

## ELECTRONIC LETTER

# Relation of type 2 diabetes to individual admixture and candidate gene polymorphisms in the Hispanic American population of San Luis Valley, Colorado

E J Parra, C J Hoggart, C Bonilla, S Dios, J M Norris, J A Marshall, R F Hamman, R E Ferrell, P M McKeigue, M D Shriver

*J Med Genet* 2004;41:e116 (<http://www.jmedgenet.com/cgi/content/full/41/11/e116>). doi: 10.1136/jmg.2004.018887

The prevalence of type 2 diabetes is higher in populations of Native American ancestry, and in Hispanic American populations formed by admixture between Europeans and Native Americans, than in populations of European ancestry.<sup>1</sup> One approach to distinguishing between environmental and genetic explanations for this difference is to study the relationship of type 2 diabetes risk to individual admixture proportions (the proportions of an individual's genome that are of European and Native American ancestry). With only a few markers informative for ancestry, it is possible to estimate the average admixture proportions of any Hispanic American population. In such analyses, it has been possible to demonstrate that the prevalence of type 2 diabetes in Hispanic Americans in the south western United States varies with the average Native American admixture proportion of these populations.<sup>2-4</sup> In the Native American population of Gila River, Arizona, USA, European admixture is associated with lower prevalence of type 2 diabetes.<sup>5</sup> However, it has not been possible to demonstrate an association of type 2 diabetes with individual admixture proportions within an Hispanic American population. To estimate the admixture proportions of an individual accurately requires a larger panel of markers: at least 40 markers with average frequency differentials of 0.6 are required to estimate the admixture of an individual with a standard error of no more than 0.1.<sup>6</sup> It is now possible to identify relatively large numbers of such ancestry informative markers from data accumulating in the public domain. For this study we typed a panel of 21 markers chosen to have large differences in frequency between European, Native American, and West African ancestry.

The possible relationship of type 2 diabetes risk to individual admixture proportions within Hispanic American populations complicates the interpretation of associations of type 2 diabetes with candidate gene polymorphisms within these populations. If admixture proportions vary between individuals (hidden population stratification) and the risk of type 2 diabetes varies with individual admixture proportions, this will confound allelic associations with type 2 diabetes at any loci where allele frequencies differ between Europeans and Native Americans. We have shown that in recently admixed populations associations are often observed between unlinked genetic markers.<sup>7-9</sup> Thus, when carrying out association studies in admixed populations, it is necessary to control for possible confounding by population stratification. The classic approach to this has been to type parents as controls, but for a late onset disease such as type 2 diabetes parents of cases are not usually available for study. By typing ancestry informative markers, we can estimate individual admixture and control for it as a confounder. The most satisfactory approach to this is to fit a statistical model of population admixture, individual admixture, and the

## Key points

- The prevalence of type 2 diabetes is higher in Hispanic American populations than in populations of European ancestry. The objectives of this study were to distinguish between genetic and environmental explanations for this ethnic difference in disease risk, and to test candidate gene polymorphisms for association with type 2 diabetes in the Hispanic American population of San Luis Valley, Colorado, USA.
- We genotyped 11 single nucleotide polymorphisms in five candidate genes (CAPN10, GNB3, PPARG, and ABCC8/KCNJ11), together with a panel of 21 ancestry informative markers, in a sample of 261 controls and 185 diabetic individuals. The ADMIXMAP program was used to model the effects of admixture.
- Type 2 diabetes risk varies with proportion of Native American ancestry in San Luis Valley Hispanics, but this relationship is confounded by socioeconomic factors. We were able to confirm modest effects of ABCC8/KCNJ11 variants on type 2 diabetes risk, but observed no association of type 2 diabetes with four CAPN10 markers.

relationship of disease risk to individual admixture. Tests for allelic association with the disease can then be adjusted for the confounder. Although the statistical model is based on a straightforward application of the laws of mendelian genetics, to fit such a model in practice requires bayesian computationally intensive methods. We have developed a general purpose program (ADMIXMAP) for modelling admixture based on this approach, and have demonstrated the ability to distinguish associations of a trait with alleles at loci that are linked to a trait locus from associations with unlinked loci that are generated by population stratification.<sup>9-10</sup> Where two or more loci in the same gene have been typed, the program can also model the unobserved haplotypes, given phase-unknown genotype data.

In this paper, we evaluated the association of type 2 diabetes, fasting insulin, and body mass index (BMI) with polymorphisms in five candidate genes in a sample from the Hispanic American population of San Luis Valley, Colorado, USA. The candidate genes are calpain 10 (CAPN10), guanine nucleotide binding protein, beta polypeptide 3 (GNB3),

**Abbreviations:** AIMs, ancestry informative markers; LD, linkage disequilibrium; MCMC simulation, Markov chain Monte Carlo simulation; SLVDS, San Luis Valley Diabetes Study

peroxisome proliferative activated receptor, gamma (PPARG), ATP binding cassette, subfamily C, member 8 (ABCC8), and potassium inwardly rectifying channel, subfamily J, member 11 (KCNJ11). We analysed four markers within the CAPN10 gene (UCSNP 19, 43, 44, and 63 polymorphisms), one marker within the GNB3 gene (C825T polymorphism), two markers within the PPARG gene (Pro12Ala and exon 6 C→T polymorphisms), and four markers located within the ABCC8 (SUR1) and KCNJ11 genes, which are closely linked on chromosome 11 (ABCC8 exon 16 splice acceptor site, ABCC8 exon 31 G→A polymorphism, ABCC8 exon 33 G→T polymorphism, and KCNJ11-E32K polymorphism). In previous studies, associations of type 2 diabetes or related traits such as obesity and insulin resistance have been detected with these polymorphisms, or other sites in the same genes. However, with the exception of CAPN10, there have been few studies of these associations in Hispanic American populations.

## METHODS

### Populations

The study sample was selected from participants of the San Luis Valley Diabetes Study (SLVDS), a geographically based study of the natural history, incidence, and risk factors for type 2 diabetes conducted in the counties of Alamosa and Conejos in southern Colorado. These counties are 43.6% Hispanic American. Informed consent from all participants and approval by the Institutional Review Board of the University of Colorado were obtained prior to data collection. Additional approval of this work was obtained through the Penn State University Institutional Review Board (ORC# 00M0453). The procedure for selecting the SLVDS study subjects has been described in detail by Hamman *et al.*<sup>11</sup> In summary, persons with type 2 diabetes in the study area were identified through all health care facilities and through advertisement in local newspapers, presentations to local organisations, and local radio programs. Eligible subjects with a medical diagnosis of type 2 diabetes were 20–75 years of age, residents of the study area, mentally competent, and spoke either Spanish or English. The baseline data collection clinic (1984–1988) was attended by 82% of eligible subjects (n = 440). Controls were selected using a two stage sampling method. First, 57% of all occupied structures in the two county area were sampled and enumerated. Enumerated persons 20–75 years of age were the sampling frame for the second stage of control selection where subjects were randomly selected within age, sex, ethnic group, and county strata to match the age and sex distribution of persons with type 2 diabetes. Some 67% of eligible controls (n = 1351) attended the baseline clinic. The total sample described above comprised unrelated individuals of Hispanic ancestry and English or Anglo ancestry. We have focused our analyses only on the persons who identified as Hispanic (Mexican, Mexican American or Chicano, Spanish/Hispanic). The SLVDS Hispanic sample included 185 individuals with a history of diabetes confirmed by oral glucose tolerance test and 261 controls with confirmed normal glucose tolerance. A total of 128 participants reported their ethnicity (based on the 1980 US census question) as Mexican, Mexican American, or Chicano, and 318 as other Spanish/Hispanic. Data on other relevant phenotypes, such as fasting insulin and BMI, were also collected in the SLVDS. Data on household income and years of education were used to assess socioeconomic status. Age at baseline visit ranged between 21 and 75 years old. Given the relatively young age of some of the subjects, it was expected that some of the control subjects would develop type 2 diabetes as they grew older. As such our analysis is conservative with respect to testing for associations with

**Table 1** Description of the SLVDS Hispanic sample

	SLVDS cases	SLVD controls
<b>n</b>	185	261
Males	70	123
Females	115	138
<b>Age at baseline visit</b>		
Mean	57.14	52.54
SD	10.85	12.59
Minimum/maximum	26/74	21/75
<b>BMI</b>		
Mean	29.32	26.17
SD	4.95	4.44
Minimum/maximum	17.23/48.38	15.84/40.03
<b>Fasting insulin (mU/l)</b>		
Mean	22.82	13.46
SD	12.51	7.15
Minimum/maximum	7/69	3/44
<b>Income*</b>		
Mean	3.59	4.68
SD	2.36	2.38
Minimum/maximum	1/9	1/10
<b>Education (years)</b>		
Mean	9.47	10.81
SD	3.45	3.75
Minimum/maximum	0/17	0/17

\*Total household income in ten categories.

either ancestry or locus specific tests. Table 1 summarises the characteristics of this Hispanic sample from San Luis Valley.

Allele frequencies in unadmixed Europeans were estimated from samples from Spain, Germany, Ireland, and Britain. Allele frequencies in unadmixed Native Americans were estimated from Mayans and from populations in the south western United States (Pima, Cheyenne, and Pueblo). Allele frequencies in West Africans were estimated from samples from Nigeria, Sierra Leone, and Central African Republic. More information on these unadmixed samples is available in dbSNP (<http://www.ncbi.nlm.nih.gov/SNP/index.html>), under the submitter handle PSU-ANTH.

### Genetic analysis

We typed 11 polymorphisms within five candidate genes. The description of the markers in each gene, the primer sequences, and the PCR conditions are given in table 2. The sequence surrounding the polymorphisms of the Calpain-10 gene was kindly provided to us by Dr Nancy Cox. The polymorphism of UCSNP-19 is based on a difference of 32 bp between the two alternative alleles, and this marker was genotyped using conventional agarose electrophoresis. The characterisation of the remaining polymorphisms was based on the presence/absence of a restriction site. In some cases, due to the absence of a natural restriction site, primer sequences were modified to create a restriction site polymorphism (table 2). After initial denaturation for 5 min at 94°C, DNA samples were amplified at the denaturation/annealing/extension temperatures specified for each marker, followed by a final extension for 5 min at 72°C. After PCR, these markers were digested using the appropriate restriction enzyme, following the recommendations of the supplier. For details of the restriction enzymes used, see table 2. After digestion, genotypes were characterised by means of conventional agarose electrophoresis (SNP 43, SNP 44, and SNP 63), or alternatively, by melting curve analysis. Details of the melting curve analysis method employed have been described in a previous manuscript.<sup>12</sup> For a subset of the samples, genotypes were characterised using both the McSNP method and conventional agarose electrophoresis, with consistent results.

A panel of 21 ancestry informative markers (AIMs) was also genotyped in the San Luis Valley sample. These markers

**Table 2** Candidate gene polymorphisms typed

Markers	Variation	5'-3' Forward/reverse primers	PCR (35 cycles)	MgCl <sub>2</sub> (mM)	Enzyme
<b>CAPN10</b>					
UCSNP-19 (Nancy Cox: pers. comm.)	2 or 3 repeats of a 32 bp unit	F-GTTTGGTCTCTTCAGCGTGGAG R-ATGAACCTTGGCAGGGTCTAAG	94°C 30 s 60°C 40 s 72°C 40 s	1.5	
UCSNP-43 (Nancy Cox: pers. comm.)	G→A	F-CACGCTTGCTGTGAAGTAATGC R-CTCTGATCCCATGGTCTGTAG	94°C 45 s 55°C 45 s 72°C 60 s	1.5	NsiI
UCSNP-44 (Nancy Cox: pers. comm.)	C→T	F-CAGGGCGCTCACGCTTGCCG R-TCCTCACCAAGTCAAGGCTTA	94°C 45 s 61°C 45 s 72°C 45 s	2.0	BsAI
UCSNP-63 (Nancy Cox: pers. comm.)	C→T	F-AGGGCCTGACGGGGTGGCG R-AGCACTCCAGCTCCTGATC	94°C 60 s 66.5°C 60 s 72°C 60 s	3.0	HhaI
<b>GNB3</b>					
GNB3-C825T NCBI rs5443	C→T	F-CATCATCTGCGGCATCAC R-AATAGTAGGGCCACTGAG	94°C 30 s 58°C 30 s 72°C 30 s	1.5	BseDI
<b>PPARG</b>					
PPARG-Pro12Ala NCBI rs1801282	G→C	F-CAAACCCCTATTCCATGCTG R-AGTGAAGGAATCGCTTCCG	94°C 30 s 58°C 45 s 72°C 30 s	2.5	HpaII
PPARG-E6 NCBI rs3856806	C→T	F-CTCAGACAGATTGTCACGGAAC R-TTCTTGATCACCTGCAGTAGC	94°C 30 s 60°C 30 s 72°C 30 s	1.5	NlaIII
<b>ABCC8 (SUR1)</b>					
ABCC8-E16 splice acceptor site NCBI rs1799854	C→T	F-GTAATGGTTGTCAGACTCC R-CTAGAAGGAGCGAGGACT	94°C 30 s 57°C 30 s 72°C 30 s	2.5	PstI
ABCC8-E31 NCBI rs4148643	G→A	F-GGAGTACATCGGTGCATGTG R-CCATTAGGGGGTAGGTAAGG	94°C 30 s 56°C 30 s 72°C 30 s	1.5	BsII
ABCC8-E33 NCBI rs757110	G→T	F-GCTACGACAGCTCCCTGAAG R-CCCCTGCTCTGACCTCT	94°C 30 s 58°C 30 s 72°C 30 s	2.0	MwoI
<b>KCNJ11</b>					
KCNJ11-E23K NCBI rs5219	G→A	F-ATCATCCCCGAGGAATACG R-GCCCTTCTGGACACAAGGC	94°C 30 s 58°C 30 s 72°C 30 s	2.5	BanII

For some markers, primers have been modified to introduce a restriction site polymorphism, and the altered base is shown in bold letters.

show large differences in frequency between the parental populations (mainly European and Native American), and were used to control for the presence of genetic structure due to admixture. The panel of AIMs includes an Alu insertion

polymorphism (PV92), four short insertion–deletion polymorphisms (MID-575, MID-52, MID-161, and MID-93), and 16 single nucleotide polymorphisms (SNPs). Relevant information about these markers is provided in table 3. We have

**Table 3** Ancestry informative markers with estimated ancestry specific frequencies of allele 1

Marker*	NCBI rs	Chromosomal location	European	Native American	African
MID-575	rs140864	1p34.3	0.584	0.007	0.124
FY-null	rs2814778	1q23.2	0.993	1.000	0.001
F13B	rs6003	1q31.3	0.104	0.018	0.704
TSC1102055	rs2065160	1q32.1	0.921	0.137	0.487
WI-11153	rs17203	3p12.3	0.172	0.819	0.785
MID-52	rs16344	4q24	0.082	0.763	0.263
SGC30610	rs3309	5q11.2	0.300	0.699	0.401
WI-17163	rs3340	5q33.2	0.197	0.690	0.054
WI-4019	rs2161	7q22.1	0.295	0.618	0.430
WI-11909	rs2695	9q21.31	0.845	0.181	0.805
D11S429	rs594689	11q13.1	0.440	0.119	0.087
TYR	rs1042602	11q14.3	0.485	0.034	0.005
DRD2 "Taqd"	rs1800498	11q23.2	0.630	0.045	0.135
DRD2 "Bcl I"	rs1079598	11q23.2	0.135	0.665	0.063
WI-14319	rs2862	15q14	0.142	0.716	0.386
CYP19	rs4646	15q21.2	0.287	0.741	0.332
PV92	rs3138523	16q23.3	0.171	0.792	0.225
WI-7423	rs2816	17p13.1	0.517	0.058	0.000
CKM	rs4884	19q13.32	0.257	0.904	0.164
MID-161	rs16440	20q11.22	0.508	0.109	0.637
MID-93	rs16383	22q13.2	0.082	0.763	0.263

\*DRD2 "Taqd" and DRD2 "Bcl I" are closely linked on chromosome 11, located 4.7 kb apart.

reported in a recent manuscript the allele frequencies of these 21 AIMs in the sample from San Luis Valley and also in samples of the relevant parental populations.<sup>13</sup>

### Statistical analysis

To test for association of the markers located within the type 2 diabetes candidate genes with the traits under study, we used the program ADMIXMAP<sup>10</sup> (available at <http://www.lshmt.ac.uk/eu/genetics/index.html#admixmap>). This is a general purpose program for modelling population admixture with genotype and phenotype data, based on a combination of bayesian and classical methods. As we have described the statistical methods used in this program in detail previously<sup>10</sup> and demonstrated their application to studies of skin pigmentation in this Hispanic American population sample, only an outline is given here. For this analysis, the Hispanic American population was modelled as formed by admixture between three subpopulations: European, Native American, and West African. The program fits a hierarchical model for the distribution of admixture proportions in the population, the admixture proportions of each parental gamete, and the ancestry of the gene copies at each locus. The variation between three states of ancestry on chromosomes of mixed descent is modelled as the outcome of three independent Poisson arrival processes. This requires only one extra parameter—the sum of the intensities of the arrival processes—to be specified in the model. Allele and haplotype frequencies are estimated by combining information from unadmixed and admixed population samples (using the posterior distribution of allele frequencies obtained from data on unadmixed individuals as a prior distribution for the corresponding ancestry specific allele frequencies in the admixed population). Where two or more SNPs in the same gene have been typed, these loci are grouped into a single “compound locus” and the program models the unobserved haplotypes, given the observed (phase unknown) genotypes at each compound locus. A generalised linear model is specified for the relation of the dependent variable (type 2 diabetes, insulin, or BMI) to individual admixture and other covariates such as age, sex, and socioeconomic variables. For type 2 diabetes, this is a logistic regression model. For insulin and BMI, this is a linear regression model. For fasting insulin, only those individuals who were classified as controls at baseline visit were included in the regression model.

The model is specified as a bayesian full probability model, in which all unobserved variables—such as haplotypes, ancestry states at each locus, gamete admixture proportions, and population level parameters—are “missing data”. Non-informative prior distributions are specified for the distribution of admixture proportions in the population, and for the parameters of the regression model. The posterior distribution of the missing data, given the observed data, is then generated by Markov chain Monte Carlo (MCMC) simulation. Inference about the parameters of the regression model is based on the posterior distribution. In large samples, the posterior means and 95% central posterior intervals (“95% credible intervals”) are asymptotically equivalent to maximum likelihood estimates and 95% confidence intervals (95% CI).

Score tests for allelic association with the trait, conditional on individual admixture and any other covariates, are constructed as described previously.<sup>10</sup> The parameter tested is the coefficient  $\beta$  for the effect of the allele or haplotype under study (coded as 0, 1, or 2 copies) in a regression model that includes admixture and other covariates such as age and sex. For each SNP, a positive score value indicates association of the trait with the allele being tested. To test the null hypothesis that  $\beta = 0$ , the score (gradient of the log-likelihood) and the observed information (curvature of the

log-likelihood) at  $\beta = 0$  are calculated by averaging over the posterior distribution of the missing data (the haplotypes and individual admixture values). The score test correctly allows for uncertainty about haplotype assignments and estimation of individual admixture proportions, because it is based on the likelihood of the observed data as a function of the parameter ( $\beta$ ) that is being tested.

The ratio of observed to complete information in the score test can be interpreted as a measure of the efficiency of the analysis, compared to a study design in which haplotypes have been observed directly and individual admixture proportions measured without error. Where an allele or haplotype is found to show significant association with the trait, it is possible to estimate the size of the effect of that allele by fitting a model in which the allele or haplotype (coded for each individual as 0, 1, or 2 copies) is included as an explanatory variable in the regression model. Inference is then based upon the posterior distribution of the regression coefficient. This, however, requires a separate run of the sampler for each hypothesis under test, whereas the score test procedure allows all loci and all haplotypes to be tested for association in a single run of the MCMC sampler. An approximation to the maximum likelihood estimate of the effect size (as the natural logarithm of the odds ratio, for a logistic regression model) can be obtained from the score test by dividing the score by the observed information. In large samples, this is asymptotically equivalent to computing the maximum likelihood estimate directly.

The fit of the observed genotype frequencies to Hardy-Weinberg proportions was estimated by a Fisher exact test. Linkage disequilibrium (LD) between markers was estimated using the 3LOCUS.PAS program, kindly provided by Dr Jeff Long. LD is expressed as the  $D'$  coefficient, in which the observed gametic disequilibrium ( $D$ ) is standardised by the theoretical maximum disequilibrium ( $D_{\max}$ ).<sup>14</sup>

## RESULTS

### Fit of the genotype frequencies to the Hardy-Weinberg proportions

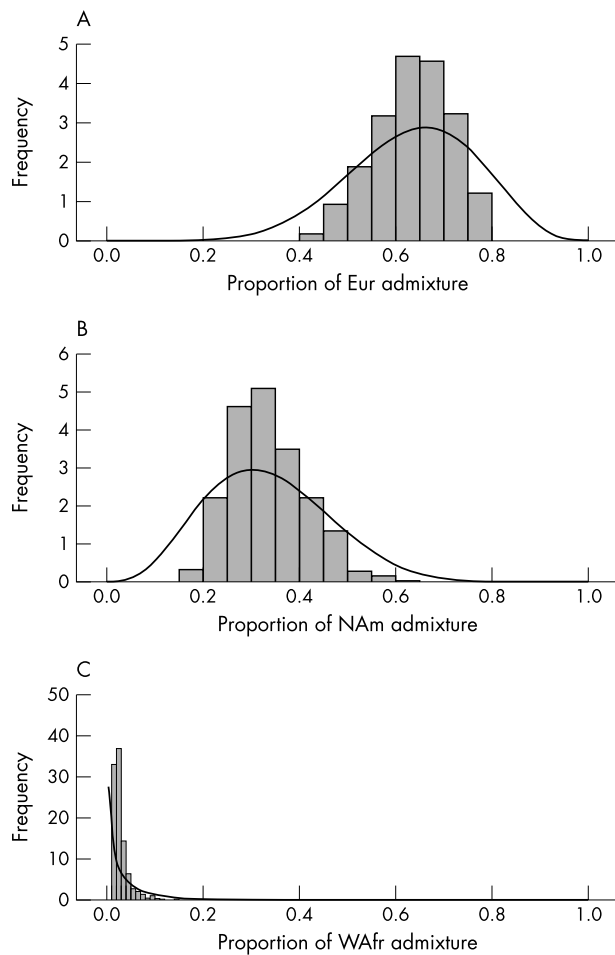
We tested independently in cases and controls if genotype frequencies deviate from the theoretical Hardy-Weinberg proportions (HW). We detected significant departures of HW in four markers in the sample of controls (FY-null,  $p = 0.01$ ; GNB3,  $p = 0.009$ ; MID-161,  $p = 0.02$ ; and MID-93,  $p = 0.015$ ). No significant deviations were observed in the sample of diabetics. Overall, the number of significant tests (four out of 64, or 6%) is very close to the expected proportion (5%).

### Associations with individual admixture

The mean admixture proportions of the population were estimated as 65% European, 34% Native American, and less than 1% African. The sum of intensities parameter was estimated as 7.1 (95% CI 4.9 to 12.4) per 100 cM, implying that the average time back to unadmixed ancestors in this population is at least seven generations. The average Native American ancestry is higher in the cases than in the controls (34.3 v 33.2%, respectively). The estimated distribution of individual admixture proportions for the total sample is shown in fig 1.

Table 4 shows the results of logistic regression analyses with type 2 diabetes as a dependent variable. In a model with age, sex, and individual admixture only, the odds ratio for type 2 diabetes associated with unit change in Native American admixture proportion (from 0 to 1) was estimated as 8.1 (95% CI 1.3 to 59). Adjustment for BMI had little effect on this odds ratio (table 4). When income and education were included as explanatory variables in the model, there was a strong inverse relation of type 2 diabetes risk to income: the odds ratio associated with an increase of one unit





**Figure 1** Distribution of individual admixture in the Hispanic American population of San Luis Valley. (A) European genetic contribution; (B) native American genetic contribution; (C) West African genetic contribution. Individual ancestry was estimated using the program ADMIXMAP, as described in the Statistical analysis section. Eur, European; Nam, North American; W Afr, West African.

in income category was 0.88 (95% CI 0.83 to 0.94). With income and education in the model, the odds ratio associated with Native American admixture was 5.5 (95% CI 0.8 to 41). With fasting insulin or BMI as dependent variable in the regression model, there was no evidence of a relationship of these variables to individual admixture (data not shown).

#### Linkage disequilibrium between markers within compound loci

We estimated the extent of linkage disequilibrium (LD) between markers within each compound locus. LD was very high between the four markers located in the CAP10 locus ( $D'$  between 85 and 100%) and also between the two markers located within the PPARG locus ( $D' = 88\%$ ). LD was also close to the maximum possible value between three of the four markers located within the ABCC8/KCNJ11 genes (ABCC8 exon 31 G→A, ABCC8 exon 33 G→T, and KCNJ11-E32K;  $D'$  between 94 and 100%). However, in spite of being located only some kilobases apart on chromosome 11 (approximately 30 kb), LD was very low between the ABCC8 exon 16 splice acceptor site and the other three markers within the ABCC8/KCNJ11 genes ( $D'$  between 9 and 19%). For this reason, in this genomic region we constructed haplotypes on the basis of the three markers showing tight linkage disequilibrium, and the exon 6 polymorphism was analysed independently.

**Table 4** Estimated odds ratios (posterior means, with 95% CI) in logistic regression models with diabetes as outcome variable

	Posterior mean	95% CI	97.5% Quantile
<b>Model 1</b>			
Age	1.033	1.024	1.044
Sex*	1.42	1.10	1.77
NAM†	8.1	1.3	58.6
<b>Model 2</b>			
Age	1.022	1.011	1.033
Sex*	1.31	1.04	1.65
Income	0.88	0.83	0.93
Education	0.96	0.92	1.00
NAM†	5.47	0.81	40.9
<b>Model 3</b>			
Age	1.040	1.029	1.051
Sex*	1.19	0.93	1.55
BMI	1.17	1.13	1.20
NAM†	7.85	1.19	58.0

\*Males = 1, females = 2; †Native American admixture proportion.

#### Associations with candidate gene polymorphisms

Table 5 shows the allele frequencies of the candidate gene polymorphisms analysed in the sample of cases and controls and table 6 depicts the results of tests for associations of type 2 diabetes with the SNPs in each compound locus, tested one at a time. Table 7 shows the results of score tests for associations of type 2 diabetes with the haplotypes estimated at each of the four compound loci. At each compound locus,

**Table 5** Allele frequency data for candidate gene polymorphisms in the sample of cases and controls from the San Luis Valley Hispanic sample

Locus	SLVDS cases		SLVDS controls	
	n	p	n	p
<b>CAPN10</b>				
CAPN10-43				
Allele G	175	0.766	255	0.733
Allele A		0.234		0.267
CAPN10-19				
3 Repeats	176	0.540	253	0.573
2 Repeats		0.460		0.427
CAPN10-63				
Allele T	177	0.175	250	0.156
Allele C		0.825		0.844
CAPN10-44				
Allele T	180	0.892	260	0.862
Allele C		0.108		0.138
<b>PPARG</b>				
PPARG-E6				
Allele T	181	0.870	247	0.913
Allele C		0.130		0.087
PPARG-Pro12Ala				
Allele Ala12	176	0.134	261	0.115
Allele Pro12		0.866		0.885
<b>ABCC8/KCNJ11</b>				
ABCC8-E16				
Allele T	182	0.533	261	0.492
Allele C		0.467		0.508
ABCC8-E31				
Allele A	176	0.392	242	0.314
Allele G		0.608		0.686
ABCC8-E33				
Allele T	182	0.690	261	0.617
Allele G		0.310		0.383
KCNJ11-E23K				
Allele A	177	0.302	250	0.382
Allele G		0.698		0.618
<b>GNB3</b>				
GNB3 C825T				
Allele T	179	0.385	250	0.361
Allele C		0.615		0.639

**Table 6** Tests for association of type 2 diabetes with SNPs within compound loci (alleles tested are indicated in parenthesis)

Locus	Adjusted for age, sex only					Adjusted for age, sex, admixture				
	Score	Obs Info	% Info	OR	p	Score	Obs Info	% Info	OR	p
CAPN10-43 (A)	-5.19	37.43	99.9	0.87	0.40	-4.28	35.20	96.7	0.89	0.47
CAPN10-19 (2 repeats)	4.95	47.09	99.9	1.11	0.47	4.09	44.42	96.9	1.10	0.54
CAPN10-63 (C)	-5.33	26.34	99.8	0.82	0.30	-3.36	23.81	94.6	0.87	0.49
CAPN10-44 (C)	-7.90	22.43	99.5	0.70	0.09	-7.59	20.88	96.1	0.70	0.10
PPARG-E6 (T)	9.16	18.31	100	1.65	0.03	7.26	15.69	90.3	1.59	0.07
PPARG-Pro12Ala (Pro12)	-4.94	22.08	99.9	0.80	0.29	-3.05	18.85	90.1	0.85	0.48
ABCC8-E16 (C)	-7.60	52.10	99.8	0.86	0.29	-4.95	47.66	95.4	0.90	0.47
ABCC8-E31 (G)	-14.35	42.98	99.9	0.72	0.03	-11.75	38.76	94.5	0.74	0.06
ABCC8-E33 (G)	-15.40	50.18	100	0.74	0.03	-13.97	46.85	96.7	0.74	0.04
KCNJ11-E23K (G)	16.12	48.37	100	1.40	0.02	14.50	44.94	96.4	1.38	0.03
GNB3 C825T (C)	-6.60	41.46	100	0.85	0.31	-6.31	39.69	98.0	0.85	0.32

OR, odds ratio.

rare haplotypes have been grouped into a single category to construct the test for association, although the statistical model evaluates all possible haplotypes. At each locus the haplotypes are tested for association one at a time, and in addition a chi-squared test statistic was calculated to test the null hypothesis that all haplotypes have no effect. Test statistics are not calculated for haplotypes where the observed information is less than 1 because for these rare haplotypes the asymptotic properties of the score test do not hold.

In a model adjusting for age and sex only, associations were significant at the 5% level for one of the two SNPs in the PPARG gene (PPARG-E6), and for three of the four SNPs in the ABCC8/KCNJ11 gene (ABCC8-E31, ABCC8-E33, and KCNJ11-E23K) (table 6). When each haplotype was tested successively, there were no significant associations with any of the four PPARG haplotypes. In the ABCC8/KCNJ11 gene, the A-T-G haplotype was positively associated with type 2 diabetes, and the G-G-A haplotype was negatively associated with type 2 diabetes. The summary test for association with all haplotypes at ABCC8/KCNJ11 was also significant ( $p=0.027$ ) (data not shown). None of the SNPs or haplotypes in CAPN10 showed any evidence of association with type 2 diabetes. In a model adjusting for individual

admixture proportions in addition to age and sex, the associations with SNPs in PPARG and ABCC8/KCNJ11 were changed only slightly (tables 6 and 7). Additional adjustment for BMI had little effect on these associations (data not shown). For two of the ancestry informative markers—MID52 and D11S429—associations with type 2 diabetes were significant at the 5% level without adjustment for admixture, but not after adjustment for individual admixture.

There were no significant associations of fasting insulin levels with SNPs or haplotypes in CAPN10, PPARG, or ABCC8/KCNJ11 (data not shown). The T allele at locus GNB3 was associated with higher insulin levels in an analysis adjusting for age and sex only ( $p=0.02$ ). This association persisted after adjustment for admixture ( $p=0.02$ ), but was weakened by adjustment for BMI (score  $-2.44$ , observed information 2.78,  $p=0.09$ ). None of the candidate gene polymorphisms showed any evidence of association with BMI.

## DISCUSSION

The prevalence of type 2 diabetes is higher in many populations of Native American ancestry than in people of European ancestry living in similar environments. In San Luis Valley, prevalence of type 2 diabetes in the admixed

**Table 7** Tests for association of type 2 diabetes with SNP alleles (or haplotypes at compound loci)

Locus	Haplotype	Adjusted for age, sex only					Adjusted for age, sex, admixture				
		Score	Obs Info	% Info	OR	p	Score	Obs Info	% Info	OR	p
<b>Candidate genes</b>											
<b>CAPN10 UCSNP 43-19-63-44</b>											
CAPN10	G-3-C-T	0.42	39.87	94	1.01	0.95	0.52	38.13	92	1.01	0.93
CAPN10	G-2-T-T	6.14	26.56	96	1.26	0.23	4.23	24.45	92	1.19	0.39
CAPN10	G-2-C-T	6.04	25.5	92	1.27	0.23	6.84	24.05	90	1.33	0.16
CAPN10	G-2-C-C	-6.80	21.26	97	0.73	0.14	-6.32	20.05	94	0.73	0.16
CAPN10	A-3-C-T	-5.03	36.9	96	0.87	0.41	-4.55	35.20	93	0.88	0.44
CAPN10	Others	-0.78	0.48				-0.71	0.46			
<b>PPARG E6-Pro12Ala</b>											
PPARG	C-Ala	1.25	5.493	93	1.26	0.59	0.73	5.57	85	1.14	0.76
PPARG	C-Pro	-8.78	24.44	99	0.70	0.08	-7.74	23.26	93	0.72	0.11
PPARG	T-Ala	5.20	17.38	97	1.35	0.21	4.34	14.90	90	1.34	0.26
PPARG	T-Pro	2.32	2.576	90	2.46	0.15	2.67	2.50	73	2.91	0.09
<b>ABCC8/KCNJ11 E31-E33-E23K</b>											
ABCC8/KCNJ11	A-T-G	13.96	44.06	96	1.37	0.03	11.28	39.25	90	1.33	0.07
ABCC8/KCNJ11	G-T-G	1.525	46.18	96	1.03	0.82	2.549	41.77	92	1.06	0.69
ABCC8/KCNJ11	G-G-A	-16.6	49.82	100	0.72	0.02	-14.87	45.83	95	0.72	0.02
ABCC8/KCNJ11	Others	1.116	0.4966				1.04	0.5004			
ABCC8-ex16		-7.60	52.10	99.8	0.86	0.29	-4.95	47.66	95.4	0.90	0.47
GNB3		-6.60	41.46	100	0.85	0.31	-6.31	39.69	98	0.85	0.32

For haplotypes where the observed information is less than 1, p values are omitted.

Hispanic American population compared with non-Hispanic American whites has been estimated to be 2.1-fold higher in men and 4.8-fold higher in women.<sup>11</sup> If the excess risk of type 2 diabetes in Hispanic Americans compared with Europeans has a genetic basis, we would expect to observe within Hispanic American populations an association of type 2 diabetes with the proportion of the genome that is of Native American ancestry. The estimated size of effect in this study (OR = 8.1 without adjustment for socioeconomic status, OR = 5.1 after adjustment for socioeconomic status) is compatible with a genetic explanation for this ethnic difference, but as the interval estimate for the adjusted odds ratio overlaps 1, we cannot exclude an environmental explanation based on factors associated with low socioeconomic status that confound the association with individual admixture. The 95% CIs for the effect of European/Native American admixture proportions on risk of type 2 diabetes are wide in this dataset: first, because individual admixture proportions vary only over a modest range in this population; and second, because with only 21 markers informative for ancestry we cannot accurately estimate the proportion of the genome that is of Native American ancestry.

Even though we cannot measure individual admixture accurately, we can still control for confounding by individual admixture when testing for associations with alleles or haplotypes after adjusting for individual admixture proportions, because the score test (by integrating over the posterior distribution of all missing data, including individual admixture) allows for uncertainty in estimates of individual admixture.<sup>10</sup> For most of the candidate gene polymorphisms that were included in this study, the confounding effect of individual admixture is weak, and thus the effect of the candidate gene polymorphism, adjusted for confounding, can be estimated accurately even though the confounder has not been measured accurately. The proportion of information extracted (ratio of observed to complete information) in the score test can be interpreted as a measure of the efficiency of the analysis, compared with the analysis of a dataset in which individual admixture is estimated accurately (with a large panel of ancestry informative markers) and individual haplotypes are assigned without error. Even though haplotypes are not observed directly in this analysis but inferred from unphased genotype data, the proportion of information extracted in score tests for association with each haplotype is greater than 70% except where the haplotype is rare. As theory shows, when studying haplotype effects on disease risk it is generally more efficient to study a large sample of unrelated individuals and to model the effects of the unobserved haplotypes than to type other family members in order to infer phase.<sup>15</sup>

In this study we typed polymorphisms in five candidate genes (CAPN10, GNB3, PPARG, ABCC8/KCNJ11) and evaluated their association with type 2 diabetes and fasting insulin. Because we have strong prior evidence of associations with these candidate gene polymorphisms, it is reasonable to interpret even modest significance levels ( $p < 0.05$ ) as evidence of association. The calpain-10 gene is of particular interest, because the association of this gene with type 2 diabetes was first detected in a Hispanic American population.<sup>16</sup> Horikawa *et al* reported that the highest diabetes risk in this population was associated with the haplotype pair 112/121 (here referred to as G2T/G3C) at the three polymorphic sites CAPN10-43, CAPN10-19, and CAPN10-63. Subsequent studies of these SNPs and others in the CAPN10 gene (for example, CAPN10-44) have not consistently replicated this result. While some studies found a modest effect on type 2 diabetes or associated phenotypes,<sup>17–20</sup> other reports indicated no association.<sup>21–24</sup> Several

factors can explain the heterogeneous findings observed in studies involving CAPN10 and type 2 diabetes: variation in statistical power between studies, ethnic differences in allele or haplotype frequencies, and the presence of gene–gene and gene–environment interactions, among others. Two recent meta-analyses of CAPN10 family based and population based studies have reported modest pooled odds ratios for the allele UCSNP-44 C (OR ~ 1.2<sup>25</sup>) and the UCSNP-43 G/G homozygote (OR ~ 1.2<sup>26</sup>). This means that very large sample sizes are required to detect the effect of these variants on type 2 diabetes risk. In the present study, we have not observed any significant association of CAPN10 alleles or haplotypes with type 2 diabetes, fasting insulin, or BMI (table 6). Additionally, no significant effect was detected when testing the G2T/G3C haplotype pair (data not shown). Although in our sample of Hispanic Americans the relative frequency of one of the haplotypes hypothesised to increase type 2 diabetes risk (112 or G2T) is higher than in populations of European ancestry (approximately 17 v 3–8%, respectively), we have not observed a significant effect of CAPN10 polymorphisms on type 2 diabetes or fasting insulin levels.

PPARG is a member of the nuclear hormone receptor subfamily of transcription factors, and has an important role in insulin action and fat metabolism.<sup>27</sup> A recent meta-analysis<sup>28</sup> including more than 3000 individuals of European ancestry indicated that the common Pro12 allele slightly increases type 2 diabetes risk. However, in a Native American population, the Oji-Cree from Western Ontario,<sup>29</sup> the Ala12 allele was associated with type 2 diabetes in women. In the San Luis Valley sample association was detected with C161T, a common silent polymorphism,<sup>30–31</sup> but not with the Pro12Ala locus. The analysis at the haplotype level did not show any significant effect.

The ABCC8 (SUR1) and KCNJ11 genes encode the two subunits that constitute the ATP sensitive potassium channel of the pancreatic beta cells. This channel is the target of the sulfonylurea class of drugs. Both genes are located 4.5 kb apart on chromosome 11, and have been widely studied in relation to risk of type 2 diabetes. We analysed three common polymorphisms in the ABCC8 gene and one site in the KCNJ11 gene in the San Luis Valley sample. The variants in the ABCC8 gene are a C/T mutation in the splice acceptor site of exon 16, which has been associated with effects on insulin secretion and type 2 diabetes,<sup>32–35</sup> a silent G/A mutation in exon 31 that has been associated with hyperinsulinaemia and type 2 diabetes,<sup>36–37</sup> and a non-synonymous G/T substitution in exon 33.<sup>38</sup> The G/A mutation in codon 23 (E23K) of the KCNJ11 gene has also been associated with type 2 diabetes in numerous studies, including two recent meta-analysis in European populations.<sup>39–42</sup> Three of these SNPs were associated with type 2 diabetes in the San Luis Valley sample (ABCC8 exon 31 allele A, ABCC8 exon 33 allele T, and KCNJ11 E23K allele G; see table 6). In fact, these three markers are in strong linkage disequilibrium (LD), with  $D'$  values between 94 and 100%, while the LD between these sites and the marker located upstream in exon 16 is very low ( $D'$  between 9 and 19%). Interestingly, the ABCC8 exon 31-A allele has been previously reported to be associated with hyperinsulinaemia in non-diabetic Mexican Americans.<sup>36</sup> Ancestry specific haplotype frequencies were estimated using ADMIXMAP, which combines the information from unadmixed “parental” samples and the SLV sample (table 8). The three common haplotypes in the compound ABCC8/KCNJ11 locus accounted for an estimated 98% of haplotypes in the San Luis Valley population. Haplotypes bearing A-T-G at the three loci listed above (Ex31-Ex33-E23K) were positively associated with type 2 diabetes, and haplotypes bearing G-G-A at the three loci listed above were inversely associated with type 2 diabetes. The haplotype increasing

**Table 8** ABCC8/KCNJ11 E31-E33-E23K ancestry specific haplotype frequencies estimated using ADMIXMAP

Haplotype	Native Americans	Europeans
A-T-A	0.0075	0.0002
A-T-G	0.4951	0.2785
A-G-A	0.0043	0.0050
A-G-G	0.0018	0.0002
G-T-A	0.0029	0.0002
G-T-G	0.1536	0.3549
G-G-A	0.3202	0.3606
G-G-G	0.0146	0.0004

type 2 diabetes risk (A-T-G) is more frequent in Native American populations than in European populations (50 v 28%). One problem with testing for haplotype effects is that without a strong prior hypothesis about which haplotypes are associated with disease risk, detection of effects relies on summary tests over all haplotypes, which have multiple degrees of freedom and thus low statistical power. It is important to note that the KCNJ11 E23K allele associated with type 2 diabetes in the San Luis Valley population (allele G) is not the same allele that has been reported to be associated with type 2 diabetes in populations of European ancestry (allele A).<sup>39-42</sup> In fact, the odds ratio associated with the presence of the E23K G allele in San Luis Valley (OR~1.40, CI: 1.05 to 1.85), does not overlap with the values reported in Europeans, where a recent meta-analysis of case control data indicated a pooled odds ratio associated with the presence of the E23K A allele of 1.23 (CI: 1.12 to 1.36<sup>40</sup>). Because the E23K A allele occurs predominantly on the same haplotype (G-G-A) in both Europeans and Native Americans, opposite directions of association with diabetes mellitus in these two populations cannot be explained by occurrence on different haplotypes. These contradictory results could be due in part to differences with respect to the ethnic background in the samples, but additional studies in Native American and Hispanic populations are needed to confirm this point.

The C825T polymorphism of the gene encoding the G protein beta-3 subunit (GNB3) has been repeatedly associated with obesity, hypertension, and type 2 diabetes.<sup>43-46</sup> This gene plays a key role in intracellular signalling and the 825T mutation creates a splice variant that is functional and is associated with enhanced G protein activation.<sup>46</sup> We observed no association with type 2 diabetes, but in non-diabetic individuals the T allele was associated with higher fasting insulin concentrations.

We have thus detected in this Hispanic American population an association of type 2 diabetes with markers located on the ABCC8 and KCNJ11 genes, closely linked on chromosome 11. We have also described evidence of association of the G825T polymorphism with fasting insulin in the non-diabetic sample. We did not observe an association of four CAPN10 markers with type 2 diabetes. In the score tests used in this study, the observed information for the log odds ratio associated with a common allele or haplotype is typically about 30, equivalent to a standard error of 0.18. We can thus estimate that our study had adequate (90%) statistical power to detect at 5% significance a log odds ratio for type 2 diabetes of about 0.6 (OR~1.8) associated with one extra copy of any common allele or haplotype. Therefore, our study would have detected effects of the magnitude described in the Mexican American sample in which the original association of CAPN10 and type 2 diabetes was reported (OR>2). However, as pointed by Song *et al.*,<sup>26</sup> neither this or, for that matter, most of the previous CAPN10 studies can individually achieve enough statistical power to detect the modest effects

that have been described in recent meta-analyses of markers such as the UCSNP-44 C allele (OR~1.19<sup>25</sup>), the UCSNP-43 G/G genotype (OR~1.19<sup>26</sup>), or the PPARG Pro12 allele (OR~1.25<sup>28</sup>).

We have also demonstrated the ability to control for confounding by population stratification when studying genetic associations within a recently admixed population, using a panel of markers informative for ancestry and Bayesian computationally intensive methods for statistical analysis. This makes it possible to study genetic associations in stratified populations using ordinary case control and cross sectional designs, rather than family based designs which require parents or sibs of affected individuals to be collected.<sup>10</sup> We note also that the San Luis Valley population, with about 34% average Native American admixture, is an ideal setting in which to apply a novel approach that exploits admixture to localise genes underlying ethnic differences in risk of type 2 diabetes. However this will require a much larger panel of markers informative for Native American versus European ancestry.

## ACKNOWLEDGEMENTS

We thank the San Luis Valley Study participants for their help.

## ELECTRONIC-DATABASE INFORMATION



The URLs mentioned in this paper are as follows: dbSNP, see NCBI website at <http://www.ncbi.nlm.nih.gov/SNP/index.html>; ADMIXMAP, available at <http://www.lshim.ac.uk/eu/genetics/index.html#admix.ls>

## Authors' affiliations

**E J Parra**, Department of Anthropology, University of Toronto at Mississauga, Mississauga, Canada ON L5L 1C6

**C J Hoggart**, **P M McKeigue**, Noncommunicable Disease Epidemiology Unit, London School of Hygiene and Tropical Medicine, London, WC1E 7HT, UK

**C Bonilla**, **S Dios**, **M D Shriver**, Department of Anthropology, Pennsylvania State University, University Park, PA 16802, USA

**J M Norris**, **J A Marshall**, **R F Hamman**, Department of Preventive Medicine and Biometrics, University of Colorado Health Sciences Center, Denver, CO 80262, USA

**R E Ferrell**, Department of Human Genetics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA, USA

This work was supported in part by grants from NIH/NIDDK (DK53958) and NIH/NHGRI (HG02154) to MDS. The development of the ADMIXMAP program was supported by NIH grant MH60343 to PMM.

Conflict of interest: none declared.

Correspondence to: Esteban J Parra, Department of Anthropology, University of Toronto at Mississauga, 3359 Mississauga Road North, Room 2002 South Building, Mississauga, Canada ON L5L 1C6; [eparra@utm.utoronto.ca](mailto:eparra@utm.utoronto.ca)

Revised version received 13 April 2004  
Accepted for publication 25 May 2004

## REFERENCES

- American Diabetes Association**, <http://www.diabetes.org>.
- Gardner Jr U**, Stern MP, Haffner SM, Gaskill SP, Hazuda HP, Relethford JH, Eifler CW. Prevalence of diabetes in Mexican Americans: relationship to percent of gene pool derived from Native American sources. *Diabetes* 1984;**33**:86-92.
- Chakraborty R**, Ferrell R-E, Stern M-P, Haffner S-M, Hazuda H-P, Rosenthal M. Relationship of prevalence of non-insulin-dependent diabetes mellitus to Amerindian admixture in the Mexican Americans of San Antonio, Texas. *Genet Epidemiol* 1986;**3**:435-54.
- Weiss K-M**. *Genetic variation and human disease: principles and evolutionary approaches*. Cambridge: Cambridge University Press, 1993.
- Williams R-C**, Long J-C, Hanson R, Sievers M-L, Knowler W-C. Individual estimates of European genetic admixture associated with lower risks of type 2



- diabetes and quantitative predictors of hyperglycemia and obesity in Pima Indians. *Am J Hum Genet* 2000;**66**:527–38.
- 6 **McKeigue P-M**, Carpenter J, Parra E-J, Shriver M-D. Estimation of admixture and detection of linkage in admixed populations by a Bayesian approach: applications to African-American populations. *Ann Hum Genet* 2000;**64**:171–86.
  - 7 **Parra EJ**, Kittles RA, Argyropoulos G, Pfaff CL, Hiester K, Bonilla C, Sylvester N, Parrish-Gause D, Garvey WT, Jin L, McKeigue PM, Kamboh MI, Ferrell RE, Pollitzer WS, Shriver MD. Ancestral proportions and admixture dynamics in geographically defined African Americans living in South Carolina. *Am J Phys Anthropol* 2001;**114**:18–29.
  - 8 **Pfaff CL**, Parra EJ, Bonilla C, Hiester K, McKeigue PM, Kamboh MI, Hutchinson RG, Ferrell RE, Boerwinkle E, Shriver MD. Population structure in admixed populations: effect of admixture dynamics on the pattern of linkage disequilibrium. *Am J Hum Genet* 2001;**68**:198–207.
  - 9 **Shriver MD**, Parra EJ, Dios S, Bonilla C, Norton H, Jovel C, Pfaff C, Jones C, Massac A, Cameron N, Baron A, Jackson T, Argyropoulos G, Jin L, Hoggart CJ, McKeigue PM, Kittles RA. Skin pigmentation, biogeographical ancestry and admixture mapping. *Hum Genet* 2003;**112**:387–99.
  - 10 **Hoggart C-J**, Parra E-J, Shriver M-D, Kittles R-A, Clayton D-G, McKeigue P-M. Control of confounding of genetic associations in stratified populations. *Am J Hum Genet* 2003;**72**:1492–504.
  - 11 **Hamman RF**, Marshall JA, Baxter J, Kahn LB, Mayer EJ, Orleans M, Murphy JR, Lezotte DC. Methods and prevalence of non-insulin-dependent diabetes mellitus in a biethnic Colorado population: the San Luis Valley diabetes study. *Am J Epidemiol* 1989;**129**:295–311.
  - 12 **Akey J-M**, Sosnoski D, Parra E, Dios S, Hiester K, Su B, Bonilla C, Jin L, Shriver MD. Melting curve analysis of SNPs (McSNP): a gel-free and inexpensive approach for SNP genotyping. *BioTechniques* 2001;**30**:358–67.
  - 13 **Bonilla C**, Parra EJ, Pfaff CL, Dios S, Marshall JA, Hamman RF, Ferrell RE, Hoggart CL, McKeigue PM, Shriver MD. Admixture in the Hispanics of the San Luis Valley, Colorado, and its implications for complex trait gene mapping. *Ann Hum Genet* 2004;**68**:139–53.
  - 14 **Thompson EA**, Deeb S, Walker D, Motulsky AG. The detection of linkage disequilibrium between closely linked markers: RFLP at the Al-CIII apolipoprotein genes. *Am J Hum Genet* 1988;**42**:113–24.
  - 15 **McKeigue PM**. Efficiency of estimation of haplotype frequencies: use of marker phenotypes of unrelated individuals versus counting of phase-known gametes. *Am J Hum Genet* 2000;**67**:1626–7.
  - 16 **Horikawa Y**, Oda N, Cox NJ, Li X, Orho-Melander M, Hara M, Hinokio Y, Lindner TH, Mashima H, Schwarz PE, del Bosque-Plata L, Horikawa Y, Oda Y, Yoshiuchi I, Colilla S, Polonsky KS, Wei S, Concannon P, Iwasaki N, Schulze J, Baier LJ, Bogardus C, Groop L, Boerwinkle E, Hanis CL, Bell GI. Genetic variation in the gene encoding calpain-10 is associated with type 2 diabetes mellitus. *Nat Genet* 2000;**26**:163–75.
  - 17 **Baier LJ**, Permana PA, Yang X, Pratlley RE, Hanson RL, Shen GQ, Mott D, Knowler WC, Cox NJ, Horikawa Y, Oda N, Bell GI, Bogardus C. A calpain-10 gene polymorphism is associated with reduced muscle mRNA levels and insulin resistance. *J Clin Invest* 2000;**106**:819–21.
  - 18 **Elbein SC**, Chu W, Ren Q, Hemphill C, Schay J, Cox NJ, Hanis CL, Hasstedt SJ. Role of calpain-10 gene variants in familial type 2 diabetes in Caucasians. *J Clin Endocrinol Metab* 2002;**87**:650–4.
  - 19 **Orho-Melander M**, Klannermark M, Svensson M-K, Ridderstrale M, Lindgren C-M, Groop L. Variants in the calpain-10 gene predispose to insulin resistance and elevated free fatty acid levels. *Diabetes* 2002;**51**:2658–64.
  - 20 **Cassell PG**, Jackson AE, North BV, Evans JC, Syndercombe-Court D, Phillips C, Ramachandran A, Snehalaitha C, Gelding SV, Vijayaravaghan S, Curtis D, Hitman GA. Haplotype combinations of calpain gene polymorphisms associate with increased risk of impaired glucose tolerance and type 2 diabetes in South Indians. *Diabetes* 2002;**51**:1622–8.
  - 21 **Hegele RA**, Harris SB, Zinman B, Hanley AJ, Cao H. Absence of association of type 2 diabetes with CAPN10 and PC-1 polymorphisms in Oji-Cree. *Diabetes Care* 2001;**24**:1498–9.
  - 22 **Tsai HJ**, Sun G, Weeks DE, Kaushal R, Wolujewicz M, McGarvey ST, Tufa J, Viali S, Deka R. Type 2 diabetes and three calpain-10 gene polymorphisms in Samoans: no evidence of association. *Am J Hum Genet* 2001;**69**:1236–44.
  - 23 **Fingerlin TE**, Erdos MR, Watanabe RM, Wiles KR, Stringham HM, Mohlke KL, Silander K, Valle TT, Buchanan TA, Tuomilehto J, Bergman RN, Boehnke M, Collins FS. Variation in three single nucleotide polymorphisms in the calpain-10 gene not associated with type 2 diabetes in a large Finish cohort. *Diabetes* 2002;**51**:1644–8.
  - 24 **Horikawa Y**, Oda N, Yu L, Imamura S, Fujiwara K, Makino M, Seino Y, Itoh M, Takeda J. Genetic variations in calpain-10 gene are not a major factor in the occurrence of type 2 diabetes in Japanese. *J Clin Endocrinol Metab* 2003;**88**:244–7.
  - 25 **Weedon MN**, Schwarz PE, Horikawa Y, Iwasaki N, Illig T, Holle R, Rathmann W, Selisko T, Schulze J, Owen KR, Evans J, Del Bosque-Plata L, Hitman G, Walker M, Levy JC, Sampson M, Bell GI, McCarthy MI, Hattersley AT, Frayling TM. Meta-analysis and a large association study confirm a role for calpain-10 variation in type 2 diabetes susceptibility. *Am J Hum Genet* 2003;**73**:1208–12.
  - 26 **Song Y**, Niu T, Manson JE, Kwiatkowski DJ, Liu S. Are variants in the CAPN10 gene related to risk of type 2 diabetes? A quantitative assessment of population and family-based association studies. *Am J Hum Genet* 2004;**74**:208–22.
  - 27 **Schwartz M-W**, Kahn S-E. Insulin resistance and obesity. *Nature* 1999;**402**:860–1.
  - 28 **Altshuler D**, Hirschhorn JN, Klannermark M, Lindgren CM, Vohl MC, Nemesh J, Lane CR, Schaffner SF, Bolk S, Brewer C, Tuomi T, Gudut D, Hudson TJ, Daly M, Groop L, Lander ES. The common PPAR $\gamma$  Pro12Ala polymorphism is associated with decreased risk of type 2 diabetes. *Nat Genet* 2000;**26**:76–9.
  - 29 **Hegele R-A**, Cao H, Harris S-B, Zinman B, Hanley A-J, Anderson C-M. Peroxisome proliferator-activated receptor-gamma2 P12A and type 2 diabetes in Canadian Oji-Cree. *J Clin Endocrinol Metab* 2000;**85**:2014–9.
  - 30 **Meirhaeghe A**, Fajas L, Helbecque N, Cotel D, Lebel P, Dallongeville J, Deeb S, Auwerx J, Amouyel P. A genetic polymorphism of the peroxisome proliferator-activated receptor  $\gamma$  gene influences plasma leptin levels in obese humans. *Hum Mol Genet* 1998;**7**:435–40.
  - 31 **Koch M**, Rett K, Maerker E, Volk A, Haist K, Deninger M, Rettig A, Renn W, Haring HU. The silent PPARgamma exon 6 CAC(His) $\rightarrow$ CAT(His) polymorphism does not affect the plasma leptin levels in a collective of first degree relatives of type 2 diabetes patients from South West Germany. *Exp Clin Endocrinol Diabetes* 2000;**108**:341–6.
  - 32 **Inoue H**, Ferrer J, Welling CM, Elbein SC, Hoffman M, Mayorga R, Warren-Perry M, Zhang Y, Mills H, Turner R, Province M, Bryan J, Permutt MA, Aguilar-Bryan L. Sequence variants in the sulfonylurea receptor (SUR) gene are associated with NIDDM in Caucasians. *Diabetes* 1996;**45**:825–31.
  - 33 **Hart LM**, de Krijff P, Dekker JM, Stolk RP, Niipels G, van der Does FE, Ruige JB, Grobbee DE, Heine RJ, Maassen JA. Variants in the sulfonylurea receptor gene: association of the exon 16-3t variant with type II diabetes mellitus in Dutch Caucasians. *Diabetologia* 1999;**42**:617–20.
  - 34 **Hart LM**, Dekker JM, van Haefen TW, Ruige JB, Stehouwer CD, Erkelens DW, Heine RJ, Maassen JA. Reduced second phase insulin secretions in carriers of a sulfonylurea receptor gene variant associating with type 2 diabetes mellitus. *Diabetologia* 2000;**43**:515–9.
  - 35 **Meirhaeghe A**, Helbecque N, Cotel D, Arveiler D, Ruidavets JB, Haas B, Ferrieres J, Tauber JP, Bingham A, Amouyel P. Impact of sulfonylurea receptor 1 genetic variability on non-insulin-dependent diabetes mellitus prevalence and treatment: a population study. *Am J Med Genet* 2001;**101**:4–8.
  - 36 **Goksel D-L**, Fischbach K, Duggirala R, Mitchell BD, Aguilar-Bryan L, Blangero J, Stern MP, O'Connell P. Variant in sulfonylurea receptor-1 gene is associated with high insulin concentrations in non-diabetic Mexican Americans: SUR-1 gene variant and hyperinsulinemia. *Hum Genet* 1998;**103**:280–5.
  - 37 **Reis A-F**, Ye W-Z, Dubois-Laforgue D, Bellanne-Chantelot C, Timsit J, Velho G. Association of a variant in exon 31 of the sulfonylurea receptor 1 (SUR1) gene with type 2 diabetes mellitus in French Caucasians. *Hum Genet* 2000;**107**:138–44.
  - 38 **Rissanen J**, Markkanen A, Karkkainen P, Pihlajamaki J, Kekalainen P, Mykkanen L, Kuusisto J, Karhapaa P, Niskanen L, Laakso M. Sulfonylurea receptor 1 gene variants are associated with gestational diabetes and type 2 diabetes but not with altered secretion of insulin. *Diabetes Care* 2000;**23**:70–3.
  - 39 **Nielsen EM**, Hansen L, Carstensen B, Echwald SM, Drivsholm T, Glumer C, Thorsteinsson B, Borch-Johnsen K, Hansen T, Pedersen O. The E23K variant of Kir6.2 associates with impaired post-OGTT serum insulin response and increased risk of type 2 diabetes. *Diabetes* 2003;**52**:573–7.
  - 40 **Gloyd AL**, Weedon MN, Owen KR, Turner MJ, Knight BA, Hitman G, Walker M, Levy JC, Sampson M, Halford S, McCarthy MI, Hattersley AT, Frayling TM. Large-scale association studies of variants in genes encoding the pancreatic beta-cell KATP channel subunits Kir6.2 (KCNJ11) and SUR1 (ABCC8) confirm that the KCNJ11 E23K variant is associated with type 2 diabetes. *Diabetes* 2003;**52**:568–72.
  - 41 **Love-Gregory L**, Wasson J, Lin J, Skolnick G, Suarez B, Permutt M-A. E23K single nucleotide polymorphism in the islet ATP-sensitive potassium channel gene (Kir6.2) contributes as much to the risk of type II diabetes in Caucasians as the PPARgamma Pro12Ala variant. *Diabetologia* 2003;**46**:136–7.
  - 42 **Barroso I**, Luan J, Middleberg RP, Harding AH, Frans PW, Jakes RW, Clayton D, Schafer AJ, O'Rahilly S, Wareham NJ. Candidate gene association study in type 2 diabetes indicates a role for genes involved in  $\beta$ -cell function as well as insulin action. *PLoS Biol* 2003;**1**:41–55.
  - 43 **Siffert W**, Rosskopf D, Siffert G, Busch S, Moritz A, Erbel R, Sharma AM, Ritz E, Wichmann HE, Jakobs KH, Horsthemke B. Association of a human G-protein  $\beta$ 3 subunit variant with hypertension. *Nat Genet* 1998;**18**:45–8.
  - 44 **Siffert W**, Forster P, Jockel KH, Mvere DA, Brinkmann B, Naber C, Crookes R, Du P Heyns A, Epplen JT, Fridley J, Freedman BI, Muller N, Stolke D, Sharma AM, Al Moutaery K, Grosse-Wilde H, Buerbaum B, Ehrlich T, Ahmad NR, Horsthemke B, Du Toit ED, Tiilikainen A, Ge J, Wang Y, Yang D, Husing J, Rosskopf D. Worldwide ethnic distribution of the G protein beta3 subunit 825T allele and its association with obesity in Caucasian, Chinese, and Black African individuals. *J Am Soc Nephrol* 1999;**10**:1921–30.
  - 45 **Roskopf D**, Frey U, Eckhardt S, Schmidt S, Ritz E, Hofmann S, Jaksch M, Muller N, Husing J, Siffert W, Jocke KH. Interaction of the G protein beta 3 subunit T825 allele and the IRS-1 Arg972 variant in type 2 diabetes. *Eur J Med Res* 2000;**5**:484–90.
  - 46 **Fernandez-Real JM**, Penarroja G, Richart C, Castro A, Vendrell J, Broch M, Lopez-Bermejo A, Ricart W. G protein beta3 gene variant, vascular function, and insulin sensitivity in type 2 diabetes. *Hypertension* 2003;**41**:124–9.