© 2000 Nature America Inc. • http://genetics.nature.com

The common PPARγ Pro12Ala polymorphism is associated with decreased risk of type 2 diabetes

David Altshuler[1,2,3*], Joel N. Hirschhorn[1,3,4*], Mia Klannemark[5], Cecilia M. Lindgren[1,5], Marie-Claude Vohl[6], James Nemesh[1], Charles R. Lane[1], Stephen F. Schaffner[1], Stacey Bolk[1], Carl Brewer[6], Tiinamaija Tuomi[5,7], Daniel Gaudet[8], Thomas J. Hudson[1,6], Mark Daly[1], Leif Groop[5] & Eric S. Lander[1,9]

These authors contributed equally to this report.

Genetic association studies are viewed as problematic and plagued by irreproducibility[1]. Many associations have been reported for type 2 diabetes[2–17], but none have been confirmed in multiple samples and with comprehensive controls. We evaluated 16 published genetic associations to type 2 diabetes and related sub-phenotypes using a family-based design to control for population stratification, and replication samples to increase power. We were able to confirm only one association, that of the common Pro12Ala polymorphism in peroxisome proliferator-activated receptor-γ (PPARγ) with type 2 diabetes. By analysing over 3,000 individuals, we found a modest (1.25-fold) but significant (*P*=0.002) increase in diabetes risk associated with the more common proline allele (~85% frequency). Moreover, our results resolve a controversy about common variation in PPARγ. An initial study found a threefold effect[12], but four of five subsequent publications[18–22] failed to confirm the association. All six studies are consistent with the odds ratio we describe. The data implicate inherited variation in PPARγ in the pathogenesis of type 2 diabetes. Because the risk allele occurs at such high frequency, its modest effect translates into a large population attributable risk—influencing as much as 25% of type 2 diabetes in the general population.

Many factors may contribute to variable association results: multiple hypothesis testing, publication bias, ethnic stratification, population-specific linkage disequilibrium between markers and causal variants, inadequate statistical power, and gene-gene and gene-environment interactions. To evaluate reported associations to type 2 diabetes and test newly discovered alleles, we designed a study to address as many of these confounders as possible.

We used a multi-layered design (Table 1). Associations were first tested in 333 Scandinavian parent-offspring trios with type 2 diabetes or abnormal glucose homeostasis using transmission disequilibrium testing (TDT; ref. 23) to eliminate false positives due to population stratification. Alleles showing nominal association in the initial sample (*P*<0.05 in the direction of the original report) were tested for replication in three additional samples: 1,130 individuals from Scandinavian sibships discordant for type 2 diabetes, 481 case-control pairs from Scandinavia and 127 case-control pairs from the Saguenay-Lac-Saint-Jean region of Quebec, Canada. The use of sibships further addresses stratification (by sib-TDT (ref. 24) or related methods), whereas the case-control sample assesses population risk.

At least 16 common single-nucleotide polymorphisms (SNPs) have been associated with type 2 diabetes or related sub-phenotypes[2–17]. Notably, family-based controls were not used (except in a single report[16]). We genotyped the 16 variants in the initial 333 parent-offspring trios (Table 2). Three variants were rare or absent (<1% allele frequency). Our data cannot exclude a biological effect of these variants, but such low frequency rules out a significant impact on diabetes risk in our population. Of the remaining 13 variants, 11 showed either non-significant deviation from 50:50 transmission from heterozygous parents or a trend in the opposite direction from the original report. Because some of these variants were originally associated with quantitative phenotypes rather than diabetes itself, we genotyped an independent sample of 379 parent-offspring trios in which the offspring had normal glucose tolerance; using the TDTQ5 method[25], we found no association with the relevant phenotype (Table 2). Of course, such data do not rule out small effects; the statistical power

Table 1 • Characteristics of study populations

	Sex (male/ female)	Age (y)	Age at onset of diabetes (y)	BMI (kg/m^2)	HbA$_{1c}$ (%)	Fasting plasma glucose (mmol/l)	Plasma glucose at 2h OGTT (mmol/l)
TDT trios (Scandinavia)							
DM/IGT/IFG	176/157	39±9	37±9	27±5	5.9±1.8	7.2±2.6	8.5±2.9
NGT	187/192	31±10	–	24±5	5.1±0.5	5.2±0.5	5.6±1.1
Sibships (Scandinavia)							
DM/severe IGT sibling	283/329	65±10	56±11	29±5	7.3±1.7	9.4±3.5	14.2±5.5
NGT sibling	216/302	62±10	–	26±3	5.4±0.6	5.4±0.4	6.0±1.0
Case control (Scandinavia)							
DM/severe IGT subjects	252/229	61±10	54±11	28±5	7.5±1.8	9.7±3.2	15.3±5.5
NGT subjects	252/229	60±10	–	27±4	5.4±0.5	5.5±0.6	6.2±1.5
Case control (SLSJ)							
DM subjects	70/57	53±8	n/d	29±5	6.5±1.9	6.4±1.8	12.8±2.1
NGT subjects	70/57	52±8		29±4	5.1±0.6	5.1±0.6	6.1±1.1

Data are presented as mean ±s.d. Plasma glucose was measured at baseline (fasting) and 2 h after an oral glucose tolerance test (OGTT). DM, type 2 diabetes; IGT, impaired glucose tolerance; IFG, impaired fasting glucose; NGT, normal glucose tolerance; severe IGT, 10.0 mmol/l >120 min blood glucose >8.5 mmol/l; SLSJ, Saguenay-Lac-St-Jean. n/d, not determined.

[1]Whitehead Institute/MIT Center for Genome Research, Cambridge, Massachusetts, USA. [2]Diabetes Unit, Massachusetts General Hospital, Boston, Massachusetts, USA. [3]Harvard Medical School, Boston, Massachusetts, USA. [4]Division of Endocrinology, Children's Hospital, Boston, Massachusetts, USA. [5]Department of Endocrinology, Wallenberg Laboratory, Malmö University Hospital, University of Lund, Malmö, Sweden. [6]Montreal Genome Centre, McGill University Health Centre Research Institute, Montréal, Canada. [7]Department of Internal Medicine, Helsinki University Central Hospital, Helsinki, Finland. [8]Chicoutimi Hospital, Chicoutimi, Quebec, Canada. [9]Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA. Correspondence should be addressed to E.S.L. (e-mail: lander@genome.wi.mit.edu).

© 2000 Nature America Inc. • http://genetics.nature.com

obtained is represented in the upper limits of the confidence intervals (Table 2).

Two variants showed nominally significant replication in the initial sample: a missense variant in *PPARG* (Pro12Ala) and a silent C/T polymorphism in exon 22 of *ABCC8* (Table 2). *PPARG* encodes PPARγ, a nuclear hormone receptor that regulates adipogenesis[26] and is a target for thiazoladinediones, medications used to treat diabetes[27]. The less frequent alanine allele (16%) was associated with decreased diabetes risk (81:104, odds ratio 0.78, $P<0.045$ one-tailed). *ABCC8* encodes the drug target for sulphonylurea antidiabetic medications[28]. The less frequent 'T' allele (3%) was associated with increased diabetes risk (26:12, odds ratio 2.2, $P<0.012$ one-tailed; Table 2). For both variants, transmission ratios were not significantly different from 50:50 in non-diabetic trios, ruling out transmission ration distortion (data not shown).

Table 2 • Association results of 16 published variants in 333 DM/IGT/IFG trios

Gene	Variant	Reported effect	T	U	Ratio	95% CI	*P* value	TDTQ5
ABCC8	C/T (exon 22)	higher risk DM[9]	26	12	2.17	1.18–4.71	0.012	
PPARG	Pro12Ala	lower risk DM[12]	81	104	0.78	0.59–1.05	0.045	
IRS1	Gly972Arg	higher risk DM[8]	30	26	1.15	0.71–2.03	0.30	
ADRB2	Gln27Glu	higher risk DM[14]	96	124	0.77	0.60–1.01	NS	
INS	–23A/T (HphI)	higher risk DM[16]	104	115	0.90	0.70–1.18	NS	
		paternal transmissions only	39	45	0.76	0.57–1.36	NS	
IRS1	Ala512Pro	higher risk DM[2]	14	21	0.67	0.36–1.37	NS	
KCNJ11	Glu23Lys	higher risk DM[13]	138	154	0.90	0.73–1.15	NS	
ABCC8	T/C (intron 24)	higher risk DM[9]	127	139	0.91	0.73–1.18	NS	
TNF	–238A/G	higher HOMA[17]	14	13	1.08	0.35–2.48	0.42	NS
ADRB3	Trp64Arg	higher 2 h insulin[3]	43	43	1.00	0.67–1.56	0.50	0.20
PON2	Ala148Gly	higher fasting glucose[10]*	122	133	0.92	0.72–1.18	NS	0.20
FABP2	Ala54Thr	higher fasting insulin[4]	112	120	0.93	0.73–1.21	NS	NS
GYS1	Met416Val	higher fasting insulin[11]	16	20	0.80	0.44–1.62	NS	NS
GCGR	Gly40Ser	higher risk DM[5]	1	4		too rare to analyse		
INSR	Val985Met	higher risk DM[15]	2	6		too rare to analyse		
IAPP	Ser20Gly	higher risk DM[7]	0	0		absent from our population		

T, Number of variant alleles transmitted from heterozygous parents of diabetes/impaired glucose homeostasis offspring; U, number of variant alleles not transmitted; Ratio, transmission ratio; 95% CI, 95% confidence intervals around the transmission ratio distortion; *P* value, one-tailed nominal *P* values are reported for trends in the same direction as the original report; NS, not significant, data trended in the opposite direction from the original report; TDTQ5, for variants originally reported to be associated with a quantitative trait, the nominal one-tailed *P* value is shown using TDTQ5 method[25] to analyse the relevant trait in 379 nondiabetic trios; DM, type 2 diabetes mellitus; OGTT, oral glucose tolerance test; HOMA (a measure of insulin sensitivity) is calculated as (fasting insulin)*(fasting glucose)/22.5 (ref. 17). *The *PON2* variant was associated with higher fasting glucose in diabetic subjects only.

We attempted to further replicate these results in additional samples. The *ABCC8* association was not confirmed, as the trend in the additional samples was in the opposite direction from our initial results (sib-TDT Z-score=0.84; Scandinavian case genotypes were CC:375, CT:24, TT:1 and control genotypes were CC:363, CT:35, TT:2). These data suggest that the initial result might represent a statistical fluctuation, not unexpected when testing 13 hypotheses. By contrast, we confirmed the *PPARG* association (Table 3): the three follow-up samples demonstrated similar odds ratios (0.71, 0.74 and 0.88), with a pooled *P* value of 0.012. Combining our initial and replication samples, the *P* value is 0.002. Because our analysis was prompted by a previously reported association, and given the biological role of PPARγ, we believe these data constitute strong replication.

We then estimated the genotype relative risk (GRR) and population attributable risk of PPARγ Pro12Ala. The transmission ratio (TDT) and genotype counts (case-control) of the proline allele are consistent with codominant, additive, multiplicative or recessive models. Under any of these scenarios, the GRR of the proline allele is estimated as approximately 1.25. Because this risk allele is so common (frequency=83–87% in our samples), a GRR of 1.25 corresponds to a population attributable risk of approximately 25%. That is, these data indicate that if the population were monomorphic for the protective alanine allele, the prevalence of type 2 diabetes would be 25% lower.

We also examined whether Pro12Ala is associated with other phenotypes by applying the TDTQ5 method[25] to the 712 offspring in the 2 trio populations. No significant alteration in body mass index (BMI) or metabolic profile was observed (Table 4). We also compared the phenotypes of 206 sibling pairs concordant for diabetic status but discordant for *PPARG* genotype, and found no significant association to BMI (data not shown).

Our results resolve a controversy concerning the association between PPARγ Pro12Ala and type 2 diabetes. An initial report indicated that the alanine allele reduced diabetes risk by 75% (ref. 12). Of five[18–22] subsequent studies, four[18,20–22] failed to detect a statistically significant association, leading the authors to dismiss a role for Pro12Ala in diabetes risk; however, all published data are consistent with the modest effect that we have described (Fig. 1), although most individual studies lack sufficient sample size to reliably detect the association. In our study, a combination of samples was required to achieve adequate power. Combining all published studies yields an estimated risk ratio for the alanine allele of 0.79 (P=0.00007). Furthermore, our two family-based samples rule out population stratification as the cause of these results.

Although PPARγ Pro12Ala is reproducibly associated with type 2 diabetes, this polymorphism may not be the aetiologic variant, but rather in linkage disequilibrium with it. The case for Pro12Ala being causative is strengthened by functional differences between the proline-containing and ala-

Table 3 • Replication of the PPARγ Pro12Ala association with diabetes risk in three independent samples

	Trios* (Scandinavia)	Sibships (Scandinavia)	Case-control (Scandinavia)	Case-control (SLSJ)	
Ala alleles transmitted	81				
Ala alleles not transmitted	104				
Ala alleles in DM sibs (DAT)		37			
Ala alleles in NGT sibs (DAT)		50			
z score (Sib TDT)		1.47			
Ala allele frequency in cases			14.6%	9.4%	
Ala allele frequency in controls			16.8%	13.5%	
Genotype frequencies, cases			336:117:9	105:20:2	
Genotype frequencies, controls			324:121:17	98:24:5	
Estimated risk ratio	0.78	0.74	0.88	0.71	
P value	0.045	0.071	0.10	0.08	
P value, three replication samples					0.012
P value, all data from this study					0.002

See Table 1 for descriptions of populations. DAT, discordant allele test. *Trio data is from Table 2, presented here for comparison. Genotype frequencies are presented as proline homozygotes:heterozygotes:alanine homozygotes. All *P* values are one-tailed in the direction of Deeb *et al.*[12]

© 2000 Nature America Inc. • http://genetics.nature.com

© 2000 Nature America Inc. • http://genetics.nature.com

© 2000 Nature America Inc. • http://genetics.nature.com

nine-containing protein products[12,29] and by the lack of additional missense variants found in our systematic screening of 70 diabetic individuals (data not shown). Nonetheless, the causal variant may be some distance away (for example, in adjacent regulatory regions or conceivably in a nearby gene). To address this issue, we are identifying and testing haplotype associations in and around the gene.

We suggest that these data have several general implications. First, the low rate of replication (1/13 associations) indicates the importance of family-based controls and large, tiered samples to decrease false-positive reports[1]. Second, the previous difficulty demonstrating association of Pro12Ala with diabetes highlights the converse problem: that studies with modest sample sizes can fail to detect true associations. Third, although much can be learned from rare genetic variants (such as seven patients carrying rare amino acid changes in PPARγ; refs 30,31), PPARγ Pro12Ala highlights the potential importance of common alleles of weak effect. Pro12Ala (or possibly a variant in linkage disequilibrium with it) has a modest impact on individuals, but a dramatic effect on the human population: the risk allele is carried by billions of people, with a correspondingly high population attributable risk (25%). Critically, despite the population impact of common risk alleles, their contribution will be impossible to discover by linkage analysis: in a case such as Pro12Ala, the risk allele will typically be transmitted from both parents, requiring a genome scan of roughly 3 million sib pairs to obtain a lod score of 3. Thus, the genetic dissection of common diseases will surely involve association studies performed on large population samples.

Table 4 • Analysis of PPARγ Pro12Ala and metabolic phenotypes

Trait	Trio Population	Trait value (mean±s.d.) Pro/Pro	Pro/Ala	Ala/Ala	TDT Q5 *P* value
body mass index	DM/IGT/IFG	27.6±5.2	27.6±5.1	24.3±5.3	0.12
(kg/M²)	NGT	24.8±4.4	25.8±4.7	28.3±7.6	0.78
systolic BP	DM/IGT/IFG	129±16	131±19	120±17	0.06
(mm Hg)	NGT	120 ±13	122±16	129±2	0.65
diastolic BP	DM/IGT/IFG	79±10	81±12	72±11	0.005
(mm Hg)	NGT	74±11	75±13	77±4	0.98
cholesterol	DM/IGT/IFG	5.4±1.1	5.5±1.0	5.2±1.1	0.13
(mM)	NGT	5.0±0.9	5.0±1.0	5.0±0.6	0.52
HDL	DM/IGT/IFG	1.2±0.3	1.2±0.3	1.4±0.4	0.02
(mM)	NGT	1.4±0.4	1.4±0.3	1.3±0.5	0.65
triglycerides	DM/IGT/IFG	1.9±1.5	1.6±0.8	1.6±1.4	0.85
(mM)	NGT	1.1±0.6	1.1±0.6	1.0±0.4	0.62
fasting glucose	DM/IGT/IFG	7.3± 2.7	7.2±2.3	5.5±0.6	0.19
(mM)	NGT	5.3±0.5	5.3±0.5	5.5±0.1	0.52
glucose 120 OGTT	DM/IGT/IFG	8.8±3.1	8.3±3.1	8.2±1.0	0.82
(mM)	NGT	5.5±1.2	5.8±1.0	6.2±1.2	0.97
fasting insulin	DM/IGT/IFG	10.7±6.9	14.0±17.0	6.8±2.8	0.16
(mU/L)	NGT	7.1±3.9	7.8±3.6	5.9±1.9	0.46
insulin 120 OGTT	DM/IGT/IFG	56.7±44.0	58.4±42.3	42.4±24.2	0.86
(mU/l)	NGT	33.9±27.0	34.6±23.0	21.0±13.0	0.44
HOMA	DM/IGT/IFG	3.7±3.6	5.3±11.8	1.7±0.8	0.12
	NGT	1.7±1.0	1.8±0.8	1.4±0.4	0.59

Metabolic phenotypes were evaluated in parent-offspring trios with both abnormal glucose homeostasis (DM/IGT/IFG, n=333) and normal glucose tolerance (NGT, n=379), in an oral glucose tolerance test (OGTT). For comparison purposes, the mean trait value and standard deviation for each genotypic class is presented. Note that the number of Ala/Ala homozygotes is very small, and thus the accuracy of trait measurements in this group is limited. HOMA, a measure of insulin sensitivity, is calculated as in Table 2. The reported *P* values are calculated using the family-based TDTQ5 (ref. 25). None of the *P* values are significant after correction for the number of hypotheses tested.

Methods

Patient populations. The clinical characteristics of the subjects included in the study are shown (Table 1). We selected Scandinavian families from three distinct family collections in Sweden and Finland: the Botnia collection, consisting of type 2 diabetic individuals from the Botnia region in western Finland[32]; the Helsinki Collection, including 290 multiplex families from northern and eastern Finland; and the Malmö collection, including 275 multiplex families from southern Sweden (A. Parker *et al.*, manuscript submitted). All subjects gave informed consent and the study was been approved by each of the three local ethics committees. The subjects underwent extensive phenotyping including oral glucose tolerance tests (OGTT), with venous measurements of blood glucose, serum insulin and C-peptide following ingestion of 75 g glucose. The WHO98 definitions of type 2 diabetes, impaired glucose tolerance (IGT) and impaired fasting glucose (IFG) were used[33]. To avoid confounding with type 1 diabetes, patients with age of onset under 18 years, GAD-ab ≥5 reference index units[32] and/or measured c-peptide levels ≤0.2 were excluded. Families with segregating mutations known to cause MODY diabetes were excluded.

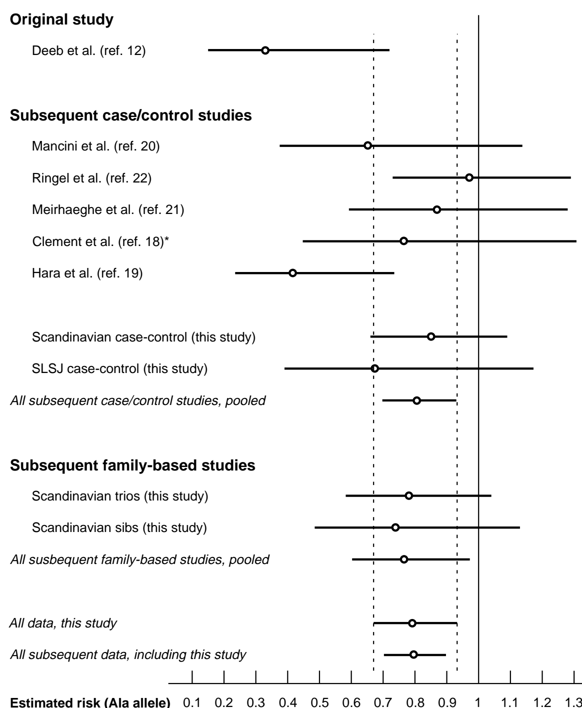

Fig. 1 Estimated risk (with 95% confidence intervals) for PPARγ Pro12Ala. For each study, the circle represents the estimated risk for the alanine allele and the line indicates the extent of the 95% confidence interval around this estimate. The dashed lines indicate the 95% confidence interval for pooled data in this study. Data are shown for odds ratios based on allele counts (for case-control studies) and the transmission ratios in trios or allele counts in phenotypically discordant sibs (for the family-based studies). Under a multiplicative model, these are all estimators of genotype relative risk. *Clement *et al.*[18] also studied severely obese individuals (BMI>40). As 56% of this group had abnormal glucose homeostasis at initial evaluation, and the likelihood of conversion to diabetes in the non-diabetic 'control' individuals was high, data from this group were not included. Inclusion of this group does not significantly affect either the pooled odds ratios or 95% confidence intervals. SLSJ, Saguenay-Lac-St-Jean.

© 2000 Nature America Inc. • http://genetics.nature.com

© 2000 Nature America Inc. • http://genetics.nature.com

Trio populations. Because transmission disequilibrium testing requires two living parents, a parent-offspring trio population is biased towards patients of young age. When studied, such individuals may not yet manifest the complete disease phenotype for a late-onset disorder such as type 2 diabetes. For this reason, we chose to include patients with IGT and IFG in our trio population. Epidemiological studies have shown that these disorders often constitute a 'pre-diabetic' state: between 25 and 50% of all subjects with IGT convert to type 2 diabetes within 10 years[34], and patients with IFG convert to type 2 diabetes at a similar rate[35]. Because our patients with IFG or IGT are young (average age 30) and are selected from families with multiple diabetic members, their abnormal glucose homeostasis serves as a strong predictor of eventual development of diabetes. We identified 333 such trios: 99 offspring had IFG, 108 had IGT and 126 had type 2 diabetes. To investigate phenotypes in a non-diabetic population, we studied a non-overlapping sample of 379 parent-offspring trios; all offspring were NGT. The offspring have a waist/hip ratio from either the upper quintile or the lowest decile as defined in a healthy control population.

Replication samples. To evaluate findings with a nominal *P* value <0.05 in the initial trio population, we also selected two independent replication samples from the same geographic areas. After excluding all subjects with age at onset of type 2 diabetes ≤30 y, we identified 481 of cases with type 2 diabetes or severe IGT (10.0 mmol/l > 120 min blood glucose ≥ 8.5 mmol/l). Because there was no requirement for parental DNA in these samples, older subjects and a more stringent definition of affectation status were used. In our experience, patients with 'severe IGT' have a very high rate (25%) of conversion to overt type 2 diabetes within a three-year follow-up period (P. Almgren and L.G., unpublished data). We then assembled 481 age-, gender- and geographically matched normal glucose tolerance (NGT) controls. Specifically, affected individuals were matched to controls from the same geographic regions (Botnian cases were matched with controls from Botnia, and so on). In addition, 612 diabetic and 518 NGT sibs from 367 sibships discordant for diabetes or severe IGT were selected. The type 2 diabetic siblings had ages of onset ≥30 y and the non-diabetic (NGT) siblings were all older than 45. The same criteria for excluding MODY and type 1 diabetes were used for these samples. In addition, we tested a case-control sample from a geographically distinct population. These samples were unrelated adults (above 18 y) recruited from the region of Saguenay-Lac-St-Jean (Northeastern Quebec) and of French Canadian descent. Patients newly diagnosed with type 2 diabetes (using WHO98 criteria following a 75 g oral glucose load) were included, and each patient was age- and sex-matched with an individual from the same population having a normal glucose tolerance.

Genotyping. Genotyping was performed by either single-base extension with fluorescence resonance energy transfer (SBE-FRET; ref. 36) or single-base extension with fluorescence polarization (SBE-FP; ref. 37), using modified protocols (S.B. *et al.*, manuscript in preparation). PCR primers, SBE primers and PCR conditions are available on request. A mix (5 μl; containing 1 U shrimp alkaline phosphatase (Roche), 2 U exonuclease I (Epicentre), 150 mM Tris HCl, pH 8, and 15 mM $MgCl_2$) was added to 10 μl PCR product in a 384-well black polypropylene plate (MJ Research) and incubated at 37 °C for 45 min and 80 °C for 15 min. SBE reaction mixture (5 μl) was added (containing 0.5 U Thermosequenase DNA polymerase (Amersham), 200 mM Tris HCl, pH 9.5, 8 mM $MgCl_2$, 1.6 μM SBE primer and 0.2 μM each of ROX- and TAMRA-labelled ddNTPs specific to the alleles being assayed). Reactions were denatured for 2 min at 92 °C, followed by 50 cycles of 92 °C for 10 s and 50 °C for 30 s. Plates were read on an Analyst Fluorescence Plate-reader (LJL Biosystems); polarized fluorescence signals (mP values) for ROX and TAMRA were used to assign genotypes. For SBE-FRET, the SBE mix contained 5´-FAM-labelled SBE primer (0.4 μM) and ROX- and TAMRA-labelled ddNTPs (0.4 μM each). Six cycles of SBE (96 °C×15 s, 50 °C×30 s, 60 °C×30 s) were performed in an ABI 7700 (Perkin Elmer), reading during the extension phase. The difference in fluorescence between cycle 6 and cycle 1 was calculated for ROX and TAMRA after matrix correction for spectral overlap and these differences used to assign genotypes. In both methods, genotypes were assigned by clustering data from 96 to 960 individuals and assignments were reviewed by at least two individuals.

Statistical analysis. All phenotype data are expressed as mean±s.d. Transmission ratios from heterozygous parents in the trio population were calculated for transmission disequilibrium testing (TDT; ref. 23). The discordant sib sample was evaluated using the sib TDT comparing allele frequencies between affected and unaffected siblings[24]. To estimate the relative risk, we also performed a discordant allele test (DAT) with a randomly chosen affected sib and the oldest unaffected sib from each sib-ship[38]. Because only one sib-pair was used from each sibship, this is also a 'pure' test of association. χ^2 distributions were used to calculate *P* values for the TDT and DAT, and 95% confidence intervals were calculated using a binomial distribution. For the case-control samples, standard allelic odds ratios were calculated, and *P* values for differences in allele frequencies between cases and controls were determined by using a χ^2 distribution with one degree of freedom. For replication tests, one-tailed *P* values were calculated based on the direction of effect described in the original report. To pool *P* values, we calculated the expected mean and variance of the number of risk alleles found in cases given the number of samples analysed. These were combined to generate an overall mean and variance, and 2-tailed *P* values were calculated. This approach also allows combination of case-control studies with TDT and sib-TDT data, similar to the approach described[26]. To evaluate the effect of alleles on quantitative phenotypes in the trio populations, we used the TDTQ5 method[25]. TDTQ5 is a family-based test used to analyse the effect of offspring genotype on phenotype; by first regressing against parental genotype, stratification is avoided. Quantitative phenotypes were separately normalized for age for each gender before analysis; for all traits reported with TDTQ5, the phenotype distributions were found to be consistent with normality. Population attributable risk was calculated in standard fashion as $PAR=(X-1)/X$, where $X=(1-f)^2 + 2f(1-f)\gamma + f^2\gamma^2$, γ is the estimated GRR (in this case 1.25), f is the frequency of the risk allele (in this case ~0.85) and a multiplicative model is assumed. As noted in the text, similar values of PAR are obtained assuming an additive or codominant model.

Acknowledgements

We thank B. Langhorst and P. Almgren for contributions to patient and genotype databases; A. Berglund, L. Roslund and M. Svensson for skilful laboratory assistance; M. Erdos for PCR primers to IRS1; C. Cellier, J. Faith, P. Perron, G. Houde, C. Betard and M.-E. Allard for their contributions to the Canadian study; The Botnia Research Team for clinical contributions; and members of the Whitehead Institute Center for Genome Research for helpful discussions. J.N.H. and D.M.A. are recipients of the Post-doctoral Fellowship for Physicians from the Howard Hughes Medical Institute. C.M.L. is sponsored by the Fund for Strategic Research through NNCR. L.G. receives support from the Sigrid Juselius Foundation, the JDF Wallenberg Foundation, the Finnish Diabetes Research Foundation, The Swedish Medical Research Council, the Novo-Nordisk Foundation and an EC Paradigm grant (BH K99JD-12812-01A; L.C.G). This work was supported in part by grants from Affymetrix Inc., Millennium Pharmaceuticals Inc. and Bristol-Myers Squibb Company to E.S.L.

Received 19 April; accepted 2 August 2000.

© 2000 Nature America Inc. • http://genetics.nature.com

1. Editorial. Freely associating. *Nature Genet.* **22**, 1–2 (1999).
2. Almind, K. *et al.* Aminoacid polymorphisms of insulin receptor substrate-1 in non-insulin- dependent diabetes mellitus. *Lancet* **342**, 828–832 (1993).
3. Widen, E. *et al.* Association of a polymorphism in the β 3-adrenergic-receptor gene with features of the insulin resistance syndrome in Finns. *N. Engl. J. Med.* **333**, 348–351 (1995).
4. Baier, L.J. *et al.* An amino acid substitution in the human intestinal fatty acid binding protein is associated with increased fatty acid binding, increased fat oxidation, and insulin resistance. *J. Clin. Invest.* **95**, 1281–1287 (1995).
5. Hager, J. *et al.* A missense mutation in the glucagon receptor gene is associated with non-insulin-dependent diabetes mellitus. *Nature Genet.* **9**, 299–304 (1995).
6. Kan, M. *et al.* Frequency of mutations of insulin receptor gene in Japanese patients with NIDDM. *Diabetes* **44**, 1081–1086 (1995).
7. Sakagashira, S. *et al.* Missense mutation of amylin gene (S20G) in Japanese NIDDM patients. *Diabetes* **45**, 1279–1281 (1996).
8. Sigal, R.J., Doria, A., Warram, J.H. & Krolewski, A.S. Codon 972 polymorphism in the insulin receptor substrate-1 gene, obesity, and risk of noninsulin-dependent diabetes mellitus. *J. Clin. Endocrinol. Metab.* **81**, 1657–1659 (1996).
9. Inoue, H. *et al.* Sequence variants in the sulfonylurea receptor (SUR) gene are associated with NIDDM in Caucasians. *Diabetes* **45**, 825–831 (1996).
10. Hegele, R.A. *et al.* Paraoxonase-2 gene (PON2) G148 variant associated with elevated fasting plasma glucose in noninsulin-dependent diabetes mellitus. *J. Clin. Endocrinol. Metab.* **82**, 3373–3377 (1997).
11. Shimomura, H. *et al.* A missense mutation of the muscle glycogen synthase gene (M416V) is associated with insulin resistance in the Japanese population. *Diabetologia* **40**, 947–952 (1997).
12. Deeb, S.S. *et al.* A Pro12Ala substitution in PPARγ2 associated with decreased receptor activity, lower body mass index and improved insulin sensitivity. *Nature Genet.* **20**, 284–287 (1998).
13. Hani, E.H. *et al.* Missense mutations in the pancreatic islet β cell inwardly rectifying K+ channel gene (KIR6.2/BIR): a meta-analysis suggests a role in the polygenic basis of type II diabetes mellitus in Caucasians. *Diabetologia* **41**, 1511–1515 (1998).
14. Ishiyama-Shigemoto, S., Yamada, K., Yuan, X., Ichikawa, F. & Nonaka, K. Association of polymorphisms in the β2-adrenergic receptor gene with obesity, hypertriglyceridaemia, and diabetes mellitus. *Diabetologia* **42**, 98–101 (1999).
15. Hart, L.M. *et al.* Prevalence of variants in candidate genes for type 2 diabetes mellitus in The Netherlands: the Rotterdam study and the Hoorn study. *J. Clin. Endocrinol. Metab.* **84**, 1002–1006 (1999).
16. Huxtable, S.J. *et al.* Analysis of parent-offspring trios provides evidence for linkage and association between the insulin gene and type 2 diabetes mediated exclusively through paternally transmitted class III variable number tandem repeat alleles. *Diabetes* **49**, 126–130 (2000).
17. Day, C.P. *et al.* Tumour necrosis factor-α gene promoter polymorphism and decreased insulin resistance. *Diabetologia* **41**, 430–434 (1998).
18. Clement, K. *et al.* The Pro115Gln and Pro12Ala PPAR γ gene mutations in obesity and type 2 diabetes. *Int. J. Obes. Relat. Metab. Disord.* **24**, 391–393 (2000).
19. Hara, K. *et al.* The Pro12Ala polymorphism in PPAR γ2 may confer resistance to type 2 diabetes. *Biochem. Biophys. Res. Commun.* **271**, 212–216 (2000).
20. Mancini, F.P. *et al.* Pro12Ala substitution in the peroxisome proliferator-activated receptor-γ2 is not associated with type 2 diabetes. *Diabetes* **48**, 1466–1468 (1999).
21. Meirhaeghe, A. *et al.* Impact of the peroxisome proliferator activated receptor γ2 Pro12Ala polymorphism on adiposity, lipids and non-insulin-dependent diabetes mellitus. *Int. J. Obes. Relat. Metab. Disord.* **24**, 195–199 (2000).
22. Ringel, J., Engeli, S., Distler, A. & Sharma, A.M. Pro12Ala missense mutation of the peroxisome proliferator activated receptor γ and diabetes mellitus. *Biochem. Biophys. Res. Commun.* **254**, 450–453 (1999).
23. Spielman, R.S., McGinnis, R.E. & Ewens, W.J. Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). *Am. J. Hum. Genet.* **52**, 506–516 (1993).
24. Spielman, R.S. & Ewens, W.J. A sibship test for linkage in the presence of association: the sib transmission/disequilibrium test. *Am. J. Hum. Genet.* **62**, 450–458 (1998).
25. Allison, D.B. Transmission-disequilibrium tests for quantitative traits. *Am. J. Hum. Genet.* **60**, 676–690 (1997).
26. Tontonoz, P., Hu, E. & Spiegelman, B.M. Stimulation of adipogenesis in fibroblasts by PPAR γ 2, a lipid-activated transcription factor. *Cell* **79**, 1147–1156 (1994).
27. Lehmann, J.M. *et al.* An antidiabetic thiazolidinedione is a high affinity ligand for peroxisome proliferator-activated receptor γ (PPAR γ). *J. Biol. Chem.* **270**, 12953–12956 (1995).
28. Aguilar-Bryan, L. *et al.* Cloning of the β cell high-affinity sulfonylurea receptor: a regulator of insulin secretion. *Science* **268**, 423–426 (1995).
29. Masugi, J., Tamori, Y., Mori, H., Koike, T. & Kasuga, M. Inhibitory effect of a proline-to-alanine substitution at codon 12 of peroxisome proliferator-activated receptor-γ 2 on thiazolidinedione-induced adipogenesis. *Biochem. Biophys. Res. Commun.* **268**, 178–182 (2000).
30. Barroso, I. *et al.* Dominant negative mutations in human PPARγ associated with severe insulin resistance, diabetes mellitus and hypertension. *Nature* **402**, 880–883 (1999).
31. Ristow, M., Muller-Wieland, D., Pfeiffer, A., Krone, W. & Kahn, C.R. Obesity associated with a mutation in a genetic regulator of adipocyte differentiation. *N. Engl. J. Med.* **339**, 953–959 (1998).
32. Groop, L. *et al.* Metabolic consequences of a family history of NIDDM (the Botnia study): evidence for sex-specific parental effects. *Diabetes* **45**, 1585–1593 (1996).
33. Alberti, K.G. & Zimmet, P.Z. Definition, diagnosis and classification of diabetes mellitus and its complications. Part 1: diagnosis and classification of diabetes mellitus provisional report of a WHO consultation. *Diabet. Med.* **15**, 539–553 (1998).
34. Edelstein, S.L. *et al.* Predictors of progression from impaired glucose tolerance to NIDDM: an analysis of six prospective studies. *Diabetes* **46**, 701–710 (1997).
35. Shaw, J.E. *et al.* Impaired fasting glucose or impaired glucose tolerance. What best predicts future diabetes in Mauritius? *Diabetes Care* **22**, 399–402 (1999).
36. Chen, X., Zehnbauer, B., Gnirke, A. & Kwok, P.Y. Fluorescence energy transfer detection as a homogeneous DNA diagnostic method. *Proc. Natl Acad. Sci. USA* **94**, 10756–10761 (1997).
37. Chen, X., Levine, L. & Kwok, P.Y. Fluorescence polarization in homogeneous nucleic acid analysis. *Genome Res.* **9**, 492–498 (1999).
38. Boehnke, M. & Langefeld, C.D. Genetic association mapping based on discordant sib pairs: the discordant-alleles test. *Am. J. Hum. Genet.* **62**, 950–961 (1998).

© 2000 Nature America Inc. • http://genetics.nature.com