0.66	0.0004
rs721729	3	174671967		A/G	−	119	285	0.43	0.56	0.0002	0.65	0.0002
rs1492908	3	159869201	*GFM1*	A/G	+	122	291	0.20	0.29	0.0002	0.62	0.0004
rs10484725	6	78325880		A/G	+	124	294	0.19	0.11	0.0002	1.75	0.0002
rs10487440	7	125763796		G/T	−	121	284	0.17	0.23	0.0002	0.59	0.0002
rs10498881	6	72999327	*RIMS1*	C/G	−	121	289	0.88	0.95	0.0002	0.48	0.0019
rs9317821	13	69008405		C/T	−	123	290	0.52	0.64	0.0002	0.67	0.0008
rs10487442	7	125850492		A/G	+	124	295	0.83	0.76	0.0003	1.68	0.0003
rs723397	5	107375942	*FBXL17*	C/T	+	123	295	0.24	0.36	0.0003	0.65	0.0014
rs6897150	5	165833376		A/G	−	123	294	0.26	0.14	0.0003	1.60	0.0003
rs1458405	6	78321198		G/T	−	112	270	0.51	0.40	0.0003	1.54	0.0002
rs10521205	17	12809867	*KIAA0672*	A/G	−	122	292	0.16	0.26	0.0003	0.60	0.0007
rs1023738	3	174685924		C/T	−	123	294	0.44	0.55	0.0003	0.67	0.0002
rs1367313	3	159848232	*GFM1*	C/T	−	124	295	0.79	0.71	0.0003	1.57	0.0005
rs10521204	17	12809385	*KIAA0672*	A/C	+	124	293	0.83	0.74	0.0003	1.65	0.0007
rs2034531	4	137364060		A/G	−	123	289	0.80	0.88	0.0003	0.59	0.0006
rs1282090	3	112779804	*CD96*	A/G	+	124	295	0.82	0.68	0.0004	1.58	0.0011
rs9328099	6	2382544		A/G	−	122	294	0.74	0.84	0.0004	0.63	0.0004
rs10497567	2	181111709	*KIAA1604*	A/G	+	124	295	0.93	0.85	0.0004	1.98	0.0005

*Genome build 36.1. †Genic region that contains associated SNPs. ‡*P* values derived using variance components analysis under an additive genetic model, adjusted for age, sex, and family structure. §OR calculated from a liability threshold model in SOLAR and estimated as allele 2 versus allele 1. ‖*P* values derived using variance components analysis under an additive genetic model, adjusted for age, sex, BMI, and family structure. Our complete dataset with results from all models are available online (available at http://www.medschool.umaryland.edu/amishstudies/index.asp).

Table 4 highlights our most consistent overall findings. We present 21 type 2 diabetes–associated SNPs in the Amish ($P < 0.005$) that also demonstrated either *1*) association with a diabetes-related quantitative trait ($P < 0.005$) in the Amish or *2*) in silico replication of type 2 diabetes association in one independent population ($P < 0.005$). Of interest, the T-allele in rs3845971 in *FHIT* was associated with increased risk of type 2 diabetes in the Amish (OR 1.42, $P = 0.004$) and also in Mexican Americans (1.46, $P = 0.004$) and with increased GAUC ($P = 4.0 \times 10^{-4}$) in nondiabetic Amish subjects.

DISCUSSION

In this article, we described the results of a GWAS of type 2 diabetes of 82,485 SNPs in the Old Order Amish, a genetically closed founder population with a homogeneous lifestyle. We reasoned that this population is likely

TABLE 3
SNPs associated with type 2 diabetes ($P < 0.01$) and at least one OGTT-derived trait ($P < 0.01$) in nondiabetic Amish subjects

SNP	Chromosome	Position*	Gene†	Strand	Alleles 1/2	Frequency allele 2	Type 2 diabetes		OGTT trait analysis				
							P‡	OR§	Trait	Mean 11‖	Mean 12‖	Mean 22‖	P‡
rs667222	1	55090696	*DHCR24*	−	A/G	0.75	0.005	0.71	GAUC	212.21	199.85	192.58	0.008
rs6588186	1	66319597	*PDE4B*	−	A/G	0.78	0.008	1.45	GAUC	192.09	185.94	201.07	0.002
rs570021	1	71217978	*PTGER3*	+	A/T	0.27	0.007	0.70	IAUC	550.64	613.59	690.13	0.007
rs6424414	1	71228058	*PTGER3*	+	C/T	0.27	0.007	0.70	IAUC	548.85	614.86	683.35	0.007
rs211706	1	75802445	*SLC44A5*	−	A/G	0.08	0.002	0.52	Fasting glucose	2.83	2.74	2.68	0.006
rs10493580	1	76277500	*LOC729766*	+	C/G	0.15	0.003	0.60	GAUC	199.25	186.85	189.20	0.005
rs1030414	1	81241030		−	C/T	0.07	0.005	0.51	GAUC	197.22	182.88	154.78	0.003
rs2257963	1	105672501		−	C/T	0.61	0.009	0.74	GAUC	208.52	195.36	190.49	0.009
rs1516150	1	105758114		+	A/G	0.39	0.010	1.36	GAUC	190.76	195.70	208.75	0.009
rs2576216	1	215202099	*ESRRG*	+	A/G	0.16	0.006	0.64	GAUC	200.81	187.83	190.21	0.002
rs2818781	1	215203396	*ESRRG*	+	C/T	0.84	0.003	1.61	GAUC	168.17	189.21	201.24	2.2×10^{-4}
rs2576212	1	215204637	*ESRRG*	−	A/G	0.84	0.005	1.58	GAUC	178.70	188.25	200.71	0.001
rs1343747	1	240794106		+	A/C	0.23	0.002	1.43	HOMA-IR	2.22	2.53	2.72	0.002
									IAUC	555.29	646.59	694.29	0.002
rs10495824	2	34604049		−	C/T	0.76	0.008	1.41	IAUC	677.25	635.10	545.83	0.002
rs10490049	2	40426854	*SLC8A1*	−	A/C	0.17	0.002	1.54	Fasting glucose	2.79	2.88	2.89	0.001
									GAUC	192.09	204.49	216.33	0.001
rs897097	2	55539162		+	A/G	0.61	0.001	1.46	HOMA-IR	2.09	2.37	2.49	0.010
rs2540317	2	102716239	*MFSD9*	−	C/T	0.78	0.007	0.72	GAUC	216.62	199.07	189.80	3.4×10^{-4}
rs4324336	2	105593460		−	A/T	0.78	0.005	1.44	Fasting glucose	2.78	2.76	2.85	0.002
rs2321201	2	134136406		−	C/T	0.37	0.006	1.35	IAUC	539.03	627.77	628.08	0.006
rs10510530	3	22459609		+	A/T	0.86	0.003	0.65	Fasting glucose	2.90	2.86	2.79	0.007
rs3845971	3	59975712	*FHIT*	+	C/T	0.72	0.004	1.42	GAUC	185.46	188.48	203.71	1.0×10^{-4}
rs1373340	3	66984977		−	A/C	0.86	0.007	0.67	HOMA-IR	1.89	2.14	2.43	0.006
									IAUC	529.61	519.62	610.68	0.009
rs2587015	3	147740690	*PLSCR1*	−	C/T	0.76	0.007	1.40	IAUC	751.72	609.81	547.07	0.001
rs2587014	3	147741153	*PLSCR1*	+	A/G	0.24	0.009	0.72	IAUC	548.12	610.88	743.82	0.001
rs2587012	3	147741360	*PLSCR1*	−	G/T	0.24	0.009	0.72	IAUC	549.20	618.96	758.40	4.0×10^{-4}
rs837678	3	191168575	*LEPREL1*	−	C/T	0.23	0.007	1.39	Fasting glucose	2.79	2.82	3.02	0.008
rs10517351	4	35352269		−	A/G	0.83	0.010	0.71	ISI	0.41	0.47	0.59	0.010
rs4128879	4	162098442		−	A/G	0.85	0.007	0.67	ISI	1.46	0.62	0.52	0.009
rs10521005	5	33638471	*ADAMTS1*	−	C/T	0.46	0.004	0.74	IAUC	659.01	568.57	550.30	0.007
rs9292501	5	33639167	*ADAMTS1*	+	A/G	0.45	0.005	0.75	GAUC	201.46	197.18	186.88	0.006
									IAUC	655.02	570.18	550.12	0.008
rs9292502	5	33639516	*ADAMTS1*	+	A/G	0.46	0.003	0.74	GAUC	201.78	197.24	186.94	0.006
									IAUC	661.14	571.39	551.03	0.006
rs709668	5	96199942		−	C/T	0.28	0.009	0.73	GAUC	201.47	191.03	183.01	0.002
rs950664	5	103381835		+	C/T	0.62	0.009	1.35	GAUC	188.26	190.43	205.91	1.4×10^{-4}
rs990133	5	108073268		+	C/T	0.86	0.004	1.58	GAUC	169.38	189.40	197.90	0.006
rs7720835	5	117269389		+	C/T	0.20	0.006	1.41	HOMA-IR	2.26	2.48	3.25	0.002
rs4365869	5	117282887		−	A/T	0.79	0.009	0.72	HOMA-IR	3.24	2.46	2.27	0.003
rs1423003	5	147427736	*SPINK5*	−	A/G	0.45	0.004	1.38	HOMA-IR	2.51	2.32	2.12	0.008
rs1862446	5	147460749	*SPINK5*	+	A/C	0.45	0.006	1.37	HOMA-IR	2.50	2.32	2.12	0.009
rs1422930	5	167330171	*ODZ2*	+	C/T	0.09	0.003	0.52	HOMA-IR	2.27	2.72	1.45	0.002
rs7770797	6	6566961	*LY86*	−	C/T	0.43	0.005	0.73	IAUC	536.16	579.90	649.95	0.010
rs10484908	6	108410287		+	C/G	0.11	0.009	1.53	GAUC	192.64	207.68	212.03	0.001
rs10237701	7	8860569		+	C/T	0.45	0.009	1.33	ISI	0.67	0.49	0.51	0.008
rs10487563	7	48905713		−	C/T	0.23	0.002	0.68	ISI	0.49	0.60	0.80	1.4×10^{-4}
rs2237457	7	50693638	*Grb10*	−	A/G	0.63	1.1×10^{-5}	0.61	GAUC	214.90	197.81	191.93	0.001
rs10499761	7	56161137	*LOC650200*	−	A/G	0.75	0.007	1.42	GAUC	178.50	192.32	201.72	0.001
rs3753107	7	91467087	*AKAP9*	−	C/T	0.77	0.006	0.72	HOMA-IR	2.88	2.50	2.27	0.007
rs10488510	7	91498161	*AKAP9*	+	G/T	0.23	0.006	1.39	HOMA-IR	2.29	2.54	2.76	0.010
rs647055	7	114159262			A/G	0.25	0.007	1.41	IAUC	551.83	620.21	780.97	3.8×10^{-4}
									ISI	0.49	0.65	0.68	0.001
rs10488284	7	119742580	*KCND2*	−	A/G	0.15	0.001	0.60	HOMA-IR	2.49	2.14	2.53	0.003
rs192392	8	53094411		+	A/C	0.54	0.004	0.74	IAUC	650.81	600.14	540.14	0.007
rs2450148	8	53112873		−	C/T	0.47	0.010	1.31	IAUC	537.54	580.77	669.94	0.002
rs10504133	8	53152586		+	C/T	0.51	0.002	0.72	IAUC	659.06	595.05	525.29	0.001
rs7001645	8	105192526	*RIMS2*	+	C/G	0.44	0.007	1.34	ISI	0.63	0.53	0.46	0.008
rs9297357	8	106211509		−	A/G	0.23	0.006	0.68	GAUC	201.81	185.25	199.13	0.001
rs10505229	8	115908561		+	C/T	0.18	0.002	1.50	GAUC	193.84	200.43	230.74	0.004

Continued on facing page

TABLE 3
Continued

SNP	Chromosome	Position*	Gene†	Strand	Alleles 1/2	Frequency allele 2	Type 2 diabetes		OGTT trait analysis				
							P‡	OR§	Trait	Mean 11‖	Mean 12‖	Mean 22‖	*P*‡
rs10511574	9	11941204		+	A/C	0.09	0.002	1.82	GAUC	193.05	209.84	170.35	0.007
rs10511777	9	26773463		−	A/G	0.94	0.006	0.53	Fasting glucose	3.19	2.92	2.80	0.002
rs10491665	9	29401568		+	A/C	0.09	0.006	0.58	GAUC	198.29	185.73	171.34	0.007
rs2804498	10	33660713	*NRP1*	−	A/G	0.40	0.001	1.42	Fasting glucose	2.78	2.82	2.89	0.003
rs768676	10	44022687		+	A/T	0.89	0.002	0.61	Fasting glucose	3.00	2.89	2.80	3.9×10^{-4}
rs1111803	10	72426094		+	C/T	0.55	0.008	0.75	GAUC	208.11	193.50	191.12	0.004
rs2437871	10	90268360	*C10orf5*	+	A/C	0.54	0.003	0.72	ISI	0.47	0.51	0.68	0.001
rs10509589	10	91804145		−	A/C	0.11	0.004	1.61	HOMA-IR	2.34	2.74	2.51	0.010
rs1887979	10	110133240		+	A/T	0.41	0.009	0.74	GAUC	201.35	193.00	187.75	0.010
rs10500651	11	5530156		+	G/T	0.73	0.006	1.42	HOMA-IR	2.88	2.46	2.27	0.005
rs7119814	11	107941856	*EXPH5*	+	C/T	0.14	0.007	1.50	ISI	0.58	0.46	0.29	0.008
rs10506173	12	39431860	*CNTN1*	+	A/G	0.95	0.007	0.55	Fasting glucose	2.97	2.91	2.81	0.007
rs312272	12	39534200	*CNTN1*	+	C/T	0.69	0.007	1.37	Fasting glucose	2.77	2.78	2.86	0.002
rs192852	12	39537359	*CNTN1*	−	C/T	0.32	0.009	0.74	Fasting glucose	2.86	2.77	2.77	0.001
rs2289522	12	39616798	*CNTN1*	+	C/T	0.55	0.001	1.46	Fasting glucose	2.76	2.81	2.86	0.004
rs3794247	12	39646028	*CNTN1*	+	A/C	0.57	0.002	1.42	Fasting glucose	2.76	2.81	2.86	0.003
rs10521210	17	12966928		+	C/T	0.65	0.009	1.34	GAUC	184.59	193.63	199.65	0.009
rs9915220	17	13630703		+	A/G	0.95	0.001	2.72	Fasting glucose	2.69	2.73	2.83	0.009
rs530205	18	42639148		+	C/T	0.61	0.009	0.76	IAUC	642.24	605.90	538.52	0.008
rs2953271	18	47696031		−	A/T	0.38	0.002	0.69	GAUC	200.87	195.67	183.77	0.004
rs10502971	18	49141959	*DCC*	+	A/G	0.59	0.009	1.34	GAUC	186.36	193.47	203.52	0.001
rs615696	18	49946910	*MBD2*	+	C/T	0.79	0.006	0.69	Fasting glucose	2.89	2.84	2.78	0.010
rs739453	19	40771776		+	A/G	0.60	0.005	0.73	IAUC	685.23	586.92	536.51	0.002
rs3745718	19	53406965	*CARD8*	−	A/G	0.40	0.010	0.75	GAUC	202.60	193.05	186.48	0.003
rs297765	20	4435111		+	A/G	0.67	2.1×10^{-4}	1.55	Fasting glucose	2.74	2.80	2.85	0.004
									GAUC	186.37	191.30	203.12	0.001
rs2255140	21	31975858	*SFRS15*	+	C/T	0.87	0.009	1.57	ISI	0.37	0.48	0.60	0.004

SNPs with $P < 0.01$ for type 2 diabetes associations were tested for consistency in a sample of nondiabetic individuals (295 of whom overlapped with the type 2 diabetes association dataset). Direction of association for glucose traits was required to be higher for diabetes risk allele. Neighboring SNPs in bold are in high LD ($r^2 > 0.80$). *Genic region that contains associated SNPs. ‡*P* values derived using the additive genetic model, adjusted for age, sex, and family structure. The complete dataset including results for dominant and recessive models are available online (available at http://www.medschool.umaryland.edu/amishstudies/index.asp). §OR calculated from a liability threshold model for allele 2 versus allele 1. ‖Mean values for traits are presented by genotype, with alleles shown in alphabetical order as "1/2." All insulin traits (IAUC, HOMA-IR, and ISI) were natural log transformed prior to analysis.

to carry a subset of the same common type 2 diabetes susceptibility variants as those found in the general population and that these variants might be easier to identify.

GWAS studies are prone to false-positives due to the very large number of statistical tests that must be performed. We were restricted by our relatively modest sample size and also computationally in our attempts to define a genome-wide significance level for which follow-up was justified (i.e., variance components tests were not feasible for the many replications needed for case-control permuted family datasets in the Amish). Thus, we relied heavily on a prioritization of SNPs worthy of follow-up by testing for *1*) internal consistency of type 2 diabetes–associated SNPs with OGTT-derived quantitative traits in nondiabetic Amish individuals, *2*) external replication of type 2 diabetes associations in three independent non-Amish 100K SNP GWAS studies, and *3*) external replication in a 500K SNP GWAS of type 2 diabetes in a large population of Scandinavians.

We found that no single SNP replicated consistently and in the same direction across all GWAS studies, nor were all SNPs associated with type 2 diabetes also associated with quantitative traits in nondiabetic individuals (supplementary Table 4). This is not particularly surprising since we expect that an appreciable number of type 2 diabetes–associated SNPs will be false-positives. Furthermore, a true susceptibility gene in one population might not be readily discernible in other populations due to inadequate sample sizes as well as differences in genetic background, LD, and environmental exposures. Similarly, a true susceptibility gene for type 2 diabetes might not show association with diabetes-related quantitative traits in nondiabetic individuals, especially since our OGTT-derived traits are only surrogates for gold-standard measures of insulin sensitivity and insulin secretion. Nevertheless, we were able to identify a number of candidate genes and loci that showed evidence for association with type 2 diabetes in

TABLE 4
SNPs associated with type 2 diabetes in the Amish ($P < 0.005$) and providing evidence for either internal consistency (association with an OGTT-derived quantitative trait in the Amish) or external replication (associated with type 2 diabetes in an independent population)

					Associated with type 2 diabetes in Amish ($P < 0.005$)				Internal consistency with OGTT-derived quantitative traits in Amish ($P < 0.005$)		External replication with type 2 diabetes in independent populations ($P < 0.005$)		
SNP	Chromosome	Position*	Gene†	Strand	Allele 1/2	Allele 2 case subjects	Allele 2 control subjects	*P*	OR‡	Trait	*P*	Population	*P*§	OR‡
rs689157	1	183074135	*C1orf24*	+	A/T	0.69	0.80	0.001	0.67			Pima Indians	0.002	0.65
rs2818781	1	215203396	*ESRRG*	+	C/T	0.89	0.81	0.003	1.61	GAUC	0.002			
rs1343747	1	240794106	—	+	A/C	0.30	0.21	0.002	1.43	HOMA-IR	0.002			
										IAUC	0.002			
rs10490049	2	40426854	*SLC8A1*	−	A/C	0.23	0.14	0.002	1.54	Fasting glucose	0.001			
										GAUC	0.001			
rs3845971	3	59975712	*FHIT*	+	C/T	0.79	0.68	0.004	1.42	GAUC	4.0×10^{-4}	Mexican Americans	0.004	1.46
rs9312113	4	63643792	—	+	C/G	0.75	0.64	0.004	1.41			FHS	0.013, 0.003	1.66
rs3775745	4	71147663	*CSN3*	−	A/C	0.26	0.39	0.002	0.70			Mexican Americans	0.003	0.68
rs1422930	5	167330171	*ODZ2*	+	C/T	0.04	0.11	0.003	0.52	HOMA-IR	0.002			
rs9321743	6	140005650	—	−	A/C	0.42	0.32	0.003	1.42			FHS	0.002, 0.024	1.60
rs10487563	7	48905713	—	−	C/T	0.15	0.27	0.002	0.68	ISI	0.001			
rs2237457	7	50693638	*Grb10*	−	A/G	0.53	0.68	1.1×10^{-5}	0.61	GAUC	0.001			
rs10488284	7	119742580	*KCND2*	−	A/G	0.12	0.17	0.001	0.60	HOMA-IR	0.003			
rs10504133	8	53152586	—	+	C/T	0.43	0.54	0.002	0.72	IAUC	0.001			
rs10505229	8	115908561	—	+	C/T	0.27	0.15	0.002	1.50	GAUC	0.004			
rs2804498	10	33660713	*NRP1*	−	A/G	0.49	0.36	0.001	1.42	Fasting glucose	0.003			
rs768676	10	44022687	—	+	A/T	0.85	0.91	0.002	0.61	Fasting glucose	0.004			
rs2437871	10	90268360	*C10orf59*	+	A/C	0.47	0.57	0.003	0.72	ISI	0.001			
rs2289522	12	39616798	*CNTN1*	+	C/T	0.64	0.52	0.001	1.46	Fasting glucose	0.004			
rs7986010	13	91517260	*GPC5*	+	A/G	0.44	0.34	0.002	1.41			FHS	0.004, 0.042	1.59
rs2953271	18	47696031	—	−	A/T	0.27	0.43	0.002	0.69	GAUC	0.004			
rs297765	20	4435111	—	+	A/G	0.75	0.63	2.1×10^{-4}	1.55	Fasting glucose	0.004			

*Genome build 36.1. †Genic region that contains associated SNPs. ‡OR calculated from a liability threshold model in SOLAR and estimated as allele 2 versus allele 1. §Case-control general estimating equation and family-based association test *P* values given for Framingham Heart Study (FHS) data; summary *P* value given for Pima Indians. All insulin traits (HOMA-IR, IAUC, and ISI) were natural log transformed prior to analysis.

more than one population and/or were also associated with OGTT-derived quantitative traits. These results are intriguing but must be interpreted with caution. None of these loci fall within previously identified linkage regions for type 2 diabetes (chromosomes 1 and 14) in the Amish.

Our strongest type 2 diabetes association signal in the Amish was observed on chromosome 7 in a functionally relevant type 2 diabetes candidate gene, *Grb10*. *Grb10* encodes growth factor–binding protein 10 and has been shown to bind to activated insulin receptor and act as a negative regulator of insulin action and glucose uptake (26–28). Overexpression of Grb10 in mice causes postnatal growth retardation and insulin resistance (29). Our 100K GWAS contained a total of 12 SNPs in *Grb10*, 6 of which were associated with type 2 diabetes ($P < 0.05$) and were in partial LD with each other ($r^2 = 0.16–0.78$). Rs2237457, located in intron 4, provided the lowest *P* value for association (OR 0.61 for the G- vs. A-allele, $P = 1.07 \times 10^{-5}$). This SNP was also strongly associated with OGTT GAUC in nondiabetic Amish individuals ($P = 0.001$). Rs2237457 was not associated with type 2 diabetes in the other three populations in which this SNP was genotyped or in the 500K SNP Scandinavian type 2 diabetes GWAS; however, three SNPs (rs2190496, rs2237478, and rs7805310) in *Grb10* that were genotyped in the Scandinavian cohort were associated with type 2 diabetes ($P = 0.029$, $P = 0.01$, and $P = 0.004$, respectively) and are in partial LD with rs2237457 ($r^2 = 0.12–0.49$ in HapMap CEU). Lack of replication could suggest a false-positive or that variation in *Grb10* is a true positive specific to the Amish due to a founder effect or context-dependent phenotypic expression of the variant due to genetic background or environmental influences. Alternatively, this variant could be in LD with a functional variant, and extended LD in the Amish enabled a type 2 diabetes association to be detected in this population and not the others.

In a recent report by Di Paola et al. (30), the A-allele of rs4947710, a synonymous coding SNP in *Grb10*, was associated with decreased risk of type 2 diabetes in a relatively homogeneous population of Italian Caucasians ($P = 0.0001$). This SNP was not part of the 100K SNP panel nor was our most highly type 2 diabetes–associated SNP (rs2237457) genotyped in the Italian sample. We found that rs2237457 and rs4947710 are not in LD ($r^2 = 0$) in HapMap CEU samples. However, rs10486757, another *Grb10* SNP associated with type 2 diabetes in the Amish ($P = 0.024$), is in LD with rs4947710 ($r^2 = 0.64$ in HapMap CEU). Further investigation of *Grb10* is currently underway.

Our GWAS and replication strategy have several limitations. First, the relatively small sample size limits our ability to detect gene variants of modest effect size. Second, we recognize that the definition of external replication of our top SNPs across three independent 100K studies of type 2 diabetes might represent a skewed distribution of the overall results since replication was limited to our ~1,000 most highly type 2 diabetes–associated SNPs. This approach was used to facilitate comparisons across populations and also to limit the number of false-positive replications due to multiple comparisons. To the extent that we attempted to pursue signals that represent the "lowest hanging fruit," we believe that the approach we have taken is reasonable. A formal meta-analysis of the entire set of data from all four 100K studies is currently underway. Third, our replication approach was focused at the level of the SNP in order to avoid additional multiple comparisons. However, it is possible that we did not identify significantly associated SNPs in other populations that were in LD with our top SNPs. This is particularly relevant for our comparisons with the Scandinavian 500K GWAS, for which only 27% of the SNPs identified in the Amish with $P < 0.01$ were identified in the 500K SHP panel.

The likelihood that we missed common variants important to type 2 diabetes is high due to the relatively sparse density of the 100K SNP panel (mean intermarker distance = 29 kb) compared with other denser GWAS SNP panels. For example, SNPs in well-replicated genes (*SLC30A8*, *IGF2BP2*, *CKDAL1*, *CDKN2A/CKDN2B*, and *HHEX/IDE*) found in four recently published type 2 diabetes GWAS studies (15–18), as well as previously known type 2 diabetes–associated variants in *TCF7L2*, *KCNJ11*, *HNF4A*, or *CAPN10* (9), were not adequately covered on the 100K genotyping panel (i.e., $r^2 < 0.8$ between the SNP of interest and SNPs on the 100K panel). As a positive control, we previously demonstrated that *TCF7L2* SNP rs7903146 and the *HNF4A* promoter SNP rs2425640, neither of which is present on the 100K panel, were associated with type 2 diabetes and impaired glucose tolerance in the Amish Family Diabetes Study (OR 1.57, $P = 0.008$; 1.60, $P = 0.04$, respectively) (12,25). Interestingly, rs10509645 in *HHEX* on the 100K panel ($r^2 = 0.7$ with rs7923837 found previously to be strongly associated with type 2 diabetes in other GWAS studies) was significantly associated with type 2 diabetes in the Amish (OR 1.30 for the G-allele; $P = 0.02$). Rs9300039 on chromosome 11, shown to be associated with type 2 diabetes in the other GWAS studies (17), was present on the 100K panel but was not significantly associated with type 2 diabetes in the Amish (OR 1.09 for the C-allele; $P = 0.67$).

In summary, we presented results from our initial examination of a GWAS of type 2 diabetes in the Amish. Although we did not identify any genes associated with type 2 diabetes that reached genome-wide significance, we report a number of genes and loci that are worthy of further study based on replication in other studies or on quantitative trait loci consistency. This report (and the three companion articles) provides a valuable resource for other investigators to utilize in the search for the pathogenic variants for type 2 diabetes.

ACKNOWLEDGMENTS

This work was supported by the following National Institutes of Health Research Grants: Training Grant in Cardiac and Vascular Cell Biology (T32 HL072751), R01 DK54261, the University of Maryland General Clinical Research Center (M01 RR16500), Hopkins Bayview General Clinical Research Center (M01 RR02719), the Maryland Clinical Nutrition Research Unit (P30 DK072488), and the Baltimore Veterans Administration Geriatric Research and Education Clinical Center.

We extend our thanks to our collaborators in the Type 2 Diabetes 100K GWAS Consortium for sharing prepublication results from the Starr County Health Studies, Framingham Heart Study, and the Pima Indians as well as to the Broad-Lund-Novartis Diabetes Genomic Initiative for access to results from their GWAS.

We also thank Soren Snitker for helpful comments and Adam Naj for help with formatting figures for this manuscript. Lastly, we gratefully acknowledge our Amish liai-

sons and research staff and the extraordinary cooperation and support of the Amish community, without whom these studies would not be possible.

REFERENCES

1. Elbein SC, Hoffman MD, Teng K, Leppert MF, Hasstedt SJ: A genome-wide search for type 2 diabetes susceptibility genes in Utah Caucasians. *Diabetes* 48:1175–1182, 1999
2. Wiltshire S, Hattersley AT, Hitman GA, Walker M, Levy JC, Sampson M, O'Rahilly S, Frayling TM, Bell JI, Lathrop GM, Bennett A, Dhillon R, Fletcher C, Groves CJ, Jones E, Prestwich P, Simecek N, Rao PV, Wishart M, Foxon R, Howell S, Smedley D, Cardon LR, Menzel S, McCarthy MI: A genomewide scan for loci predisposing to type 2 diabetes in a U.K. population (the Diabetes UK Warren 2 Repository): analysis of 573 pedigrees provides independent replication of a susceptibility locus on chromosome 1q. *Am J Hum Genet* 69:553–569, 2001
3. Du W, Sun H, Wang H, Qiang B, Shen Y, Yao Z, Gu J, Xiong M, Huang W, Chen Z, Zuo J, Hua X, Gao W, Sun Q, Fang F: Confirmation of susceptibility gene loci on chromosome 1 in northern China Han families with type 2 diabetes. *Chin Med J (Engl)* 114:876–878, 2001
4. Meigs JB, Panhuysen CI, Myers RH, Wilson PW, Cupples LA: A genome-wide scan for loci linked to plasma levels of glucose and HbA_{1c} in a community-based sample of Caucasian pedigrees: the Framingham Offspring Study. *Diabetes* 51:833–840, 2002
5. Hanson RL, Ehm MG, Pettitt DJ, Prochazka M, Thompson DB, Timberlake D, Foroud T, Kobes S, Baier L, Burns DK, Almasy L, Blangero J, Garvey WT, Bennett PH, Knowler WC: An autosomal genomic scan for loci linked to type II diabetes mellitus and body-mass index in Pima Indians. *Am J Hum Genet* 63:1130–1138, 1998
6. Hsueh WC, St. Jean PL, Mitchell BD, Pollin TI, Knowler WC, Ehm MG, Bell CJ, Sakul H, Wagner MJ, Burns DK, Shuldiner AR: Genome-wide and fine-mapping linkage studies of type 2 diabetes and glucose traits in the Old Order Amish: evidence for a new diabetes locus on chromosome 14q11 and confirmation of a locus on chromosome 1q21–q24. *Diabetes* 52:550–557, 2003
7. Ng MC, So WY, Cox NJ, Lam VK, Cockram CS, Critchley JA, Bell GI, Chan JC: Genome-wide scan for type 2 diabetes loci in Hong Kong Chinese and confirmation of a susceptibility locus on chromosome 1q21–q25. *Diabetes* 53:1609–1613, 2004
8. McCarthy MI: Growing evidence for diabetes susceptibility genes from genome scan data. *Curr Diab Rep* 3:159–167, 2003
9. Owen KR, McCarthy MI: Genetics of type 2 diabetes. *Curr Opin Genet Dev* 17:239–244, 2007
10. Hsueh W-C, Mitchell BD, Aburomia R, Pollin T, Sakul H, Ehm MG, Michelsen BK, Wagner MJ, St. Jean PL, Knowler WC, Burns DK, Bell CJ, Shuldiner AR: Diabetes in the old order Amish: characterization and heritability analysis of the Amish Family Diabetes Study. *Diabetes Care* 23:595–601, 2000
11. Damcott CM, Pollin TI, Reinhart LJ, Ott SH, Shen H, Silver KD, Mitchell BD, Shuldiner AR: Polymorphisms in the transcription factor 7-like 2 (TCF7L2) gene are associated with type 2 diabetes in the Amish: replication and evidence for a role in both insulin secretion and insulin resistance. *Diabetes* 55:2654–2659, 2006
12. Damcott CM, Hoppman N, Ott SH, Reinhart LJ, Wang J, Pollin TI, O'Connell JR, Mitchell BD, Shuldiner AR: Polymorphisms in both promoters of hepatocyte nuclear factor 4-α are associated with type 2 diabetes in the Amish. *Diabetes* 53:3337–3341, 2004
13. Klein RJ, Zeiss C, Chew EY, Tsai JY, Sackler RS, Haynes C, Henning AK, San Giovanni JP, Mane SM, Mayne ST, Bracken MB, Ferris FL, Ott J, Barnstable C, Hoh J: Complement factor H polymorphism in age-related macular degeneration. *Science* 308:385–389, 2005
14. Gibbs JR, Singleton A: Application of genome-wide single nucleotide polymorphism typing: simple association and beyond. *PLoS Genet* 2:e150, 2006
15. Zeggini E, Weedon MN, Lindgren CM, Frayling TM, Elliott KS, Lango H, Timpson NJ, Perry JR, Rayner NW, Freathy RM, Barrett JC, Shields B, Morris AP, Ellard S, Groves CJ, Harries LW, Marchini JL, Owen KR, Knight B, Cardon LR, Walker M, Hitman GA, Morris AD, Doney AS, the Wellcome Trust Case Control Consortium (WTCCC), McCarthy MI, Hattersley AT: Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. *Science* 316:1336–1341, 2007
16. 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* 39:770–775, 2007
17. 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* 316:1341–1345, 2007
18. Saxena R, Voight BF, Lyssenko V, Burtt NP, de Bakker PI, Chen H, Roix JJ, Kathiresan S, Hirschhorn JN, Daly MJ, Hughes TE, Groop L, Altshuler D, Almgren P, Florez JC, Meyer J, Ardlie K, Bengtsson BK, Isomaa B, Lettre G, Lindblad U, Lyon HN, Melander O, Newton-Cheh C, Nilsson P, Orho-Melander M, Rastam L, Speliotes EK, Taskinen MR, Tuomi T, Guiducci C, Berglund A, Carlson J, Gianniny L, Hackett R, Hall L, Holmkvist J, Laurila E, Sjogren M, Sterner M, Surti A, Svensson M, Svensson M, Tewhey R, Blumenstiel B, Parkin M, Defelice M, Barry R, Brodeur W, Camarata J, Chia N, Fava M, Gibbons J, Handsaker B, Healy C, Nguyen K, Gates C, Sougnez C, Gage D, Nizzari M, Gabriel SB, Chirn GW, Ma Q, Parikh H, Richardson D, Ricke D, Purcell S: Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. *Science* 316:1331–1336, 2007
19. BRLMM: an improved genotype calling method for the GeneChip Human Mapping 500K Array Set [article online], 2006. Available from http://www.affymetrix.com/support/technical/whitepapers/brlmm_whitepaper.pdf. Accessed 5 September 2007
20. Almasy L, Blangero J: Multipoint quantitative-trait linkage analysis in general pedigrees. *Am J Hum Genet* 62:1198–1211, 1998
21. Boerwinkle E, Chakraborty R, Sing CF: The use of measured genotype information in the analysis of quantitative phenotypes in man. I. Models and analytical methods. *Ann Intern Med* 50:181–194, 1986
22. Purcell S, Cherny SS, Sham PC: Genetic Power Calculator: design of linkage and association genetic mapping studies of complex traits. *Bioinformatics* 19:149–150, 2003
23. Florez JC, Manning AK, Dupuis J, McAteer J, Irenze K, Gianniny L, Mirel DB, Fox CS, Cupples LA, Meigs JB: A 100K genome-wide association scan for diabetes and related traits in the Framingham Heart Study: replication and integration with other genome-wide datasets. *Diabetes* 56:3063–3074, 2007
24. Hanson RL, Bogardus C, Duggan D, Kobes S, Knowlton M, Infante AM, Marovich L, Benitez D, Baier LJ, Knowler WC: A search for variants associated with young-onset type 2 diabetes in American Indians in a 100K genotyping array. *Diabetes* 56:3045–3052, 2007
25. Hayes MG, Pluzhnikov A, Miyake K, Sun Y, Ng MCY, Roe CA, Below JE, Nicolae RI, Konkashbaev A, Bell GI, Cox NJ, Hanis CL: Identification of type 2 diabetes genes in Mexican Americans through genome-wide association studies. *Diabetes* 56:3033–3042, 2007
26. Deng Y, Bhattacharya S, Swamy OR, Tandon R, Wang Y, Janda R, Riedel H: Growth factor receptor-binding protein 10 (Grb10) as a partner of phosphatidylinositol 3-kinase in metabolic insulin action. *J Biol Chem* 278: 39311–39322, 2003
27. Langlais P, Dong LQ, Ramos FJ, Hu D, Li Y, Quon MJ, Liu F: Negative regulation of insulin-stimulated mitogen-activated protein kinase signaling by Grb10. *Mol Endocrinol* 18:350–358, 2004
28. Mounier C, Lavoie L, Dumas V, Mohammad-Ali K, Wu J, Nantel A, Bergeron JJ, Thomas DY, Posner BI: Specific inhibition by hGrb10zeta of insulin-induced glycogen synthase activation: evidence for a novel signaling pathway. *Mol Cell Endocrinol* 173:15–27, 2001
29. Shiura H, Miyoshi N, Konishi A, Wakisaka-Saito N, Suzuki R, Muguruma K, Kohda T, Wakana S, Yokoyama M, Ishino F, Kaneko-Ishino T: Meg1/Grb10 overexpression causes postnatal growth retardation and insulin resistance via negative modulation of the IGF1R and IR cascades. *Biochem Biophys Res Commun* 329:909–916, 2005
30. Di Paola R, Ciociola E, Boonyasrisawat W, Nolan D, Duffy J, Miscio G, Cisternino C, Fini G, Tassi V, Doria A, Trischitta V: Association of hGrb10 genetic variations with type 2 diabetes in Caucasian subjects. *Diabetes Care* 29:1181–1183, 2006