Brief Genetics Report

Calsquestrin 1 (*CASQ1*) Gene Polymorphisms Under Chromosome 1q21 Linkage Peak Are Associated With Type 2 Diabetes in Northern European Caucasians

Swapan Kumar Das,[1] Winston Chu,[1] Zhengxian Zhang,[1] Sandra J. Hasstedt,[2] and Steven C. Elbein[1,3]

Genome-wide scans in multiple populations have identified chromosome 1q21-q24 as one susceptibility region for type 2 diabetes. To map the susceptibility genes, we first placed a dense single nucleotide polymorphism (SNP) map across the linked region. We identified two SNPs that showed strong associations, and both mapped to within intron 2 of the calsequestrin 1 (*CASQ1*) gene. We tested the hypothesis that sequence variation in or near *CASQ1* contributed to type 2 diabetes susceptibility in Northern European Caucasians by identifying additional SNPs from the public database and by screening the *CASQ1* gene for additional variation. In addition to 15 known SNPs in this region, we found 8 new SNPs, 3 of which were in exons. A single rare nonsynonymous SNP in exon 11 (A348V) was not associated with type 2 diabetes. The associated SNPs were localized to the region between −1,404 in the 5′ flanking region and 2,949 in intron 2 ($P = 0.002$ to $P = 0.034$). No SNP 3′ to intron 2, including the adjacent gene *PEA15*, showed an association. The strongest associations were restricted to individuals of Northern European ancestry ascertained in Utah. A six-marker haplotype was also associated with type 2 diabetes ($P = 0.008$), but neither transmission disequilibrium test nor family-based association studies were significant for the most strongly associated SNP in intron 2 (SNP CASQ2312). An independent association of SNPs in introns 2 and 4 with type 2 diabetes is reported in Amish families with linkage to chromosome 1q21-q24. Our findings suggest that noncoding SNPs in *CASQ1* alter diabetes susceptibility, either by a direct effect on *CASQ1* gene expression or perhaps by regulating a nearby gene such as *PEA15*. *Diabetes* 53:3300–3306, 2004

Considerable data support a genetic etiology for type 2 diabetes, but multiple susceptibility loci on different chromosomes are likely involved (1). We previously mapped one type 2 diabetes susceptibility locus in extended North European Caucasian families to chromosome 1q21-q24 (2), a region that is now well replicated in Old Order Amish, Pima Indian, Chinese, French, and British families (1). Recently, we completed a dense linkage map and identified two linkage peaks centered on locations 157 and 164 Mb (3). Additionally, we have previously reported the association of polymorphisms of the liver pyruvate kinase enzyme with type 2 diabetes, located centromeric to the first linkage peak (4). However, analysis of many strong candidate genes under the linkage peaks have not yet identified an association with type 2 diabetes that could explain the linkage.

To further map the susceptibility loci within the region of linkage, we collaboratively placed 580 single nucleotide polymorphisms (SNPs) over a 20-Mb region that encompassed both linkage peaks using MassArray MALDI-TOF mass spectrometry (Sequenome) (5,6). We typed DNA pools constructed from individual samples in three populations: Utah Caucasians, Amish Caucasians, and Pima Indians. Of 92 SNPs that showed evidence for association in the original sample, we selected 14 SNPs to test in individual samples based on significance on replication in pools ($P < 0.05$). Two adjacent SNPs were highly significant in pools ($P = 0.00008$ and 0.003), subsequently confirmed in individual samples by a different method, and mapped to the first and larger of the two linkage peaks (157 Mb). The two SNPs were localized to intron 2 of the calsequestrin 1 (*CASQ1*) gene in a region that was also associated with type 2 diabetes in the Amish Family Diabetes Study (AFDS) (7) and that also included the gene *PEA15*, which was reported to show increased expression in tissues from diabetic subjects (8).

To test the hypothesis that sequence variants in or near the *CASQ1* gene increase susceptibility to type 2 diabetes and to define the region of association, we undertook a two-stage study. We first evaluated SNPs reported in the public database and extending 5′ and 3′ of *CASQ1*, including adjacent genes. Subsequently, we screened the 11 exons, untranslated regions, 3′ and 5′ flanking regions, and all of intron 2 of the *CASQ1* gene for additional sequence

variation in North European Caucasian subjects. We tested each SNP for an association with type 2 diabetes. The most promising SNPs were tested for family-based associations with glucose levels and type 2 diabetes.

RESEARCH DESIGN AND METHODS

CASQ1 gene variation was detected in 24 unrelated Caucasian individuals from Utah families linked to the 1q21 region, including 8 nondiabetic family members and 16 individuals with type 2 diabetes. Initial studies to follow up on the pooled associations were conducted in 129 individuals with type 2 diabetes and 108 control individuals from Utah that overlapped with but slightly expanded the sample constructed for the original pooling experiment. Most individuals were selected from the families used in the linkage study. We subsequently sought to expand the sample size. Because insufficient control subjects were available from Utah, we examined two sample sets. The first set comprised 190 diabetic individuals and 119 control individuals, all ascertained in Utah for Northern European ancestry and including the 129 case and 108 control subjects used in the first follow-up studies. To improve our power, albeit at the expense of maintaining homogeneity, we also examined an expanded Caucasian sample of 191 Caucasian individuals with diabetes and a first-degree relative with diabetes and 191 control individuals with a normal glucose tolerance test or a fasting or random glucose <5.6 mmol/l and no family history of diabetes in a first-degree relative. The expanded sample included additional similarly ascertained individuals from Arkansas with mixed Caucasian ancestry (1 diabetic subject and 72 nondiabetic control subjects). Of the diabetic subjects for both analyses, 70 individuals were ascertained from Utah families used in the linkage studies and the remainder were ascertained for diabetes by history and medication or documented by oral glucose tolerance test using World Health Organization criteria. All individuals for the case-control studies were unrelated. Because some samples were no longer available or were of poor quality, some case and control samples from Utah in the original association study were not included in the expanded sample.

Transmission disequilibrium test and family-based association studies were conducted on 704 members of 68 families from Utah and of Northern European ancestry, 292 of whom were considered affected, as previously described (2,3).

Subjects ascertained in Utah provided written informed consent under a protocol approved by the University of Utah Institutional Review Board. Subjects studied in Arkansas provided written informed consent under protocols approved by the University of Arkansas for Medical Sciences Human Research Advisory Committee.

Genotypic analysis and marker selection. Additional markers in the *CASQ1* region were chosen from dbSNP (9) for the region from 42 kb upstream of *CASQ1* to 20 kb downstream of the gene. All SNPs except rs617599 were typed by pyrosequencing using the PSQ-96 (Pyrosquencing, Uppsala, Sweden) according to the manufacturer methods but modified to use a universal biotinylated primer (sequences available from authors). SNP rs617599 was typed by using oligonucleotide ligation assay with radioactive labeling and detection with a Storm optical scanner (Molecular Dynamics, Sunnyvale, CA), as described previously (4). The insertion/deletion variant rs3838216 was typed using infrared dye–labeled M13 primers on a LI-COR GR4200 sequencer and scored using Gene ImagIR software (version 3.5.6; Scanalytics, Fairfax, VA). All variants were in Hardy-Weinberg equilibrium.

Detection of *CASQ1* sequence variants. We screened a total of 7,647 bp of sequence, including 1,326 bp of the 5′ flanking sequence (1,489 bp upstream of ATG start site), 832 bp of the 3′ flanking sequence, the 5′ and 3′ untranslated regions, all 11 exons, and 100–200 bp of the sequence flanking each exon, for mutations using 20 sets of primers (sequences available from authors). Amplicons were designed using WAVEMAKER software (version 4.0; Transgenomic, Omaha, Nebraska), with sizes of 247–627 bp, and were screened at up to three denaturing temperatures using denaturing high-pressure liquid chromatography on a WAVE[HT] DNA Fragment Analysis System (Transgenomic). Both strands of fragments showing altered migration were sequenced using the CEQ 8000 Genetic analysis system (Beckman-Coulter, Fullerton, CA). Additionally, we directly sequenced intron 2 (860 bp) of *CASQ1* in two overlapping fragments to catalog additional variation near the associated markers in six individuals selected to represent each of the three observed haplotypes constructed from the two intron 2 SNPs, rs617599 and rs617698.

Putative transcription factor binding sites in intron 2 were detected using TFSEARCH (version 1.3; available at http://www.cbrc.jp/research/db/TFSEARCH.html) using the TRANSFAC database (10).

Statistical analysis. Allele frequencies were compared using the Fisher's exact test. For all statistical tests, we considered $P < 0.05$ to be evidence of significance. We report the P values without Bonferroni correction. Secondary analyses of genotypic association were computed using Pearson's χ^2 test under additive, dominant, and recessive models. Pairwise haplotypes were estimated from the combined case and control population data using the expectation maximization algorithm, from which two linkage disequilibrium (LD) statistics, D′ and r^2, were calculated. Extended haplotypes were predicted using the expectation maximization algorithm, as implemented in the Arlequin program. Because SNPs CASQ −1404, 1074, 2312, 2351, 2399, and 2949 all showed an association or a trend to an association with type 2 diabetes in our Utah Caucasian sample set, we compared the frequency of this six-marker haplotype among case and control subjects. Studies of fasting and 30-, 60-, 90-, and 120-min postchallenge glucose levels were performed on the full set of families included in the linkage study (2), including families not showing linkage to chromosome 1q21, using the Pedigree Analysis Package (11). The total number of subjects was 567 for fasting glucose, 244 for 30-min glucose, 483 for 60-min glucose, 242 for 90-min glucose, and 459 for 120-min glucose. The number of unaffected individuals was 372 for fasting glucose, 178 for 30-min glucose, 360 for 60-min glucose, 176 for 90-min glucose, and 360 for 120-min glucose. Excess transmission of alleles from parents to affected offspring was tested by a maximum likelihood implementation of the transmission disequilibrium test, as previously described (4,12). Genotype associations with type 2 diabetes were tested using the logit of the probability of type 2 diabetes as the dependent variable. The genotype at each SNP was coded as 1, 2, or 3 (homozygous common allele, heterozygous rare allele, and homozygous rare allele, respectively) for the independent variable. The logit of the probability of type 2 diabetes was assumed to equal $\alpha_i + \beta$ (age − 45) + γ (BMI − 30), where i is 1, 2, or 3. Nonindependence of observations was accounted for by a polygenic component, and penetrance at age 45 years and a BMI of 30 kg/m² was estimated for each genotype as $f_i = e^{\alpha i}/(1 + e^{\alpha i})$. SNP effects on type 2 diabetes were tested as the difference in penetrance between genotypes (2 degrees of freedom).

RESULTS

We initially evaluated 36 SNPs from the public database over the 100-kb region that surrounded the initial observation, encompassing genes *ATP1A4*, *CASQ1*, *PEA15*, and *H326*. Of these, 15 SNPs and one insertion/deletion variant were both polymorphic and able to be assayed, spanning the region from −41,280 to 20,346 bp relative to the ATG start of *CASQ1* (Table 1), for an average interval between polymorphic markers of 6.3 kb. The only associations were identified within *CASQ1*; hence, we searched for additional variants by screening the exons, the immediate 5′ flanking region, and the immediate 3′ region. We confirmed the seven SNPs previously selected and identified seven new variants (Table 1 and Fig. 1). We failed to confirm one database SNP (rs822451, intron 6). We identified rare (frequency <5%) synonymous variants in exons 3 and 4 and a rare single nonsynonymous SNP in exon 11 (CASQ1 10476; A348V) that was not associated with type 2 diabetes. We were unable to develop assays for two of the new SNPs identified, and three of these were rare and not typed in the full sample (Table 1).

The results are summarized in Table 1 and Fig. 1. Significant association or trends to an association were limited to intron 2 and to SNPs in the proximal 5′ flanking region and were therefore most consistent with an association in *CASQ1* rather than the neighboring genes *ATP1A4* or *PEA15*. We found the most significant associations in samples used to confirm the initial pooled findings, for which SNPs −1404, 2312, 2399, and 2949 all showed nominal significance ($P = 0.0016$–0.04). The results when all available Utah samples were tested, including additional case and control subjects, are shown in Table 1. When additional Caucasian samples of non–Northern European ancestry were included, the significant allelic association was limited to SNP CASQ-1404 in the 5′ flanking region. Based on the concentration of associated

TABLE 1
CASQ1 region variation and association with type 2 diabetes

Lab name	rs number	Location	Position (bp)	Variant	Frequency, Utah cases	Frequency, Utah control subjects	P	Frequency, all cases	Frequency, all control subjects	P
CASQ −41280	rs911229	5′ flank	−41280	C/G	0.140	0.185	0.140	0.141	0.179	0.138
CASQ −34422	rs927203	5′ flank	−34422	T/C	0.447	0.461	NS	0.447	0.448	NS
CASQ −29131	rs953104	5′ flank	−29131	A/G	0.132	0.142	NS	0.131	0.150	NS
CASQ −18446	rs705843	5′ flank	−18446	G/C	0.243	0.244	NS	0.245	0.230	NS
CASQ −14387	rs881291	5′ flank	−14387	T/C	0.315	0.311	NS	0.318	0.325	NS
CASQ −13058	rs731852	5′ flank	−13058	G/T	0.356	0.350	NS	0.355	0.351	NS
CASQ −8503	rs1935306	5′ flank	−8503	A/G	0.479	0.457	NS	0.479	0.441	NS
CASQ −3318	rs860294	5′ flank	−3318	C/A	0.345	0.315	NS	0.346	0.313	NS
CASQ −1404	rs1186694	5′ flank	−1404	C/T	0.190	0.130	0.056	0.202	0.139	0.021
CASQ −887	New	5′ flank	−887	G/A	ND*	ND	—	ND	ND	—
CASQ −512	New	5′ flank	−512	G/A	ND†	ND	—	ND	ND	—
CASQ 1074	rs3838216	Intron 1	1074	Del/Ins	0.240	0.304	0.084	0.242	0.267	NS
CASQ 2312	rs617698	Intron 2	2312	A/G	0.434	0.349	0.038	0.434	0.403	NS
CASQ 2351	New	Intron 2	2351	G/A	0.230	0.310	0.058	0.230	0.283	0.098
CASQ 2399	rs617599	Intron 2	2399	G/A	0.449	0.317	0.004	0.437	0.404	NS
CASQ 2949	rs3747623	Intron 2	2949	C/T	0.422	0.330	0.029	0.423	0.387	NS
CASQ 3059	New	Exon 3	3059	C/T	ND†	ND	—	ND	ND	—
CASQ 4292	New	Exon 4	4292	G/A	ND†	ND	—	ND	ND	—
CASQ 4535	rs2275703	Intron 4	4535	A/C	0.460	0.509	0.241	0.461	0.474	NS
CASQ 4891	New	Intron 5	4891	C/T	ND*	ND	—	ND	ND	—
CASQ 7808	rs822450	Intron 7	7808	G/A	0.500	0.454	0.398	0.497	0.481	NS
CASQ 10171	rs3827532	Intron 10	10171	C/T	0.028	0.015	NS	0.024	0.011	NS
CASQ 10476	New	Exon 11	10476	A348V	0.005	0.004	NS	0.0027	0.0026	NS
CASQ 11413	rs686015	3′ flank	11413	T/A	0.497	0.447	0.130	0.495	0.479	NS
CASQ 11845	New	3′ flank	11845	A/C	0.016	0.015	NS	0.016	0.008	NS
CASQ 20346	rs680083	3′ flank	20346	C/T	0.505	0.454	0.186	0.503	0.479	NS

Shown are the SNPs flanking and in the *CASQ1* gene, identified either from the public database and confirmed or from screening. Lab name refers to Fig. 1, rs number is the dbSNP catalog number, and position is relative to the ATG start of *CASQ1*. Frequency is provided for the samples strictly ascertained from Utah, including 190 case and 119 control subjects, and for the 191 case and control subjects, including additional samples more broadly ascertained. Del/Ins, deletion/insertion; ND, not determined; NS, not significant ($P > 0.2$). *Not determined due to extended repetitive sequences and assay design problems; †not typed due to low allele frequency (<0.05) All variants are listed with the common allele first, and frequencies are reported for the minor allele.
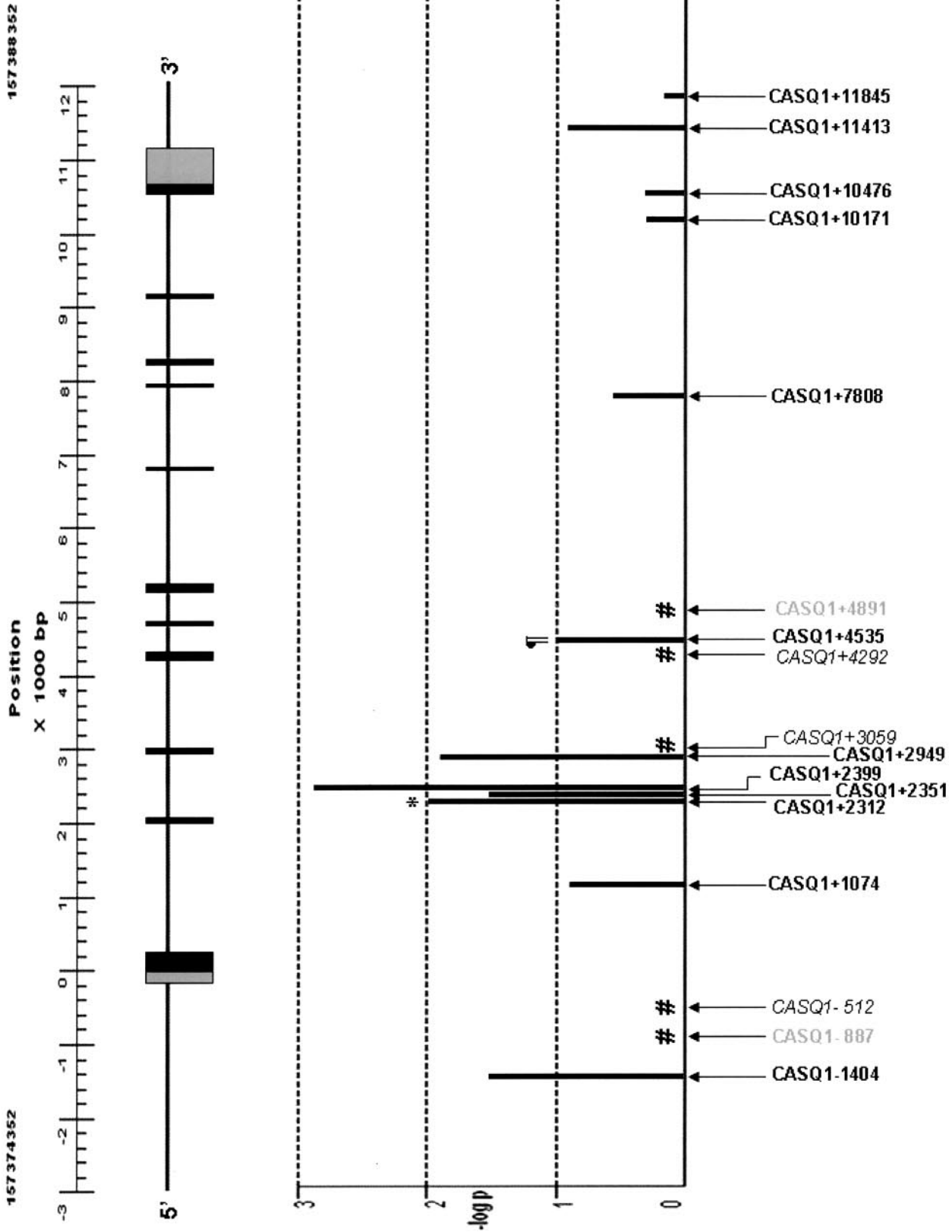
FIG. 1. Map of *CASQ1* gene. The map shows the location of the SNPs and exons for the *CASQ1* gene. The location of each SNP is shown below the figure as a bar if the SNP was typed or indicated by # if the SNP was not typed in the full case-control population. SNPs typed only to establish LD are shown in italics, and those for which assays could not be designed are shown in gray. Exons are shown as boxes, with translated regions shown in black and untranslated regions shown in light gray. SNPs that are significant (*P* < 0.05) in both Utah and Amish populations are indicated by *; SNPs significant in the Amish but not Utah population are indicated by ¶.

SNPs in intron 2, which was not originally screened, we resquenced this intron in six individuals chosen for each of the observed haplotypes. One additional SNP (SNP CASQ 2351) was identified (Table 1) that also showed a trend to an association (*P* = 0.058), with the minor allele overrepresented among cases. The haplotype distribution for six SNPs from the 5′ flanking region (−1,404) to intron 2 (2,949) was significantly different in case and control subjects (*P* = 0.008); the difference remained significant, albeit less so, among the full sample set (*P* = 0.013) (Table 2). Analysis of genotype associations for individual SNPs was most consistent with a dominant or additive model, with the minor allele as the risk allele for SNPs CASQ −1404, 1074, 2312, 2351, 2399, and 2949 (data not shown). We typed SNP CASQ 2312 (rs617698), which was associated with type 2 diabetes in the case-control study, in the full family sample. No allele was transmitted in excess from parents to affected offspring, and diabetes penetrance did not differ among genotypes, whether in all families or in only those families linked to this region.

The pairwise LD statistics across SNPs from 5′ to 3′ is shown in Figs. 2 and 3 in the online appendix (available at http://diabetes.diabetesjournals.org). When haplotype blocks were defined using D′ and the methods of Gabriel et al. (13), SNPs 1074 (in intron 1) through 2949 (just upstream of exon 3) were in a single haplotype block (Fig. 3*A*, online appendix) comprising only three common haplotypes (Fig. 3*B*, online appendix), whereas associated SNP −1,404 fell 5′ to this block. Using the r^2 statistic, which best indicates the similarity in the information about association between two SNPs, only four clusters of SNPs had r^2 values >0.65: −13058/−3318, 1074/2351, 2312/2399/2949, and 7808/11413/20346. Using the LDSelect program (14), 12 SNPs would be required to capture the haplotypes at r^2 = 0.64 and 13 SNPs at r^2 = 0.80 (Table 3, online appendix). Notably, the SNPs associated with type 2 diabetes or showing a trend to an association in this study fell into three bins by r^2, bin 8 (−1404), bin 9 (2312/2399/2949), and bin 10 (1074/2351) (Table 3, online appendix), but only two haplotype blocks (Fig. 3, online appendix). Combined, these six SNPs form the highest-risk haplotype (Table 2). Surprisingly, analysis of SNP −1404 and the block 2 tag SNPs 2312 and 2351 dropped the haplotype significance to 0.023 from 0.008 for the six-marker combination among Northern European Caucasians.

Among the Amish, CASQ1 SNPs were associated with elevated glucose. We tested for a similar association with fasting and 30-, 60-, 90-, and 2-h postchallenge glucose levels, both in all family members and in nondiabetic family members only. We found no significant association with SNP CASQ1 2312 for any glucose measure in either group.

DISCUSSION

Chromosome 1q is among the best replicated regions linked to type 2 diabetes. The region is also linked to other potential elements of the metabolic syndrome: hypertension (15) and familial combined hyperlipidemia (16,17). Work from our laboratory suggests that this region harbors at least three loci based on association (4) and linkage (3) studies. The current analyses were initiated based on a SNP map that encompassed both linkage peaks and which showed an association of two SNPs in this

TABLE 2

Haplotype frequencies of six-marker haplotype in case and control subjects

Haplotype −1404, 1074, 2312, 2351, 2399, 2949	Utah control subjects	Utah case subjects	All control subjects	All case subjects
C Del **G C** A T	0.194	0.269	0.191	0.269
C Del **A C** G C	0.194	0.188	0.186	0.188
C Ins **A C** G C	0.178	0.107	0.181	0.107
C Del **A T** G C	0.133	0.043	0.128	0.043
C Del **G T** A T	0.1	0.094	0.108	0.094
C Ins **A T** G C	0.067	0.68	0.059	0.068
T Del **A C** G C	0.044	0.068	0.049	0.068
T Ins **A C** G C	0.05	0.06	0.049	0.06
T Del **G C** A T	0.022	0.047	0.025	0.047
Global *P* value	0.008		0.013	

Shown are the haplotype frequencies for the six-marker haplotype covering −1404 to 2949 in both Northern European Caucasians ascertained in Utah and in all samples tested. Note that the case samples are nearly identical in the two analyses. SNPs selected as tagSNPs from HaploView are marked in bold. SNP1074 is a 17-base insertion/deletion; the larger allele is marked Ins and the smaller Del in the haplotypes.

region. The broadly defined region of association includes three genes, *ATP1A4*, *CASQ1*, and *PEA15*. *PEA15* is overexpressed in fibroblasts, adipose, and skeletal muscle from diabetic subjects (8), is a substrate from protein kinase C, alters glucose transport (8), and is thus a strong candidate gene. However, a previous study of *PEA15* in Pima Indians found no coding SNPs and no association with type 2 diabetes (18), and our association clearly does not extend 3′ of *CASQ1* into the *PEA15* region. Instead, our data restrict the association to the six SNP haplotypes that extend from the 5′ flanking region through intron 2 of *CASQ1*: −1404, 1074, 2312, 2,351, 2399, and 2949. Nonetheless, we cannot exclude the possibility that these associated variants modulate *PEA15* expression.

The associated SNPs belong to three noncontiguous bins based on the r^2 LD statistic. Based on D′ values, SNP −1404 is not in the block defined by SNPs 1074–2949 (Figs. 2 and 3, online appendix). Hence, the observed associations are partially independent and not related exclusively to strong LD between closely spaced SNPs. This observation, along with data from Fu et al. (19) from the AFDS, suggests that multiple *CASQ1* SNPs may contribute to disease susceptibility. Like calpain 10 and many other recently described complex disease genes (20–22), no associated variant alters the coding sequence and no single variant is likely to explain the association. However, we did not observe a much stronger association of predicted haplotypes than of individual SNPs.

Additional support for *CASQ1* as a candidate comes from the independent association of SNPs within *CASQ1* in the AFDS (7,19). That study also found an association at SNP 2312 (rs617698), which in both the Amish and Utah populations is in strong LD with SNP 2399 that showed the strongest initial association in our studies. However, much of the association in the Amish study maps to intron 4 (SNP 4535 in our nomenclature; rs2275703), where we have not found any evidence for association. Notably, SNP 4535 is not in LD with SNPs in intron 2. In contrast to the AFDS, where the haplotype 2312/4535 was significantly associated with type 2 diabetes, this combination showed

no association with type 2 diabetes in either our restricted population (Utah only) or our full case-control study (global *P* value >0.24).

Despite the evidence for replicated association of SNPs in the *CASQ1* gene with type 2 diabetes, as has been typical for complex disease genes, the significance of the associations for individual SNPs is modest. The predicted odds ratios are in the range of 1.5–1.8. Additionally, as has also been observed for other type 2 diabetes susceptibility genes, the SNPs associated in the two populations show only partial overlap. These differences may result from multiple susceptibility alleles, differing LD between founder (Amish) and outbred (Utah) populations, or spurious associations in one or both populations. Finally, where the same SNPs were associated in both Amish and Utah samples (intron 2), we show opposite associations (minor allele is the risk allele in Utah, major allele is the risk allele in the Amish). Possible explanations for this paradox include different susceptibility SNPs at some distance from intron 2 that were not detected in this study, different gene-gene interactions in the two populations, or an effect of gene-environment interactions in populations with different lifestyles.

In our studies, the differences between case and control subjects were observed primarily when we used the restricted population of case and control subjects ascertained in Utah. The two association studies differ primarily in the control populations. The Utah population is exclusively of Northern European origin and maps most closely to the U.K. and Scandinavia. The exact origins of our Arkansas Caucasian population are broad and may contain more Native-American and African-American admixture, although known admixture was avoided. Whether the differences observed between these populations represent true differences in allele frequency, the effects of increasing the sample size or more complex underlying gene-environment interactions are unknown. In general, Utah and Arkansas populations show similar allele frequencies.

We were unable to confirm the association with type 2 diabetes in our Utah families using family-based approaches, and unlike the AFDS, we did not find evidence that those SNPs associated with type 2 diabetes also influenced fasting or postchallenge glucose levels among either all family members or nondiabetic family members only. These failures most likely relate to the low power of family- as opposed to population-based approaches, although we cannot formally exclude population stratification or a type 2 error in the original observation. The number of subjects included in our analysis of glucose levels was considerably lower than that analyzed in the Amish, in part due to the larger number of diabetic family members in our study and the slightly smaller overall study size. Despite these uncertainties, the replication by nearby SNPs not in LD with the initial observation and independent observations in a second population both support a role for SNPs within *CASQ1* in the pathogenesis of type 2 diabetes.

CASQs are major calcium binding proteins localized in the terminal cisternae of the sarcoplasmic reticulum. *CASQ1*, the primary skeletal muscle CASQ, binds calcium with high capacity but moderate affinity (23) and modulates activity of the ryanodine receptor to control muscle calcium homeostasis and thus excitation-contraction coupling (24). Inactivating mutations of *CASQ2*, the cardiac CASQ, are known to result in cardiac arrhythmias, and overexpression of *CASQ2* in mice causes cardiac hypertrophy and cardiomyopathy (25), but little is known about genetic variation in *CASQ1*. Nonetheless, the regulation of muscle calcium may have implications for insulin action. *CASQ1* is upregulated in the streptazotocin-induced diabetic rat skeletal muscle (26). GLUT4 can be recruited by either of two pathways, one proceeding from insulin through phosphatidylinositol 3 kinase and the other through muscle contraction, exercise, calcium release, and hyperglycemia (27,28). Hence, variation in *CASQ1* expression might easily alter myocellular calcium homeostasis, GLUT4 recruitment, and insulin sensitivity.

The SNPs in intron 2 reside in a region of marked DNA conservation between mouse and human. Such conserved nongenic sequences were recently proposed as additional regulatory elements (29). Indeed, SNPs 2351, 2399, and 2949 alter putative binding sites for transcription factors TATA, MZF1 and P300, and SP1, respectively (Fig. 3, online appendix). Of these, SNP 2399 is predicted to abolish binding for MZF1 and P300 and SNP 2949 is predicted to abolish binding for SP1 transcription factors. Thus, this region might alter *CASQ1* or perhaps *PEA15* regulation and contribute to diabetes risk. Interestingly, considerable data have implicated SP1 in the downstream actions of insulin, although the role of SP1 in intronic regions is unknown (30). Additional studies are needed to explore gene expression in muscle among individuals with the high-risk haplotype in order to determine whether insulin sensitivity is altered among carriers of high-risk SNPs and to confirm the association of SNPs in this region with type 2 diabetes in additional populations with linkage to 1q21.

REFERENCES

1. McCarthy MI: Growing evidence for diabetes susceptibility genes from genome scan data. *Curr Diab Rep* 3:159–167, 2003
2. Elbein SC, Hoffman MD, Teng K, Leppert MF, Hasstedt SJ: A genome-wide search for type 2 diabetes susceptibility genes in Utah Caucasians. *Diabetes* 48:1175–1182, 1999
3. Das SK, Hasstedt SJ, Zhang Z, Elbein SC: Linkage and association mapping of a chromosome 1q21–q24 type 2 diabetes susceptibility locus in Northern European Caucasians. *Diabetes* 53:492–499, 2004
4. Wang H, Chu W, Das SK, Ren Q, Hasstedt SJ, Elbein SC: Liver pyruvate kinase polymorphisms are associated with type 2 diabetes in northern European Caucasians. *Diabetes* 51:2861–2865, 2002
5. Werner M, Sych M, Herbon N, Illig T, Konig IR, Wjst M: Large-scale determination of SNP allele frequencies in DNA pools using MALDI-TOF mass spectrometry. *Hum Mutat* 20:57–64, 2002
6. Buetow KH, Edmonson M, MacDonald R, Clifford R, Yip P, Kelley J, Little DP, Strausberg R, Koester H, Cantor CR, Braun A: High-throughput development and characterization of a genomewide collection of gene-based single nucleotide polymorphism markers by chip-based matrix-

assisted laser desorption/ionization time-of-flight mass spectrometry. *Proc Natl Acad Sci U S A* 98:581–584, 2001

7. Hsueh WC, St Jean PL, Mitchell BD, Pollin TI, Knowler WC, Ehm MG, Bell CJ, Sakul H, Wagner MJ, Burns DK, Shuldiner AR: Genome-wide and fine-mapping linkage studies of type 2 diabetes and glucose traits in the Old Order Amish: evidence for a new diabetes locus on chromosome 14q11 and confirmation of a locus on chromosome 1q21–q24. *Diabetes* 52:550–557, 2003

8. Condorelli G, Vigliotta G, Iavarone C, Caruso M, Tocchetti CG, Andreozzi F, Cafieri A, Tecce MF, Formisano P, Beguinot L, Beguinot F: PED/PEA-15 gene controls glucose transport and is overexpressed in type 2 diabetes mellitus. *EMBO J* 17:3858–3866, 1998

9. Smigielski EM, Sirotkin K, Ward M, Sherry ST: dbSNP: a database of single nucleotide polymorphisms. *Nucleic Acid Res* 28:352–355, 2000

10. Heinemeyer T, Wingender E, Reuter I, Hermjakob H, Kel AE, Kel OV, Ignatieva EV, Ananko EA, Podkolodnaya OA, Kolpakov FA, Podkolodny NL, Kolchanov NA: Databases on transcriptional regulation: TRANSFAC, TRRD and COMPEL. *Nucleic Acid Res* 26:362–367, 1998

11. Hasstedt SJ: *PAP: Pedigree Analysis Package*. Version 5. Salt Lake City, UT, University of Utah, 2001

12. 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* 87:650–654, 2002

13. Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B, Higgins J, DeFelice M, Lochner A, Faggart M, Liu-Cordero SN, Rotimi C, Adeyemo A, Cooper R, Ward R, Lander ES, Daly MJ, Altshuler D: The structure of haplotype blocks in the human genome. *Science* 296:2225–2229, 2002

14. Carlson CS, Eberle MA, Rieder MJ, Yi Q, Kruglyak L, Nickerson DA: Selecting a maximally informative set of single-nucleotide polymorphisms for association analyses using linkage disequilibrium. *Am J Hum Genet* 74:106–120, 2004

15. Hunt SC, Ellison RC, Atwood LD, Pankow JS, Province MA, Leppert MF: Genome scans for blood pressure and hypertension: the National Heart, Lung, and Blood Institute Family Heart Study. *Hypertension* 40:1–6, 2002

16. Coon H, Myers RH, Borecki IB, Arnett DK, Hunt SC, Province MA, Djousse L, Leppert MF: Replication of linkage of familial combined hyperlipidemia to chromosome 1q with additional heterogeneous effect of apolipoprotein A-I/C-III/A-IV locus: the NHLBI Family Heart Study. *Arterioscler Thromb Vasc Biol* 20:2275–2280, 2000

17. Pajukanta P, Nuotio I, Terwilliger JD, Porkka KV, Ylitalo K, Pihlajamaki J, Suomalainen AJ, Syvanen AC, Lehtimaki T, Viikari JS, Laakso M, Taskinen MR, Ehnholm C, Peltonen L: Linkage of familial combined hyperlipidaemia to chromosome 1q21- q23. *Nat Genet* 18:369–373, 1998

18. Wolford JK, Bogardus C, Ossowski V, Prochazka M: Molecular characterization of the human PEA15 gene on 1q21–q22 and association with type 2 diabetes mellitus in Pima Indians. *Gene* 241:143–148, 2000

19. Fu M, Damcott C, Sabra M, Pollin TI, Ott SH, Wang J, Garant MJ, O'Connell JR, Mitchell BD, Shuldiner AR: Polymorphism in the calsequestrin 1 (*CASQ1*) gene on chromosome 1q21 is associated with type 2 diabetes in the Old Order Amish. *Diabetes* 53:3292–3299, 2004

20. Cox NJ, Hayes MG, Roe CA, Tsuchiya T, Bell GI: Linkage of calpain 10 to type 2 diabetes: the biological rationale. *Diabetes* 53 (Suppl. 1):S19–S25, 2004

21. Helgadottir A, Manolescu A, Thorleifsson G, Gretarsdottir S, Jonsdottir H, Thorsteinsdottir U, Samani NJ, Gudmundsson G, Grant SF, Thorgeirsson G, Sveinbjornsdottir S, Valdimarsson EM, Matthiasson SE, Johannsson H, Gudmundsdottir O, Gurney ME, Sainz J, Thorhallsdottir M, Andresdottir M, Frigge ML, Topol EJ, Kong A, Gudnason V, Hakonarson H, Gulcher JR, Stefansson K: The gene encoding 5-lipoxygenase activating protein confers risk of myocardial infarction and stroke. *Nat Genet* 36:233–239, 2004

22. Barroso I, Luan J, Middelberg RP, Harding AH, Franks 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* 1:E20, 2003

23. Yano K, Zarain-Herzberg A: Sarcoplasmic reticulum calsequestrins: structural and functional properties. *Mol Cell Biochem* 135:61–70, 1994

24. Shin DW, Pan Z, Kim EK, Lee JM, Bhat MB, Parness J, Kim dH, Ma J: A retrograde signal from calsequestrin for the regulation of store-operated Ca2+ entry in skeletal muscle. *J Biol Chem* 278:3286–3292, 2003

25. MacLennan DH: Ca2+ signalling and muscle disease. *Eur J Biochem* 267:5291–5297, 2000

26. Howarth FC, Glover L, Culligan K, Qureshi MA, Ohlendieck K: Calsequestrin expression and calcium binding is increased in streptozotocin-induced diabetic rat skeletal muscle though not in cardiac muscle. *Pflugers Arch* 444:52–58, 2002

27. Hayashi T, Wojtaszewski JF, Goodyear LJ: Exercise regulation of glucose transport in skeletal muscle. *Am J Physiol* 273:E1039–E1051, 1997

28. Nolte LA, Rincon J, Wahlstrom EO, Craig BW, Zierath JR, Wallberg-Henriksson H: Hyperglycemia activates glucose transport in rat skeletal muscle via a Ca^{2+}-dependent mechanism. *Diabetes* 44:1345–1348, 1995

29. Dermitzakis ET, Reymond A, Scamuffa N, Ucla C, Kirkness E, Rossier C, Antonarakis SE: Evolutionary discrimination of mammalian conserved non-genic sequences (CNGs). *Science* 302:1033–1035, 2003

30. Samson SL, Wong NC: Role of Sp1 in insulin regulation of gene expression. *J Mol Endocrinol* 29:265–279, 2002