Am. J. Hum. Genet. 69:1236–1244, 2001

Type 2 Diabetes and Three *Calpain-10* Gene Polymorphisms in Samoans: No Evidence of Association

Hui-Ju Tsai,[1] Guangyun Sun,[2] Daniel E. Weeks,[1] Ritesh Kaushal,[2] Michael Wolujewicz,[2] Stephen T. McGarvey,[3] Joseph Tufa,[4] Satupaitea Viali,[5] and Ranjan Deka[2]

[1] Department of Human Genetics, University of Pittsburgh, Pittsburgh; [2]Department of Environmental Health, University of Cincinnati, Cincinnati; [3]International Health Institute, Brown University, Providence, RI; [4]Department of Health, American Samoa Government, Pago Pago, American Samoa; and [5]Department of Health, Government of Samoa, Apia, Samoa

Although genomewide scans have identified several potential chromosomal susceptibility regions in several human populations, finding a causative gene for type 2 diabetes has remained elusive. Others have reported a novel gene, *calpain-10* (*CAPN10*), located in a previously identified region on chromosome 2q37.3, as a putative susceptibility gene for type 2 diabetes. Three single-nucleotide polymorphisms (SNPs) (UCSNP43, UCSNP19, and UCSNP63) were shown to be involved in increased risk of the disease among Mexican Americans. We have tested the association of these three SNPs with type 2 diabetes among the Samoans of Polynesia, who have a very high prevalence of the disease. In the U.S. territory of American Samoa, prevalence is 25% and 15% in men and women, respectively, whereas, in the independent nation of Samoa, prevalence is 3% and 5% in men and women, respectively. In our study sample, which consisted of 172 unrelated affected case subjects and 96 control subjects, we failed to detect any association between case subjects and control subjects in allele frequencies, haplotype frequencies, or haplotype combinations of UCSNP43, -19, and -63. Also, our data showed no evidence of linkage, among 201 affected sib pairs, in the region of chromosome 2 that contains these SNPs. Three plausible scenarios could explain these observations. (1) *CAPN10* is a susceptibility gene only in particular ethnic groups; (2) our study lacks power to detect the effects of *CAPN10* polymorphisms (but our sample size is comparable to that of earlier reports); or (3) the underlying biological mechanism is too complex and requires further research.

Introduction

Non–insulin-dependent diabetes mellitus (NIDDM), also called "type 2 diabetes" (MIM 125853), is a common disease that affects ~135 million people on a global scale (King et al. 1998). Substantial evidence has been presented that type 2 diabetes has a complex genetic etiology, and, in recent years, genomewide scans have identified several potential chromosomal susceptibility regions (e.g., Hanis et al. 1996; Mahtani et al. 1996; Hanson et al. 1998; Elbein et al. 1999; Ghosh et al. 1999; Hegele et al. 1999; Ehm et al. 2000; Watanabe et al. 2000; Permutt et al. 2001). Nevertheless, finding a causative gene has remained elusive. Hanis et al. (1996) reported linkage to a region on chromosome 2q37.3 among Mexican Americans and identified a major susceptibility locus (*NIDDM1* [MIM 601283]), located in the interval that spans markers D2S125–D2S140. Following a combined strategy of positional cloning and a newly developed statistical method of partitioning linkage, these investigators identified a novel gene, *calpain-10* (*CAPN10* [MIM 605286]), as a putative type 2 diabetes susceptibility gene in this region (Horikawa et al. 2000). Three single-nucleotide polymorphisms (SNPs) (UCSNP43, UCSNP19, and UCSNP63), all located in intronic sequences, were found to be involved in increased risk of the disease. Although the common G allele at SNP43 (henceforth, the SNPs are abbreviated without the prefix "UC") was initially found to be significantly associated with the phenotype in families that showed linkage to *NIDDM1* on chromosome 2, haplotype 112/121 (defined by the three SNP sites) had shown the highest risk, an almost threefold risk of developing diabetes.

In the present study, we tested the association of the three polymorphisms (SNP43, 19, and 63) at the *CAPN10* gene with type 2 diabetes among the Samoans of the western Pacific. The Samoan participants are distributed over two polities, the independent nation of Samoa (formerly Western Samoa) and the U.S. territory of American Samoa. Samoa is a relatively underdeveloped country with a largely agricultural economy. Approximately 15% of the adults are engaged in wage earning, and the remainder are involved in local subsistence farming and fishing. On the other hand, American Samoa has received substantial financial support from the U.S. federal government since the 1950s; the

Received June 21, 2001; accepted for publication September 28, 2001; electronically published November 5, 2001.

Address for correspondence and reprints: Dr. Ranjan Deka, Department of Environmental Health, University of Cincinnati, 3223 Eden Avenue, Cincinnati, OH 45267-0056. E-mail: ranjan.deka@uc.edu

0002-9297/2001/6906-0009$02.00

economy is far less agricultural than that of Samoa, with ~60% of males and ~42% of women in the labor force (U.S. Department of Commerce 1990). In spite of great economic disparity between the two polities and exposure to different levels of the processes of modernization, they form one sociocultural unit, with the frequent exchange of mates and the maintenance of extended familial relationships across the border. We also note that our earlier population genetic studies demonstrated no genetic difference between the two locales (Deka et al. 1994). The contemporary adult Samoan population has a very high prevalence of overweight, obesity, and type 2 diabetes (McGarvey and Baker 1979; Zimmet 1979; McGarvey 1991, 1993; Collins et al. 1994; Hodge et al. 1994). The prevalence and four-year incidence rates of type 2 diabetes (defined as fasting glucose levels >126 mg/dl) are significantly higher in American Samoan men and women 29–62 years of age than in their Samoan counterparts. In American Samoa, prevalence is 25% and 15% in men and women, respectively, compared with 3.3% and 5.4% in men and women, respectively, from Samoa. Four-year incidence, measured from 1990–1991 to 1994–1995, is 12.4% and 8.5% in American Samoan men and women, respectively, compared with 2% and 2.6% in men and women, respectively, from Samoa (McGarvey et al. 2000; S. T. McGarvey, unpublished data). It is reasonable to conceive that the differences in diabetes prevalence in Samoa and American Samoa largely reflect an interaction between genetic susceptibility and environmental and lifestyle factors—in this instance, differential exposure to the processes of modernization.

Methods

Subjects

Affected subjects.—A sample of 299 affected sib pairs and 1 half-sib pair from 371 individuals in 176 families was recruited in American Samoa, using the Department of Health diabetes registry. The diagnosis of type 2 diabetes was confirmed by medical records. It should be noted that the registry also includes case subjects who were considered to have type 1 diabetes and were therefore excluded from the present study. Mean age of the 371 case subjects was 55.0 ± 11.5 years, with a range of 30–80 years. Age at onset of type 2 diabetes was calculated on the basis of the individual's self-report of the number of years since diagnosis. The mean age at onset of type 2 diabetes was 47.6 ± 10.7 years, with a range of 30–75 years. Extensive medical histories were collected, with a focus on information about how and when the subjects' type 2 diabetes was diagnosed. The interview included items about diabetes medication and use of injected insulin. We selected the first affected sib identified in each family to form the unrelated case subjects ($N = 172$). The first affected sib in each family was our proband from the diabetes registry through whom the family initially came to our attention. Subjects were asked about their Samoan ancestry, to limit study participation to those who reported that all four grand parents were of Samoan ethnicity without European or Asian ancestry. Study protocols were approved by the institutional review board of the Miriam Hospital, Providence, RI. Written informed consent was obtained from all participants.

Control subjects.—We established a control sample of 96 individuals who did not have type 2 diabetes; 40 were recruited from American Samoa, and 56 were recruited from Samoa. These samples were derived from our earlier longitudinal studies of adiposity and risk factors for cardiovascular disease in American Samoa and Samoa in 1990–1995 (McGarvey et al. 1993; Galanis et al. 1999). The control subjects were ⩾50 years of age and had fasting glucose levels of <120 mg/dl and body mass indexes <30 kg/m^2. These two measures were below these levels at both of two consecutive time points, four years apart, of data collection in our earlier longitudinal studies.

Genotyping

Genomic DNA was extracted from buffy coats, using Puregene DNA isolation kit (Gentra Systems), according to standard protocols. Primer sequences and protocols for PCR amplification and analysis of SNP43, -19, and -63 were kindly provided by Drs. Nancy J. Cox and Craig L. Hanis. In brief, 10–20 ng of genomic DNA were amplified in a total volume of 10 μl of reaction mixture containing standard PCR buffer, 10 pmol of each primer, 200 μm of each dNTP, and 0.5 U of *Taq* DNA polymerase. For SNP43 and -63, the amplified products were digested with restriction enzymes *Nsi*I and *Hha*I, respectively. The products were separated by electrophoresis on 2% agarose gel.

The genome scan was conducted using the ABI 10-cM panel of microsatellite markers, according to standard protocols, on an ABI 377 automated DNA sequencer (Applied Biosystems). A negative control sample and two positive control samples of known genotype (CEPH sample 1347-02) were tested, as a quality control, on each gel. Genescan 672 software was used to collect the data, to track lanes, to estimate fragment sizes, and to check internal size standards. Genotypes were assigned using the Genotyper software.

Statistical Analyses

Relationship testing and error checking.—We used the programs Relpair (version 0.9) and Prest to check the accuracy of relationships between individuals within families without removing genotyping errors from the

original data (Boehnke and Cox 1997; McPeek and Sun 2000). These analyses were based on 125 markers typed throughout the genome. Relpair implements a likelihood-based method to infer the most likely relationship for each pair. It applies a hidden-Markov–model algorithm to calculate the identity-by-descent (IBD) distribution of each pair. Then, under the assumption of no genetic interference, it estimates the multipoint likelihood of the marker data for each pair and infers the most likely relationship between each pair. Because Relpair estimates the likelihood conditional only on each of four possible relationships (full sibs, monozygotic twins, half sibs, and unrelated pairs), we then checked the more distant relationships, such as avuncular and first-cousin relationships, by using the program Prest. Instead of using a hidden-Markov method, Prest uses an augmented IBD Markov process to calculate the likelihood of the marker data. There are two new test statistics, conditional expected IBD (EIBD) and adjusted identity-by-state (AIBS), in Prest. We initially screened the data by use of these two statistics and then performed the maximum-likelihood–ratio test (MLRT) in Prest to infer the most likely relationship for each pair. The program PedCheck was used to identify the genotyping errors in the data (O'Connell and Weeks 1998).

Allele frequencies.—Allele frequencies for each SNP in the control subjects and the unrelated case subjects were computed by gene counting. We then tested, by use of a 2 × 2 contingency χ^2 test, whether there was a significant difference in allele frequencies between the control group recruited from American Samoa and that from Samoa, as well as between the control subjects and the unrelated case subjects. Conformity of genotype proportions to Hardy-Weinberg expectations was tested by using the program Gen (Lazzeroni and Lange 1998).

Haplotype estimation.—Haplotype frequencies were estimated by the programs EH, Zaplo, and Phase (Xie and Ott 1993; Terwilliger and Ott 1994; O'Connell 2000; Stephens et al. 2001). All three programs implement the expectation-maximization (EM) algorithm to estimate haplotype frequencies. The EH program uses the method of gene counting, which provides maximum-likelihood estimates of the haplotype frequencies. Both EH and Phase are applicable only to unrelated individuals. In contrast, Zaplo uses a pedigree-based likelihood method to estimate haplotype frequencies and can therefore be applied to the intact family structures, using all individuals in the family. Under the assumption of no recombination between the markers, the initial step is to compute the likelihood of the marker data by assigning equal allele frequencies for each haplotype. Zaplo then computes the expected haplotype frequencies by using the probability of each genotype to weight the counts, and the program updates the likelihood by use of the expected values. It then iterates until convergence. Because our unrelated case subjects are drawn from families that contain multiple affected siblings, it is important to use Zaplo to verify that the haplotype frequencies obtained using all members of the families are very similar to the frequencies obtained when one affected case subject was selected from each family. This is especially relevant, since our set of families contains more-precise information about haplotype frequencies than does our subset of singleton unrelated case subjects.

Test of linkage disequilibrium.—Linkage disequilibrium (LD) was evaluated by the program GOLD for each pairwise comparison of the SNPs in the case subjects and the control subjects (Abecasis and Cookson 2000). We calculated the standardized disequilibrium coefficient D' and the Δ^2 measure of disequilibrium, as described for each pair of the SNPs. The association of each pairwise comparison was tested by Fisher's exact test (Hedrick 1987; Devlin and Risch 1995).

Association analyses.—We used a nonparametric T5 statistic, which is implemented in the EH program, to test whether haplotype frequencies were significantly different between the familial case subjects and the control subjects (Zhao et al. 2000). To compute T5, we ran EH three times—on the case subjects alone, on the control subjects alone, and on the combination of case subjects and controls—to obtain the following three log-likelihoods: $\ln L_{case}$, $\ln L_{control}$, and $\ln L_{combined}$. Under the hypothesis that the allelic association was allowed, we computed the T5 statistic, which is defined as $2[\ln(L_{case}) + \ln(L_{control}) - \ln(L_{combined})]$, which has an approximate χ^2 distribution, to test the association of haplotype frequencies between the case subjects and the control subjects.

Besides testing for association on the basis of haplotype frequencies, we also examined the association between specific haplotype combinations and type 2 diabetes. We used the program Phase, which deploys a Bayesian approach, to estimate the most likely haplotypes of the three SNPs in the case subjects and the control subjects separately. In the first step, Phase estimates the haplotype frequencies, using the EM algorithm from 100 different starting points, and then chooses the estimate of haplotype frequencies with the highest likelihood. Next, it uses these haplotype frequencies as a priori expectations, performs Gibbs sampling to obtain an approximate sample from the posterior distribution of haplotypes given genotypes, and then chooses the most probable haplotype assignment for each individual.

Linkage analyses.—Nonparametric linkage analyses of 30 microsatellite markers (ABI panels 3 and 4) spanning chromosome 2 were performed using the program Allegro (Gudbjartsson et al. 2000). We calculated multipoint S_{all} LOD scores under the linear model.

Table 1

Estimated Allele Frequencies for SNP43, -19, and -63 in the 40 American Samoa Control Subjects and the 56 Samoan Control Subjects

SNP[a]	Allele 1 Frequency (No./Total) in[b]		χ^2	*P*
	AS Controls	S Controls		
43	.897 (70/78)	.920 (103/112)	.073	.788
19	.375 (30/80)	.366 (41/112)	.001	.980
63	.821 (64/78)	.848 (95/112)	.095	.758

[a] We used the allele labeling scheme of Horikawa et al. (2000), in which allele 1 at SNP43 is the G allele, allele 1 at SNP19 is the two repeats of the 32-bp sequence, and allele 1 at SNP63 is the C allele.
[b] AS = American Samoan; S = Samoan.

Results

Our relationship testing, which was based on 125 markers typed throughout the genome, revealed that the specified relationships of the members within some families were inconsistent with the relationships inferred from the marker data. Thus, we omitted all the pedigrees with unresolved structures and generated a conservative data set consisting of 139 families containing 201 affected sib pairs and 1 half-sib pair. This data set was used for subsequent analyses that depend on pedigree structure (i.e., Zaplo and the multipoint LOD scores). However, the majority of the association analyses presented here are not affected by uncertain pedigree structures, since they are based only on the proband from each family. In this context, it should be noted that there are no differences in sociodemographic characteristics or diabetes status between the group of probands identified from the diabetes registry and the group of probands' siblings. For example, the ratio of women to men among the probands is 50:50, and that among the siblings is 45:55. Among women, the mean age of the proband is 53.4 years and that of other siblings is 53.3 years. Among men, the mean age of the proband is 57.5 years and that of other siblings is 55.7 years. The average length of time since diagnosis among women is 7.0 years and 6.5 years for probands and other siblings, respectively; and, among men, average length of time since diagnosis is 8.3 years and 7.7 years for probands and other siblings, respectively. More than 85% of the probands currently use diabetic medication, compared with 83% of the other siblings. There is no apparent bias in sex, age, or diabetes severity in choosing the proband (first sibling recruited) in each family for the genetic analyses.

Allele Frequencies

Tests for genetic homogeneity in the allele frequencies of SNP43, -19, and -63 between the American Samoan and Samoan control subjects showed no significant difference in the allele frequencies at each of the three polymorphic sites between these two control groups (table 1). Our earlier population genetic study also had demonstrated that there is no difference in allele frequencies at nine hypervariable and four serum protein loci between these two geographical groups of Samoans (Deka et al. 1994). Hence, we combined the control subjects from both places to form one control group, which was used in the further analyses. We then compared the allele frequencies between the familial case subjects (the probands) and the control subjects. The results showed that there was no significant difference in the allele frequencies of SNP43, -19, and -63 between the familial case subjects and the control subjects (table 2). Also, the genotype frequencies of all the SNPs that we used were in Hardy-Weinberg equilibrium (data not shown).

Analyses of Association

We were interested in testing whether the haplotype frequencies of SNP43, -19, and -63 were different between the case subjects and the control subjects. Table 3 presents the haplotype frequencies in 151 case subjects, 94 control subjects, and all the siblings for whom data at all of the three polymorphisms are available. Six of the eight possible haplotypes were observed in the total data set. The three most common haplotypes were 121, 111, and 112 (allele designations are shown in table 1). The haplotype frequencies estimated by each program (EH, Zaplo, and Phase) were quite consistent (table 3). Also, the haplotype frequencies estimated on the basis of the family data were very close to the results estimated on the basis of the unrelated subjects. We then tested whether the haplotype frequencies of SNP43, -19, and -63 were significantly different between the familial case subjects and the control subjects. The results indicated that there was no evidence of association between the familial case subjects and the control subjects ($P = .858$) (table 4).

Table 2

Estimated Allele Frequencies for SNP43, -19, and -63 in the 172 Case Subjects (Typed at One or More SNPs) and the 96 Control Subjects

SNP	Allele 1 Frequency (No./Total) in		χ^2	*P*
	Cases	Controls		
43	.914 (309/338)	.911 (173/190)	.0003	.986
19	.326 (107/328)	.370 (71/192)	.837	.360
63	.853 (273/320)	.837 (159/190)	.135	.714

Table 3

Estimated Haplotype Frequencies in the 151 Case Subjects and 94 Control Subjects, When Individuals Are Included Only If They Are Typed at All Three SNPs

Haplotype			Frequency in						
			All Subjects in Case Families, as Determined by Zaplo	Case Subjects, as Determined by			Control Subjects, as Determined by		
SNP43	SNP19	SNP63		EH	Zaplo	Phase	EH	Zaplo	Phase
1	1	1	.196	.187	.185	.185	.221	.225	.218
1	1	2	.117	.141	.136	.142	.146	.144	.149
1	2	1	.586	.581	.585	.579	.523	.522	.527
1	2	2	.000	.005	.007	.007	.019	.019	.016
2	1	1	.000	.000	.000	.000	.000	.000	.000
2	1	2	.000	.000	.000	.000	.000	.000	.000
2	2	1	.095	.086	.083	.086	.090	.089	.090
2	2	2	.006	.003	.004	.000	.000	.000	.000

Haplotype Combination and the Risk of Type 2 Diabetes

We also calculated the odds ratio (OR) for each haplotype combination relative to all the other haplotype combinations as a group, to test if there was association between the different haplotype combinations and the risk of type 2 diabetes. There was no significant difference in any haplotype combination of SNP43, -19, and -63 between the case subjects and the control subjects (table 5).

Evaluation of LD

Pairwise LD between each pair of SNPs was measured on the basis of the case subjects and the control subjects separately. We performed Fisher's exact test to check LD between each pair of SNPs. LD was observed between SNP43 and SNP19 and between SNP43 and SNP63, both in the case subjects and in the control subjects (table 6).

Multipoint Linkage Analysis

Multipoint linkage analysis in 139 families in the conservative data set for 30 microsatellite markers (ABI panels 3 and 4) indicated that there was no promising peak across chromosome 2 that could be linked to type 2 diabetes. The maximum nonparametric LOD score, which was only 0.40, occurred near D2S286, whereas the LOD score curve was 0 in the region (between D2S125 [at 260.6 cM from pter] and D2S140 [at 263.6 cM from pter]) that contained the three SNPs (data not shown).

Table 4

Association Testing Based on the Haplotype Frequencies of SNP43, -19, and -63 between the Case Subjects and the Control Subjects

Group	*N*	ln(L)	χ^2	*P*[a]
Case	151	−281.97	69.44	.858
Control	94	−190.35	33.87	
Combination	245	−473.96	101.04	

[a] Calculated from the T5 statistic: $2[\ln(L)_{case} + \ln(L)_{control} - \ln(L)_{case + control}]$; df = 7.

American Samoan Control Subjects

Using only the American Samoan control subjects, we repeated all the analyses presented in tables 1–6, and obtained qualitatively similar results (data not shown) that do not alter any of the conclusions presented here.

Discussion

The findings of Horikawa et al. (2000) were unexpected, given that calpains are cysteine proteases and are unlikely to be involved in the pathways of glucose homeostasis. In a commentary, Permutt et al. (2000) noted that the connection between calpain-10 and glucose metabolism is far from clear. Nonetheless, the finding that *CAPN10* was a putative type 2 diabetes susceptibility gene was significant, because it represented the first gene associated with a common and complex disease to be identified by positional cloning. However, validation of these findings in other populations is needed to establish the role of calpain-10 polymorphisms in susceptibility to type 2 diabetes. In a follow-up study among the Pima Indians, it was found that, although homozygosity for the G allele at SNP43 was not associated with a higher prevalence of type 2 diabetes, homozygosity was associated with reduced calpain-10 mRNA levels in muscle cells and with decreased rates of glucose turnover that were suggestive of insulin resistance (Baier et al. 2000). Since then, several studies conducted in various human populations have been reported in abstract form, and they present conflicting evidence of association. In a Finnish population from Botnia, the SNP43 G allele was shown to be associated with insulin resistance and type 2 diabetes (Orho-Melander et al. 2000). In two other

Table 5

Association between the Haplotype Combinations and Type 2 Diabetes

Haplotype Combination	No. of Cases	No. of Controls	OR (95% CI)
111/111	5	3	1.04 (.24–4.45)
111/112	9	7	.79 (.28–2.19)
111/121	31	23	.80 (.43–1.47)
112/112	3	2	.93 (.15–5.69)
112/121	26	12	1.42 (.68–2.98)
112/122	0	2	…
121/121	50	28	1.17 (.67–2.04)
121/122	1	1	.62 (.04–10.03)
221/221	0	1	…
221/111	6	5	.74 (.22–2.48)
221/112	2	3	.41 (.07–2.48)
221/121	17	7	1.58 (.63–3.96)
221/122	1	0	…
Total	151	94	…

samples of white populations (German and Czech), association between type 2 diabetes and calpain-10 polymorphisms was indicated (Schwarz et al. 2001). However, no association was found in a Polish population (Malecki et al. 2001) or in a Chinese population (Ng et al. 2001). In a recently published large study of white populations in Britain and Ireland, Evans et al. (2001) did not find any association, either individually or as haplotypes, with the three SNPs considered here. However, one of the alleles, allele C, at a fourth locus in the *CAPN10* gene, SNP44, was found to be preferentially transmitted to affected children in parent-offspring trios, and these affected individuals had an excess of this allele when compared with population control subjects. Although this finding is interesting, it needs further evaluation, in view of the fact that this association was not observed in a second case-control group or in a set of discordant sibs. The affected offspring in the trios represented a group of individuals who were, on average, younger and more obese than the other diabetic probands in the study. The question thus arises whether SNP44 is associated with a smaller subset of patients with earlier-onset type 2 diabetes, distinct from that found in the typical late-onset type 2 diabetes phenotype. Furthermore, it should be noted that Evans et al. (2001) found no evidence of linkage at the *CAPN10* locus.

In the Samoan populations that we studied, we failed to detect any association between case subjects and control subjects in allele frequencies, haplotype frequencies, or haplotype combinations of SNP43, -19, and -63. Also, our data showed no evidence of linkage in the region of chromosome 2 that contains these SNPs. The Samoans represent a relatively small isolated population that has been subjected to population bottlenecks and genetic drift, so we are not surprised to find that genetic variation in SNP43, -19, and -63 among this group is different from that in other populations. Allele frequencies of these three SNPs in Samoans are more similar to those in Asians than to those in Mexican Americans, Europeans, or Amerindians (Horikawa et al. 2000). Haplotype frequencies in Samoans are different from Mexican Americans, Europeans, Asians, and Amerindians. Therefore, one must consider the effects of genetic and population heterogeneity when attempting to detect susceptibility genes for type 2 diabetes in different populations. Although we also found that the 112/121 haplotype combination appears to increase the risk of type 2 diabetes in Samoans, the OR is not significant (OR = 1.42, 95% confidence interval [CI] = 0.68–2.98) (table 5). The 121/221 haplotype combination has an even higher OR, but the OR is also not significant (OR = 1.58, 95%CI = 0.63–3.96). These findings, together with observations from other worldwide populations noted above, raise the question of whether *CAPN10* is a diabetes susceptibility gene only in some particular ethnic groups. It is likely that evolutionary histories of individual populations have an important role in disease susceptibility loci of complex traits in which the susceptibility alleles have minor individual effects. Thus, the contribution of alleles at a specific locus to type 2 diabetes susceptibility will be likely to vary across different populations. It should be noted that the Samoans have a significantly reduced level of genetic variation when compared with large cosmopolitan populations (Deka et al. 1994). It is not unlikely that in this population, some alleles have become enriched at the expense of others because of genetic drift and related evolutionary forces. One likely scenario is that the calpain-10 polymorphisms are true functional variants; however, they interact with other genetic polymorphisms, whose individual effects on disease susceptibility could have been altered by differences in allele frequency among various populations. Such differences might result from the differing evolutionary histories of the populations, as noted above, or could even arise from the interaction of the calpain-10 polymorphisms with variable environmental exposures. As a result of such varied gene-gene or gene-environment

Table 6

Pairwise Linkage Disequilibrium between SNP43, -19, and -63 in the 151 Case Subjects and the 94 Control Subjects

Pairwise Comparison	Cases[a]			Controls[a]		
	P[b]	D′	Δ^2	P[b]	D′	Δ^2
SNP43 vs. SNP19	.000	1.000	.214	.000	1.000	.240
SNP43 vs. SNP63	.020	1.000	.128	.080	1.000	.140
SNP19 vs. SNP63	.000	.934	.560	.000	.847	.494

[a] D′ = standardized disequilibrium coefficient measurement; Δ^2 = measure of disequilibrium.

[b] By Fisher's exact test.

interactions, the effect of a functional variant could vary between populations, thereby leading to variation in disease susceptibility. In fact, Horikawa et al. (2000) had noted that the risk of disease could be modified by other polymorphisms located in and around the functional variants, which affect the transcriptional activity of the *CAPN10* gene. The risk may also be modified by haplotypic effects created by alleles at loci that are in LD with the functional variants (Evans et al. 2001). The ability to detect associations also depends on mutational, allele, and haplotype frequencies. Thus, it is not surprising that there have been a variety of positive and negative findings regarding *CAPN10* and its association with type 2 diabetes.

Our lack of significant results is unlikely to have resulted from low power, because the number of families in our study (176) is comparable to that (170) in the original study by Horikawa et al. (2000). Horikawa et al. (2000) observed an OR of 2.80 for the 112/121 haplotype combination in 170 Mexican American families and an OR of 3.58 in their second set of 69 Mexican American families. Given our sample sizes, we have >80% power to detect an OR of 2.80 at the significance level of .025, and we have >90% power to detect an OR of 3.58 at the significance level of .010. Thus, if the 112/121 haplotype combination had an effect in the Samoan population similar to that seen in the Mexican American population, we would have had relatively high power to detect it. Even so, the possibility remains that our results may be consistent with those reported by Horikawa et al. (2000), because our OR of 1.42 is within the 95% CI of their populations, except for that of the second set of Mexican Americans. However, we note that our OR may not be strictly comparable, because some of the previous studies selected control subjects at random, whereas we used older nonobese nondiabetic individuals as control subjects.

Another plausible explanation for apparently contradictory findings across studies is that the underlying biological mechanism of the *CAPN10* gene is complex and its functional relevance and physiological roles in glucose metabolism are yet to be understood. In the discussion above, we have noted the findings of Baier et al. (2000), which are initial steps in determining the functional role of the *CAPN10* gene. However, a more recent study contradicts these implications. In a study of 74 nondiabetic German subjects, Stumvoll et al. (2001) found that the G allele at SNP43 was associated with higher glucose-stimulated insulin secretion and more-efficient proinsulin processing. The question remains whether other polymorphisms, within or outside the *CAPN10* gene, act in concert with these identified alleles to enhance disease predisposition (see the discussion above of the findings of Evans et al. [2001]). Because we believe the loci that influence type 2 diabetes susceptibility in the Samoans may be different from those in other populations, we plan to first complete our genomewide scan for such loci and then, on the basis of our results, conduct further fine mapping and detailed case-control studies in the indicated regions.

Acknowledgments

We thank Drs. Nancy J. Cox and Craig L. Hanis for providing the primer sequences and the basic protocols for analysis of the SNPs. We thank Diane Smelser for technical support. We thank Matthew Steele, Pasa Turituri, Asaf Bitton, Abigail Robin, Kirsten Becker, and Dottie Savi'i for assistance in the fieldwork. This study was supported by NIH grants AG09375, HL52611, DK55406, DK59642, and ES06096.

Electronic-Database Information

Accession numbers and URLs for data in this article are as follows:

Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/Omim/ (for type 2 diabetes mellitus [MIM 125853]; *NIDDM1* [MIM 601283]; and *CAPN10* [MIM 605286])

References

Abecasis GR, Cookson WO (2000) GOLD: graphical overview of linkage disequilibrium. Bioinformatics 16:182–183

Baier LJ, Permana PA, Yang X, Pratley RE, Hanson RL, Shen G-Q, Mott D, Knowler WC, Cox NJ, Horikawa Y, Oda N, Bell GI, Bogardus C (2000) A calpain-10 gene polymorphism is associated with reduced muscle mRNA levels and insulin resistance. J Clin Invest 106:69–73

Boehnke M, Cox NJ (1997) Accurate inference of relationships in sib-pair linkage studies. Am J Hum Genet 61:423–429

Collins VR, Dowse GK, Toelupe PM, Aloaina FL, Spark RA, Zimmet PZ (1994) Increasing prevalence of NIDDM in the Pacific Island population of Western Samoa over a 13-year period. Diabetes Care 17:288–296

Deka R, McGarvey ST, Ferrell RE, Kamboh MI, Yu LM, Aston CE, Jin L, Chakraborty R (1994) Genetic characterization of American and Western Samoans. Hum Biol 66:805–822

Devlin B, Risch N (1995) A comparison of linkage disequilibrium measures for fine-scale mapping. Genomics 29:311–322

Ehm MG, Karnoub MC, Sakul H, Gottschalk K, Holt DC, Weber JL, Vaske D, Briley D, Briley L, Kopf J, McMillen P, Nguyen Q, Reisman M, Lai EH, Joslyn G, Shepherd NS, Bell C, Wagner MJ, Burns DK, American Diabetes Association GENNID Study Group (2000) Genomewide search for type 2 diabetes susceptibility genes in four American populations. Am J Hum Genet 66:1871–1881

Elbein SC, Hoffman MD, Teng K, Leppert MF, Hasstedt SJ (1999) A genome-wide search for type 2 diabetes susceptibility genes in Utah Caucasians. Diabetes 48:1175–1182

Evans JC, Frayling TM, Cassell PG, Saker PJ, Hitman GA, Walker M, Levy JC, et al (2001) Studies of association between the gene for calpain-10 and type 2 diabetes mellitus in the United Kingdom. Am J Hum Genet 69:544–552

Galanis DJ, McGarve ST, Quested C, Sio B, Afele-Fa'amuli S (1999) Dietary intake among modernizing Samoans: implications for risk of cardiovascular disease. J Am Diet Assoc 99:184–190

Ghosh S, Watanabe RM, Hauser ER, Valle T, Magnuson VL, Erdos MR, Langefeld CD, et al (1999) Type 2 diabetes: evidence for linkage on chromosome 20 in 716 Finnish affected sib pairs. Proc Natl Acad Sci USA 96:2198–2203

Gudbjartsson DF, Jonasson K, Frigge ML, Kong A (2000) Allegro, a new computer program for multipoint linkage analysis. Nat Genet 25:12–13

Hanis CL, Boerwinkle E, Chakraborty R, Ellsworth DL, Concannon P, Stirling B, Morrison VA, et al (1996) A genome-wide search for human non-insulin-dependent (type 2) diabetes genes reveals a major susceptibility locus on chromosome 2. Nat Genet 13:161–166

Hanson RL, Ehm MG, Pettitt DJ, Prochazka M, Thompson DB, Timberlake D, Foroud T, Kobes S, Baier L, Burns DK, Almasy L, Blangero J, Garvey WT, Bennett PH, Knowler WC (1998) An autosomal genomic scan for loci linked to type II diabetes mellitus and body-mass index in Pima Indians. Am J Hum Genet 63:1130–1138

Hedrick PW (1987) Gametic disequilibrium measures: proceed with caution. Genetics 117:331–341

Hegele RA, Sun F, Harris SB, Anderson C, Hanely AJ, Zinman B (1999) Genome-wide scanning for type 2 diabetes susceptibility in Canadian Oji-Cree, using 190 microsatellite markers. J Hum Genet 44:10–14

Hodge AM, Dowse GK, Toelupe P, Collins VR, Imo T, Zimmet PZ (1994) Dramatic increase in the prevalence of obesity in Western Samoa over the 13 year period 1978–1991. Int J Obes 18:419–428

Horikawa Y, Oda N, Cox NJ, Li X, Orho-Melander M, Hara M, Hinokio Y, et al (2000) Genetic variation in the gene encoding calpain-10 is associated with type 2 diabetes mellitus. Nat Genet 26:163–175

King H, Aubert RE, Herman WH (1998) Global burden of diabetes, 1995–2025. Diabetes Care 21:1414–1431

Lazzeroni LC, Lange K (1998) A conditional inference framework for extending the transmission/disequilibrium test. Hum Hered 48:67–81

Mahtani MM, Widén E, Lehto M, Thomas J, McCarthy M, Brayer J, Bryant B, Chan G, Daly M, Forsblom C, Kanninen T, Kirby A, Kruglyak L, Munnelly K, Parkkonen M, Reeve-Daly MP, Weaver A, Brettin T, Duyk G, Lander ES, Groop LC (1996) Mapping of a gene for type 2 diabetes associated with an insulin secretion defect by a genome scan in Finnish families. Nat Genet 14:90–94

Malecki MT, Klupa T, Zub A, Cyganek K, Moczulski DK, Sieradzki J (2001) Allelic variations in calpain 10, PPARγ genes and risk of type 2 diabetes in a Polish population. Diabetes Suppl 50:A243

McGarvey ST (1991) Obesity in Samoans and a perspective on its etiology in Polynesians. Am J Clin Nutr 53:1586S-1594S

McGarvey ST, Baker PT (1979) The effects of modernization and migration on Samoan blood pressure. Hum Biol 51: 461–475

McGarvey ST, Ganta DR, Turituri P, Quested C, Tufa J, Tuuau-Potoi N (2000) Type 2 diabetes prevalence an incidence in Samoa and American Samoa. Fifth International Meeting on Diabetes in Indigenous Peoples, Christchurch, New Zealand, October 2–6, 2000

McGarvey ST, Levinson PD, Bausserman L, Galanis DJ, Hornick CA (1993) Population change in adult obesity and blood lipids in American Samoa from 1976–1978 to 1990. Am J Hum Biol 5:17–30

McPeek MS, Sun L (2000) Statistical tests for detection of misspecified relationships by use of genome-screen data. Am J Hum Genet 66:1076–1094

Ng MCY, So W-Y, Critchley JAJH, Cockram CS, Chan JCN (2001) Association of calpain 10 genetic polymorphisms with type 2 diabetes and insulin response in non-diabetic Chinese. Diabetes Suppl 50:A234

O'Connell JR (2000) Zero-recombinant haplotyping: applications to fine mapping using SNPs. Genet Epidemiol Suppl 19:S64–70

O'Connell JR, Weeks DE (1998) PedCheck: a program for identification of genotype incompatibilities in linkage analysis. Am J Hum Genet 63:259–266

Orho-Melander M, Åberg M, Berglund A, Horikawa Y, Cox N, Oda N, Groop L (2000) Genetic variation in the calpain 10 gene is associated with insulin resistance and type 2 diabetes. 36th Annual Meeting of the European Association for the Study of Diabetes, Jerusalem, September 17–21, 2000 (abstract 200)

Permutt MA, Bernal-Mizrachi E, Inoue H (2000) Calpain 10: the first positional cloning of a gene for type 2 diabetes. J Clin Invest 106:819–821

Permutt MA, Wasson JC, Suarez BK, Lin J, Thomas J, Meyer J, Lewitzky S, Rennich JS, Parker A, DuPrat L, Maruti S, Chayen S, Glaser B (2001) A genome scan for type 2 diabetes susceptibility loci in a genetically isolated population. Diabetes 50:681–685

Schwarz PEH, Horikawa Y, Vcelak J, Selisko T, Rietzsch H, Bendlova B, Schulze J, Cox NJ (2001) Genetic variation of CAPN10 affects susceptibility to type 2 diabetes in German and Czech population. Diabetes Suppl 50:A232

Stephens M, Smith NJ, Donnelly P (2001) A new statistical method for haplotype reconstruction from population data. Am J Hum Genet 68:978–989

Stumvoll M, Fritsche A, Stefan N, Madaus A, Weisser M, Machicao F, Häring H (2001) Functional significance of the UCSNP-43 polymorphism in the CAPN10 gene for proinsulin processing and insulin secretion in non-diabetic humans. Diabetes Suppl 50:A239

Terwilliger J, Ott J (1994) Handbook of human genetic linkage. Johns Hopkins University Press, Baltimore

US Department of Commerce, Bureau of the Census. 1990 Census of population and housing: social, economic and housing characteristics. American Samoa. 1990 CPH-6-AS. US Government Printing Office, Washington, DC

Watanabe RM, Ghosh S, Langefeld CD, Valle TT, Hauser ER, Magnuson VL, Mohlke KL, et al (2000) The Finland-United States investigation of non-insulin-dependent diabetes mellitus genetics (FUSION) study. II. An autosomal genome scan

for diabetes-related quantitative-trait loci. Am J Hum Genet 67:1186–1200

Xie X, Ott J (1993) Testing linkage disequilibrium between a disease gene and marker loci. Am J Hum Genet 53:A1107

Zhao JH, Curtis D, Sham PC (2000) Model-free analysis and permutation tests for allelic associations. Hum Hered 50: 133–139

Zimmet P (1979) Epidemiology of diabetes and its macrovascular manifestations in Pacific populations: the medical effects of social progress. Diabetes Care 2:144–153