

Calpain 10: the first positional cloning of a gene for type 2 diabetes?

M. Alan Permutt, ... , Ernesto Bernal-Mizrachi, Hiroshi Inoue

J Clin Invest. 2000;**106**(7):819-821. <https://doi.org/10.1172/JCI11228>.

Commentary

Type 2 diabetes mellitus (T2DM) is not a single disease, but rather a genetically heterogeneous group of metabolic disorders sharing glucose intolerance. The precise biochemical defects in these conditions are unknown, but they almost certainly include impairments of both insulin secretion and insulin action. The familial nature, marked differences in prevalence among various racial groups, and different concordance rates between monozygotic and dizygotic twins indicate a genetic basis for the disease, but animal models and other data show that the disease is multifactorial. The identity and number of genes that are involved in this complex disorder are being clarified by recent genetic studies (see ref. 1 for recent review). Strategies for gene hunting Given that genetic factors contribute to the etiology of T2DM, how can the relevant genes be identified? One traditional method for gene discovery requires the cloning of candidate genes, which must then be screened for variants. In principle, a greater frequency of certain variant alleles in cases than suitably matched controls offers *prima facie* evidence that the candidate gene is involved in the disease. Considerable additional work is then required to confirm the association and to establish the biological relevance of the putative disease allele. Initially, the genes for insulin and the insulin receptor and, later, glucose transporter genes were considered as candidates for T2DM. Variants in [...]

Find the latest version:

<https://jci.me/11228/pdf>



Calpain 10: the first positional cloning of a gene for type 2 diabetes?

Commentary

See related article,
pages R69–R73.

M. Alan Permutt, Ernesto Bernal-Mizrachi, and Hiroshi Inoue

Washington University School of Medicine, Division of Endocrinology, Metabolism and Diabetes,
St. Louis, Missouri, USA

Address correspondence to: M. Alan Permutt, Washington University School of Medicine, Division of Endocrinology,
Metabolism and Diabetes, 660 S. Euclid Avenue, Campus Box 8127, St. Louis, Missouri 63110, USA.
Phone: (314) 362-8680; Fax: (314) 747-1309; E-mail: Apermutt@im.wustl.edu.

Type 2 diabetes mellitus (T2DM) is not a single disease, but rather a genetically heterogeneous group of metabolic disorders sharing glucose intolerance. The precise biochemical defects in these conditions are unknown, but they almost certainly include impairments of both insulin secretion and insulin action. The familial nature, marked differences in prevalence among various racial groups, and different concordance rates between monozygotic and dizygotic twins indicate a genetic basis for the disease, but animal models and other data show that the disease is multifactorial. The identity and number of genes that are involved in this complex disorder are being clarified by recent genetic studies (see ref. 1 for recent review).

Strategies for gene hunting

Given that genetic factors contribute to the etiology of T2DM, how can the relevant genes be identified? One traditional method for gene discovery requires the cloning of candidate genes, which must then be screened for variants. In principle, a greater frequency of certain variant alleles in cases than suitably matched controls offers *prima facie* evidence that the candidate gene is involved in the disease. Considerable additional work is then required to confirm the association and to establish the biological relevance of the putative disease allele. Initially, the genes for insulin and the insulin receptor and, later, glucose transporter genes were considered as candidates for T2DM. Variants in the gene for glucokinase, long known to be a rate-limiting enzyme in the control of insulin secretion, have been shown to result in a relatively rare type of diabetes, one form of maturity onset diabetes of the young (MODY). Mutations in the *Glucokinase* gene were not found in the more common forms of T2DM, however. Similar results for a number of candidate genes have now been almost universally

observed. Failure of candidate gene analysis in most complex diseases is likely due to low prior probability of the involvement of the gene in question in the overall risk of the disease (2).

Linkage analysis is another traditional method of gene discovery. This approach allows regions of chromosomal DNA to be identified that are shared among affected family members to an extent that would not be expected by chance alone. Thus, parents and/or siblings of affected individuals must be genotyped at a series of polymorphic markers, so that the degree of sharing of inherited alleles can be determined among affected and unaffected offspring. Markers genotyped in family members are commonly a region of polymorphic short nucleotide repeats that are called simple tandem repeats (STRs). To screen the whole genome without any prior knowledge of the gene defect, anonymous STRs with defined chromosomal locations are typed in family members. Genotyping 300–400 STRs in family members at roughly 10 centimorgan (cM) intervals has generally been adequate to define chromosomal regions sharing genes responsible for monogenic disorders.

For single gene disorders, recombinant mapping in families provides another powerful tool to refine the location of a disease locus detected by linkage analysis. In recombinant mapping, all family members are genotyped with relatively dense markers in the region of linkage, and recombinants observed in affected siblings narrows the critical region. The positional cloning of the gene for Wolfram syndrome (OMIM 222300), a rare autosomal recessive disorder defined by the combined presence of juvenile onset diabetes mellitus, optic atrophy, diabetes insipidus, and deafness, was accomplished by this method (3). Inspection of haplotypes in a small number of chromosome 4p-linked families revealed recombinants that allowed localization

of the gene to a critical region less than several hundred kilobases, ultimately leading to the identification of mutations in a novel gene, *WFS1*, in all affected individuals. The *MODY3* gene, *HNF1 α* , was identified by similar methods (4). Most of the advances in understanding the genetic basis of diabetes have been in rare conditions that are caused by mutations predominantly in a single gene. Genes for these disorders could be ascertained in a small number of families because they were the result of rare, single gene defects with strong effects. A minority of type 2 diabetic patients (2–5%) has diabetes that results from mutations in a single gene that results in either β cell dysfunction or, less frequently, insulin resistance. Mutations in five different genes, encoding the glycolytic enzyme glucokinase and the transcription factors *HNF-1 α* , *-4 α* , and *-1 β* and *PDX/IPF 1*, have been found to result in MODY (1).

Positional cloning of T2DM genes

The number of loci that harbor diabetes-predisposing genes, along with the estimated magnitude of their effects, have been assessed by genome scans in hundreds of families of various ethnic and racial groups (1). The first such study conducted a genome scan on 170 Mexican-American families and found linkage near the terminus of chromosome 2q37 (lod = 3.2, $P < 10^{-4}$) (5). This susceptibility locus, NIDDM1, originally encompassed perhaps 10–20 cM, but the target region was subsequently narrowed to 7 cM and shown to depend for its effect on the genotype at an unlinked locus, encompassing the *CYP19* gene on chromosome 15 (6). In work described in *Nature Genetics* (7) and in the current issue of the *JCI* (8), Graeme Bell's group follows up on the NIDDM1 narrative with a description of what may prove to be the first positionally cloned gene for a complex or polygenic disorder.

To positionally clone a gene for a complex disorder, one must rely on statistical probability by using a method known as linkage disequilibrium mapping. Recombinant mapping in families cannot be used to narrow the broad region defined by linkage, because, by definition, none of the genes for a complex trait is necessary or sufficient for development of the phenotype. In this case, Horikawa et al. (7) constructed a genomic contig encompassing most of the 7 cM NIDDM1 interval that is likely to include a predisposing gene for T2DM. The authors identified a total of 214 polymorphic sites across this region and tested for association with T2DM and for evidence of linkage. At NIDDM1, five single-nucleotide polymorphisms (SNPs) showed a difference for allelic frequencies between patients and controls, so Horikawa et al. sequenced approximately 60 kb in this region, revealing two genes and one expressed sequence tag. One gene, designated calpain 10 (*CAPN10*), encodes a new member of the cysteine protease family, and the other encodes G protein-coupled receptor 35 (*GPR35*). Both of these genes were shown to be ubiquitously expressed in tissues, including pancreatic islets and skeletal muscle. None of the coding region SNPs in *CAPN10* or in *GPR35* were associated with the disease or with the evidence for linkage. However, one SNP (UCSNP-43) in intron 3 of *CAPN10* showed an increase in the frequency of the common G allele in patients compared with controls, and this SNP was also associated with evidence for linkage. Because the G/G homozygotes at UCSNP-43 were not associated with the disease, the authors evaluated haplotypes across the *CAPN10* locus and found a pair of haplotypes (112 and 121), defined by specific polymorphisms at UCSNP-43 and two other sites, that was significantly associated with T2DM. Interestingly, homozygotes for either of these haplotypes were not at increased risk of disease, suggesting that the risk at *CAPN10* results from the interaction of two different risk factors, neither of which alone suffices to increase risk. In Finnish and German populations, the 112/121 genotype appears to be rare and is not significantly associated with increased risk for diabetes. Yet the 112 haplotype by itself was associated with increased risk of diabetes in Europeans, even though alone this haplotype was not associated with diabetes in Mexican-Americans.

Because of the intronic location of the putative disease mutation, it is of particular importance in this case to show that the sequence variation is biologically significant. In this issue of the *JCI*, Baier et al. (8) address the issue of the metabolic consequences of the *CAPN10* "at risk" genotype in Pima Indians. These subjects are well suited for analysis because of their high prevalence of T2DM and because they are the best characterized population for evaluating the metabolic consequences of the prediabetic state. The authors tested the hypothesis that genetic variation at *CAPN10* is associated with metabolic differences and that this variation might be the cause of these metabolic abnormalities. Appropriately, only people with normal glucose tolerance were tested. For purposes of this study the "at risk" genotype was said to be homozygosity for the UCSNP-43 G allele, although in the initial study of Mexican-Americans the G/G genotype was not a significant risk factor (odds ratio = 1.54, 95% confidence interval = 0.88–2.41) (7). The G allele is the common allele in Pimas, as in other ethnic and racial groups (G = 0.62, A = 0.38). The G/G genotype was associated with small but significant differences in fasting glucose, 2-hour insulin, endogenous glucose output, and glucose disposal and carbohydrate oxidation, relative to those individuals with the G/A or A/A genotypes. Data from respiratory chamber studies revealed lower sleeping metabolic rates and higher lipid oxidation and lower protein oxidation in the Pimas with the G/G genotype. The authors suggest that, in times of prolonged fasting, this phenotype could preserve skeletal muscle protein and glycogen, which might favor survival. In times of abundant food supply, the G/G phenotype might favor enhanced carbohydrate stores and secondary insulin resistance.

Functions of calpains

Although these studies clearly show that the *CAPN10* G/G genotype is associated with metabolic consequences in Pima Indians, it does not necessarily follow that genetic variation in the *CAPN10* gene causes these metabolic differences. The connection with glucose metabolism is still far from clear. Nevertheless, the biology of previously characterized calpains helps define some of the mechanistic questions that must now be addressed if *CAPN10* is to be considered a genuine diabetes gene.

The calpains represent a superfamily of structurally related proteins that are widely distributed in cells of various organisms, from *Caenorhabditis elegans* to humans. Many of these proteins have been shown to function as intracellular calcium-dependent cysteine proteases (9), and several members of the family have been cloned and characterized. These include ubiquitously expressed and tissue-specific isoforms, found in vertebrates as well as in some invertebrate homologues. Calpains may be implicated in a wide variety of cellular processes including apoptosis, ischemia, and inflammation, and they are known to be involved in the pathogenesis of Alzheimer's disease and limb-girdle muscular dystrophy type 2A, the latter being the first human disease that could be ascribed to a mutation in a calpain gene (*CAPN3*).

A functional calpain is composed of a heterodimer of a common S (small, regulatory, 30 kDa) subunit assembled to one of variable L (large, catalytic, 80 kDa) subunits. Calpains can be divided into at least four classes (A–D), based on their activity and their modes of regulation, and *CAPN10* shows significant homology to the recently identified mouse *CAPN8*, which belongs to the Class B calpains. While the amino acid sequence of domains I–III of *CAPN10* were conserved with other vertebrate calpains, it lacks the usual COOH-terminal calmodulin-like Ca^{2+} -binding domain IV, suggesting that *CAPN10* may not bind Ca^{2+} for activity. *CAPN10* mRNA was expressed in all human tissues examined, including muscle, liver, and islet (7). Eight splice variants have been identified, which may allow for tissue specificity in *CAPN10* activity.

Interestingly, *in vitro* experiments indicate that the intronic UCSNP-43 site may occur in a *cis*-acting element that influences *CAPN10* transcription (7). Gel shift experiments suggest that nuclear extracts can bind to the region around UCSNP-43, but the G/A polymorphism exerts no discernible effect on binding. However, Baier et al. inserted a 46-bp fragment of the UCSNP-43 region into a *luciferase* reporter construct and showed that the G allele has greater promoter activity than the A allele (8). The authors also demonstrate that the G/G genotype is associated with an altered phenotype in skeletal muscle biopsies of 18 male normoglycemic Pima subjects, where the level of *CAPN10* mRNA is 53% lower in

G/G subjects (8). Finally, they established a significant correlation between CAPN10 mRNA levels and the rate of carbohydrate oxidation.

What is the potential physiological role of CAPN10 in glucose metabolism and how could genetic variation in this gene contribute to the pathogenesis of T2DM? Unfortunately, the answer is not clear at this moment. Studies of CAPN10 in other racial groups reported at the 60th Annual Session of the American Diabetes Association leave the role of the UCSNP-43 G allele in question. Among 153 United Kingdom Caucasian parent-offspring T2DM trios, there was no transmission disequilibrium distortion for the UCSNP-43 polymorphism, although there was disequilibrium for transmission of an extended haplotype around the region. Another study of European diabetics found excess transmission of the A allele for UCSNP-43 (10). A third study reported increased frequency of the UCSNP-43 G allele in diabetics compared with controls and showed that subjects with the G allele had higher fasting insulin and were more insulin resistant (11).

Future in vitro and in vivo studies will be needed to refine CAPN10's roles in glucose and lipid metabolism, fat cell differentiation, and insulin signaling. Since CAPN10 is also expressed in pancreatic islets, it is possible that unregulated calpain proteolysis participates in β cell damage — a potentially unifying mechanistic event. Polonsky and colleagues (12, 13) have demonstrated in vitro that nonspecific calpain inhibitors ALLM (100 mM) and E-64-d (200 mM) induce alterations of glucose-induced

insulin secretion in mouse islets. Calpain inhibitors reduced insulin-mediated glucose transport in isolated rat muscle strips and adipocytes. Whether this effect is by inhibition of CAPN10 or other members of this family of proteases remains to be determined.

Final lessons

Stepping back from these mechanistic questions, the experiences of Graeme Bell and his associates serve to illustrate the enormous difficulties involved in positional cloning of genes for complex diseases. While lod scores at 2q did reach values considered significant in a genome-wide search, these linkage results have not been replicated to this extent, either in Mexican-Americans or in any other racial group. Does the NIDDM1 locus really harbor a diabetes susceptibility gene? If so, how does one proceed from a broad region of linkage to more narrow regions for dense association studies? Which subjects are appropriate for association studies, and how many should be genotyped? How can we achieve accurate, efficient, low cost SNP genotyping, and at what density should they be genotyped? Most importantly, how do we interpret the results?

The provocative results reported by Baier et al. (8) represent an initial step toward defining the metabolic and clinical phenotype associated with genetic variation in CAPN10. Most intriguing are the possible metabolic consequences of variation at this locus on tissues involved in insulin production and action. These questions must be explored in patients and in transgenic animals and will occupy investigators for some time to come.

Bell and associates' important results have taught us a great deal, but, as might have been predicted, many more questions remain for this gene and for the numerous other genes that contribute to diabetes susceptibility.

1. Permutt, M.A., and Hattersley, A. 2000. Searching for type 2 diabetes genes in the post genome era. *Trends Endocrinol. Metab.* **11**:383–393.
2. Risch, N. 2000. Searching for genetic determinants in the new millennium. *Nature.* **405**:847–856.
3. Inoue, H., et al. 1998. A gene encoding a transmembrane protein is mutated in patients with diabetes mellitus and optic atrophy (Wolfram syndrome). *Nat. Genet.* **20**:143–148.
4. Kaisaki, P.J., et al. 1997. Mutations in the hepatocyte nuclear factor-1 α gene in MODY and early-onset NIDDM: evidence for a mutational hot spot in exon. *Diabetes.* **46**:528–535.
5. Hanis, C.L., et al. 1996. A genome-wide search for human non-insulin dependent (type 2) diabetes genes reveals a major susceptibility locus on chromosome. *Nat. Genet.* **13**:161–166.
6. Cox, N.J., et al. 1999. Loci on chromosomes 2 (NIDDM1) and 15 interact to increase susceptibility to diabetes in Mexican Americans. *Nat. Genet.* **21**:213–215.
7. Horikawa, Y., et al. 2000. Genetic variation in the calpain 10 gene (CAPN10) is associated with type 2 diabetes mellitus. *Nat. Genet.* In press.
8. Baier, L., et al. 2000. A calpain-10 gene polymorphism is associated with reduced muscle mRNA levels and insulin resistance. *J. Clin. Invest.* **106**:R69–R73.
9. Sorimachi, H., Ishiura, S., and Suzuki, K. 1997. Structure and physiological function of calpains. *Biochem. J.* **328**:721–732.
10. Ren, Q., Hasstedt, S.J., Hanis, C., and Elbein, S. 2000. Increased transmission of NIDDM2 variant in Caucasian familial type 2 diabetes. *Diabetes.* **49**:A200. (Abstr.)
11. Orho-Melander, M., et al. 2000. Genetic variation in the calpain 10 gene is associated with insulin resistance and type 2 diabetes in Finland. *Diabetologia.* In Press. (Abstr.)
12. Zhou, Y.-P., et al. 2000. Calpain inhibitors impair insulin secretion after 48-hours: a model for β -cell dysfunction in type 2 diabetes? *Diabetes.* **49**:A80. (Abstr.)
13. Seamus, K., et al. 2000. Calpain-sensitive pathways in insulin secretion and action: a pathophysiological basis for type 2 diabetes? *Diabetes.* **49**:A62. (Abstr.)