Human Molecular Genetics, 2006, Vol. 15, No. 18 ***2709–2720***
doi:10.1093/hmg/ddl204
Advance Access published on August 7, 2006

TBC1D1 is a candidate for a severe obesity gene and evidence for a gene/gene interaction in obesity predisposition

Steven Stone[1,*], Victor Abkevich[1], Deanna L. Russell[1,2], Robyn Riley[1], Kirsten Timms[1], Thanh Tran[1], Deborah Trem[1,3], David Frank[1], Srikanth Jammulapati[1], Chris D. Neff[1], Diana Iliev[1], Richard Gress[4], Gongping He[1], Georges C. Frech[1], Ted D. Adams[4], Mark H. Skolnick[1], Jerry S. Lanchbury[1], Alexander Gutin[1], Steven C. Hunt[4] and Donna Shattuck[1]

[1]Myriad Genetics, Inc., Salt Lake City, 320 Wakara Way, UT 84108, USA, [2]Department of Psychiatry, Genome Research Institute, University of Cincinnati, OH 45237, USA, [3]Cardiovascular Research and Training Institute, University of Utah, Salt Lake City, UT 84112, USA and [4]Cardiovascular Genetics, Department of Medicine, University of Utah, Salt Lake City, UT 84108, USA

Received June 8, 2006; Revised and Accepted July 28, 2006

The molecular etiology of obesity predisposition is largely unknown. Here, we present evidence that genetic variation in TBC1D1 confers risk for severe obesity in females. We identified a coding variant (R125W) in TBC1D1 that segregated with the disease in 4p15–14-linked obesity pedigrees. In cases derived from pedigrees with the strongest linkage evidence, the variant was significantly associated with obesity ($P = 0.000007$) and chromosomes carrying R125W accounted for the majority of the evidence that originally linked 4p15–14 with the disease. In addition, by selecting families that segregated R125W with obesity, we were able to generate highly significant linkage evidence for an obesity predisposition locus at 4q34–35. This result provides additional and confirming evidence that R125W affects obesity susceptibility, delimits the location of an obesity gene at 4q34–35 and identifies a gene/gene interaction that influences the risk for obesity predisposition. Finally, although the function of TBC1D1 is unknown, the protein is structurally similar to a known regulator of insulin-mediated Glut4 translocation.

INTRODUCTION

The incidence of obesity has reached epidemic levels in the developed world (1,2). If unchecked, the rising rates of obesity (and its associated morbidities) pose a significant threat to human health and a costly burden to public healthcare systems. Unfortunately, patients with severe obesity [usually defined as having a body-mass index (BMI) of over 35] have very few medical options. In fact, with the exception of surgical intervention, there are no effective long-term therapies (3). Our goal is to define the genetic etiology of obesity predisposition to enable the development of effective therapeutic interventions.

Obesity is a complex genetic disease. Although it is clearly heritable (4–6) (e.g. first-degree relatives of obese patients have up to a 9-fold increase in disease risk (7)), the segregation of the disease in obesity-prone families is non-Mendelian (8). Also, many studies have provided evidence to suggest that risk for obesity is determined not only by specific genotypes but also by significant gene/gene and gene/environment interactions (9–12).

There have been many attempts to identify genes involved in the predisposition to common obesity, but presumably because of the genetic complexity, these efforts have had only mixed success (6,13). Studies focused on rare forms of obesity have been more productive. Mutations in several genes have been linked to monogenic obesity. Also, several genes have been implicated in syndromic obesity (i.e. obesity with other clinically significant phenotypes not

*To whom correspondence should be addressed. Tel: +1 8015843695; Fax: +1 8015843640; Email: sstone@myriad.com

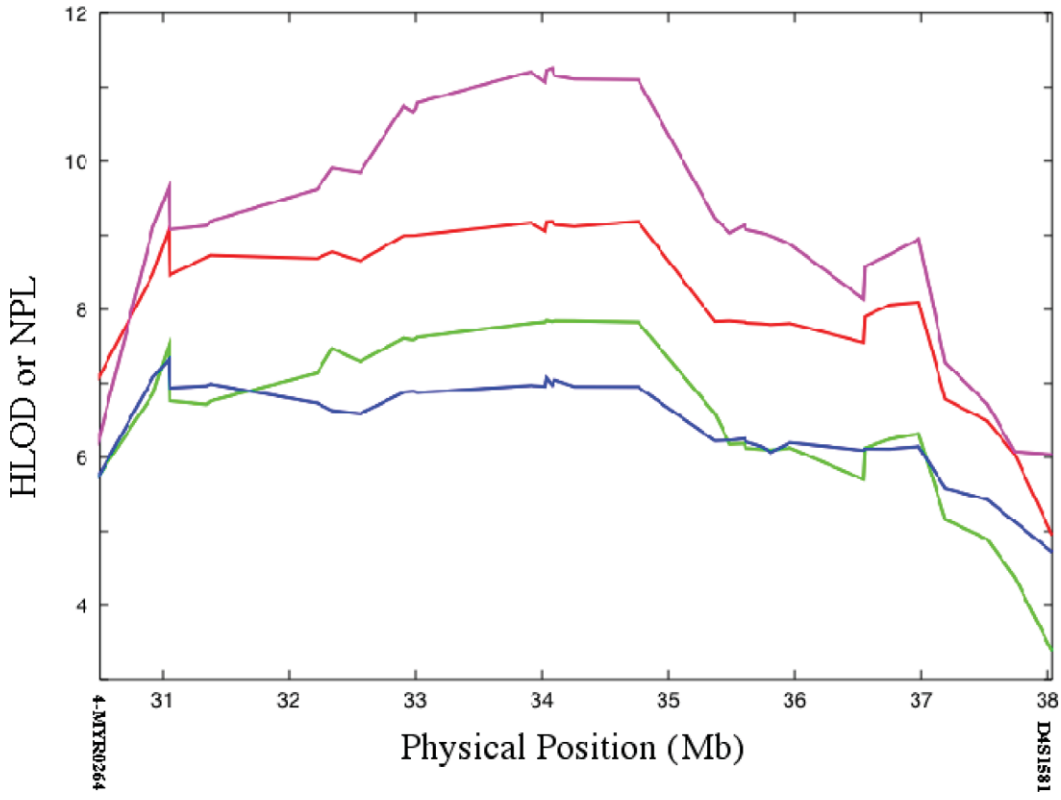

Figure 1. HLODs used to define linkage interval. The region shown is from 4-MYR0264 to D4S1581. The colored lines correspond to the results from four different linkage analyses: purple, NPL; red, codominant model; green, dominant model; blue, recessive model. The phenotype is females-only BMI (14). HLOD or NPL score is plotted on the *y*-axis and physical position (in Mb) on the *x*-axis. For all relevant figures, the physical position is based on the NCBI genome assembly Build 35.

usually associated with the disease). However, genetic causes of the common form of this disease remain poorly defined.

We previously identified a linkage to severe obesity in females at 4p15–14 (14). In that study, we increased the likelihood of finding obesity-causing genes by selecting pedigrees with multiple, severely affected members. As a result, we obtained highly significant linkage evidence to severe obesity, with heterogeneity LOD (HLOD) score over 9. Although our analysis was done in Caucasian pedigrees, the same interval (4p15–14) was subsequently linked to high BMI values in a Mexican-American cohort (15).

Here, we present evidence that variation in TBC1D1 (tre-2/USP6, BUB2, cdc16 domain family member 1) underlies the 4p15–14 obesity locus. Interestingly, the evidence presented suggests that TBC1D1 confers obesity risk only in combination with disease-promoting variation elsewhere in the genome. Finally, we show that TBC1D1 is similar structurally to AS160 (Akt substrate of 160 kDa), a known regulator of the insulin-signaling pathway.

RESULTS

Defining the genomic interval for selection of candidate genes

Typically, the exact position of a disease gene cannot be determined from the linkage analysis. However, linkage data can provide an approximate location, which can be used to define the genomic interval that is most likely to contain the gene of interest (the linkage interval).

Because obesity is a common disease, some of the affected females in the 4p15–14-linked pedigrees are likely to be sporadic cases (i.e. affected due to any factor other than the presumptive locus at 4p15–14). This made it difficult to construct a reliable recombinant map to delimit the linkage interval. Therefore, we defined the linkage interval on the basis of a straightforward analysis of the LOD scores. Central to this approach is the assumption that the position of the highest HLOD score approximates the position of the disease gene. However, for a complex disorder, this supposition may not be true. Studies have shown that errors in both phenotypes and linkage models can shift the location of the highest HLOD score away from the disease gene (16). To compensate, we defined the boundaries of our linkage interval based on a larger than typical drop in HLOD. Where possible, we used at least a two-unit reduction of HLOD score and defined an interval that was supported by both parametric and non-parametric linkage analyses. To satisfy as many of these requirements as reasonably possible, we selected the interval from 4-MYR0264 (an STR marker developed for this study at position 30 502 000 bp) (14) to D4S1581 for subsequent selection of candidate genes (Fig. 1). This region spans ~7.5 Mb according to the most recent assembly of the human genome (Build 35, May 2004).

Table 1. Pedigrees with LOD scores ≥1.0 in 4p15–14

Family	Highest dominant LOD score in linkage interval	Disease-associated haplotypes		Number of affected females (BMI ≥ 35) with R(125)W/total number of affected females in pedigree
		Number of haplotypes	Haplotype(s) contain R(125)W allele of TBC1D1	
43601	3.4	2	Yes (2 of 2)	15/16
604401	2.0	1	No	—
7380	1.4	1	Yes	6/7
7444	1.4	1	No	—
708201	1.3	1	Yes	5/5
722801	1.2	1	Yes	9/12
72561	1.0	1	Yes	6/7
11135	1.0	1	No	—
7135101	1.0	1	Yes	6/10

Resequencing candidate genes

To identify the obesity gene at 4p15–14, we first identified all positional candidate genes between 4-MYR0264 and D4S1581 by comparing the genomic sequence derived from this interval with expressed sequence in the public domain. A representative cDNA clone defined a candidate gene if the clone was derived from the interval and contained a potential open-reading frame of at least 60 amino acids. In total, we identified 16 candidate genes between 4-MYR0264 and D4S1581. For comparison, the region contains six reference genes.

Next, we searched for genetic variation that segregated with obesity by resequencing each candidate gene in obese individuals that carried one of the 10 disease-associated haplotypes (Table 1). The disease-associated haplotypes were identified from the nine families that generated a LOD ≥ 1 in the linkage interval under a dominant genetic model (see Materials and Methods and Table 1 for details). Importantly, each disease-associated haplotype generated significant statistical evidence that it cosegregated with obesity. Although any variant on a disease-associated haplotype could potentially cause obesity, we focused on variants likely to have an effect on gene function. Therefore, we resequenced only exons and intronic sequence within 50 bp of an intron/exon junction.

As a result of resequencing candidate genes, we identified 68 variable positions within the linkage interval. By resequencing two subjects from each selected pedigree, we were able to make a putative determination of which variants were on a disease-associated haplotype (see example in Fig. 2). If a variant occurred in multiple families, we compared the number of times it was assigned to a disease-associated haplotype with the total number of observations. We arbitrarily classified the variant as segregating with obesity if it was on disease-associated haplotypes at least 70% of the time. It is likely that variants (and especially common variants) were over-assigned to disease-associated haplotypes, but the bias was intended to maximize our ability to identify variants that potentially segregated with obesity and warranted further investigation.

By these criteria, 16 of the 68 variants were selected as possibly segregating with obesity. The segregating variants were candidates for disease-causing variation and were evaluated further by comparing their frequency within the 10 disease-associated haplotypes (always assuming the variant was on the disease-associated haplotype) with their frequency in 90 unrelated Utah CEPH samples (population controls). On the basis of this comparison, we identified three variants that were significantly enriched ($P < 0.05$) on disease-associated haplotypes. The first variant is located in the candidate gene represented by GenBank clone BG181206 ($P = 0.04$). The variant is located in an intron, 38 nucleotides upstream of the intron/exon junction. The second variant is located in the candidate gene represented by GenBank clone BX099393 ($P = 0.02$). Again, this variant is in an intron, 40 nucleotides

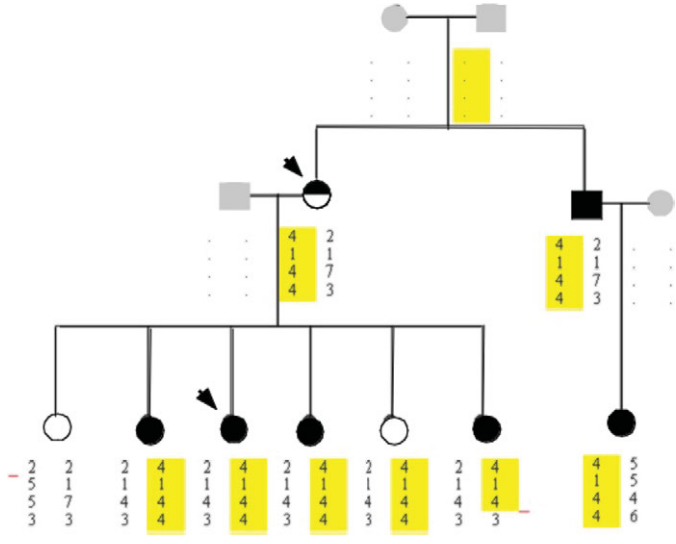

Figure 2. Fragment of 4p15–14-linked pedigree that segregates a disease-associated haplotype. The haplotype marked in yellow was selected as disease-associated and, irrespective of the drawing, could have originated from either founder. The subjects marked with an arrow (→) were resequenced and share a disease-associated haplotype. Therefore, variants carried by both subjects were assigned to the disease-associated haplotype. Black symbol, BMI ≥ 35; half-black symbol, BMI 30–35; open symbol, BMI ≤ 30; gray symbol, unknown BMI (these subjects did not participate in the study). The five-marker haplotype is derived from D4S2995, D4S3040, D4S1581, D4S2382 and D4S405. Numbers correspond to specific alleles and dots depict missing genotypes. The red lines indicate recombinant break-points.

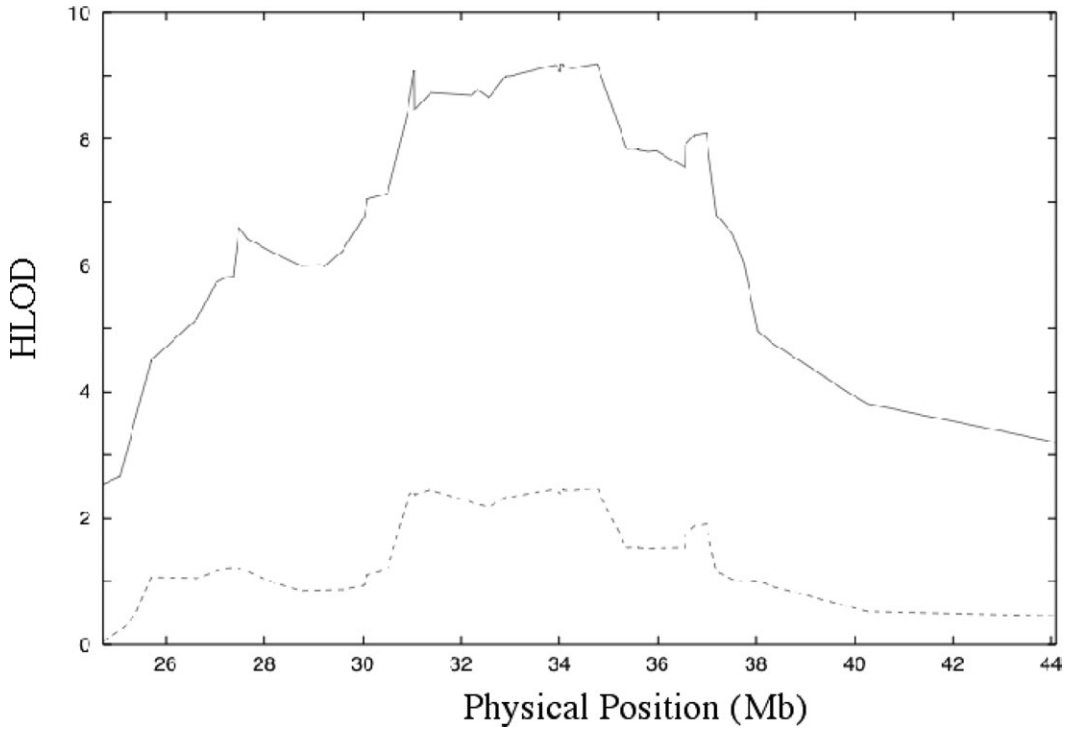

Figure 3. Evidence for an obesity linkage to 4p15–14 after female carriers of R125W were excluded from the linkage analysis. The solid line plots the HLOD for 4p15–14 using a female-only BMI phenotype and codominant inheritance model as previously described (14). The dashed line is the HLOD from the same analysis, except all the females with R125W were given an unknown phenotype. HLOD is plotted on the *y*-axis and physical position (in Mb) on the *x*-axis.

upstream of the intron/exon junction. Both BX099393 and BG181206 are derived from the spliced ESTs, but are otherwise uncharacterized. Neither gene contains a compelling ORF (68 and 71 amino acids, respectively), nor are they similar to any other gene sequence in the public domain (excluding a small segment of repetitive DNA in BX099393). The third variant is located in the candidate gene TBC1D1 ($P = 0.000008$). The variant is at position 125 in the full-length peptide and results in the substitution of tryptophan for arginine (R125W).

In the linkage interval, we identified 68 variable positions. Therefore, the *P*-values described earlier should be adjusted for 68 tests. Importantly, the *P*-value for R125W remained significant (adjusted $P = 0.00054$). The *P*-values for BX099393 and BG181206 were not significant after adjustment.

Genetic evidence linking TBC1D1 with obesity predisposition

On the basis of the results from resequencing candidate genes, we focused our analysis on TBC1D1 as a candidate for the obesity predisposition gene at 4p15–14. In particular, we identified the missense variant, R125W, as potentially responsible for predisposition to obesity. Our initial observation was that the R125W variant segregated with affected females in seven out of 10 of the disease-associated haplotypes (Table 1). The assignment of R125W to the disease-associated haplotypes was confirmed by resequencing additional subjects in the linked pedigrees. For example, in family 43601, 15 of the 16 female members with a BMI $\geq$35 carried the W allele of TBC1D1. In general, when R125W was observed in a disease pedigree, it segregated with obesity. The exception was family 7135101. In this pedigree, only six of 10 affected females carried the W allele, and as a result, although the pedigree generates a LOD score of over 1.0 in the linkage interval, at TBC1D1, the LOD score dropped to 0.5.

Two additional results strengthened the link between TBC1D1 and obesity predisposition. First, when the number of control haplotypes in the comparison was increased, the enrichment of R125W in the disease-associated haplotypes remained statistically significant. Initially, we compared the frequency of R125W in the disease-associated haplotypes with 90 unrelated CEPH samples. We increased the control set to include an additional 333 unrelated subjects (666 haplotypes) derived from the Utah population. When compared with this final set of population controls, R125W remained significantly enriched in disease-associated haplotypes with a *P*-value of 0.000007 (7/10 versus 77/846).

Second, R125W-containing haplotypes generated the majority of the evidence that originally linked obesity to 4p15–14. This is consistent with the view that R125W contributes to obesity predisposition. To evaluate this, we genotyped our entire pedigree set at position 125 in TBC1D1. HLOD scores were recalculated excluding female carriers of R125W from the analysis (i.e. their phenotype was changed to unknown). The HLOD scores in the 4p15–14 region were dramatically affected (Fig. 3). The change in phenotype to unknown only affected 24% of the females with BMIs >35, but the HLOD for the codominant model dropped from over

Table 2. Summary of pedigrees used in this study

Pedigree set (*N*)	Analysis	Result
Pedigrees with clusters of affected females (37); see Stone *et al*. (14)	Linkage	4p15–14 linkage evidence
4p15–14-linked pedigrees (9)	Association	R125W associated with obesity
4p15–14-linked pedigrees that segregate R125W with obesity (6)	Linkage (two-generation segments)	4q34–35 linkage evidence
Independent pedigrees with clusters of affected subjects that also segregate R125W with obesity (18)	Linkage (two-generation segments)	Replication of 4q34–35 linkage evidence

9.0 to 2.8. The remaining linkage evidence (HLOD of 2.8), and the fact that not all of the linked pedigrees segregated R125W (Table 1), may indicate that there are additional disease-causing variants in TBC1D1 that are yet to be discovered. If true, these variants most likely reside in genomic regions that were not resequenced (e.g. non-coding regulatory elements).

The above evidence that implicates TBC1D1 in obesity predisposition derives from an analysis of linked pedigrees. In positional cloning projects, researchers may attempt to confirm gene discoveries by looking for evidence of disease association in unrelated cases. However, those studies may lack sufficient power to detect disease associations in the context of complex genetics. This is particularly true if non-additive gene/gene interactions are an important factor in determining disease risk. Nevertheless, we attempted to show that the R125W allele was enriched in random obesity cases. We genotyped position 125 in TBC1D1 in 173 unrelated affected females (BMI $\geq$ 35) derived from Utah but not included in our linkage analysis. However, R125W was not enriched in this set of random cases (28/346 versus 77/846 in population controls, $P = 0.33$).

Evidence for gene/gene interactions in obesity predisposition

Given the negative result of our association study with R125W when cases were not derived from the linked pedigrees, we considered the possibility that R125W conferred risk for obesity only in the context of specific gene/gene interaction(s). We reasoned that it might be possible to detect these interacting loci and, in so doing, provide independent genetic evidence that TBC1D1 is involved in obesity.

To identify interacting loci, we used the six families that segregated R125W with obesity (43601, 7380, 708201, 722801, 72561 and 7135101) in a subsequent genome search for obesity predisposition genes (see Table 2 for a summary of the pedigrees used for each stage in this study). In these pedigrees, the majority of affected females carry the W allele (Table 1). As we were trying to detect interacting loci, we employed the same genetic analysis that led to the discovery of TBC1D1. Specifically, we used a female-only BMI phenotype, and a genetic model for which BMI values were used to determine the likelihood that a subject carried a disease allele (14). In addition, as interacting alleles were likely to be common (given the strong 4p15–14 linkage evidence), multigenerational families were divided into two-generation segments. The small size of each pedigree reduced the probability of intrafamilial genetic heterogeneity. The resultant HLOD scores were biased at 4p15–14 (as pedigrees were selected on the basis of genotype at that location) but were unbiased in the rest of the genome. The results of the genome search are shown in Fig. 4. In addition to the expected linkage at 4p15–14, we saw an HLOD of 2.8 at 4q34–35. These regions are genetically unlinked. As we completed four genome searches for interacting loci, an HLOD of 2.8 is only suggestive evidence for linkage. However, it should be noted that the pedigree set was very small and had limited power to detect genetic linkage.

To further investigate the 4q linkage, we developed a dense STR marker map across a 32 cM interval around the peak HLOD score from the genomic search. In addition, we ascertained 18 additional obesity families that segregated R125W with obesity. The families were selected according to the rule that each family had to contain at least two female siblings with BMIs $\geq$ 35, and these affected females had to carry the W allele of TBC1D1. These families were not included in our previous analysis. With 19 additional markers, the 24 families (18 plus the six in the genome search) were analyzed for linkage at 4q34–35 with the same models and phenotype as used in the genome search. As a result, the total HLOD in the interval increased to 5.1 (Fig. 5). The 2-LOD linkage interval is from D4S1529 to D4S3051. Without conditioning the analysis on R125W, the linkage evidence at 4q34–35 is only suggestive (HLOD of 1.7, data not shown). Interestingly, other groups also have observed suggestive linkage evidence at 4q34 for obesity-related phenotypes, including high BMI (17,18).

The discovery of a gene/gene interaction was not a prespecified condition for identifying TBC1D1 as an obesity gene. Therefore, it was critical that we replicate the evidence for the interaction in an independent set of families. The 18 families that were selected solely because they segregated R125W with obesity satisfy this requirement. When we included only these families in our linkage analysis, they generated an HLOD of 2.8 in the 4q34–25 interval (Fig. 5). This result was obtained without the multiple testing inherent in a genome search and therefore provided strong support for the presence of an interacting locus at 4q34–35 (HLOD $P = 0.0004$).

R125W is in a PTB domain

There is a good evidence to suggest that R125W might affect the TBC1D1 function. The change results in a non-conservative amino acid substitution. Also, the variant occurs in a phosphotyrosine-binding (PTB) domain (Fig. 6). PTB domains in other proteins have been extensively

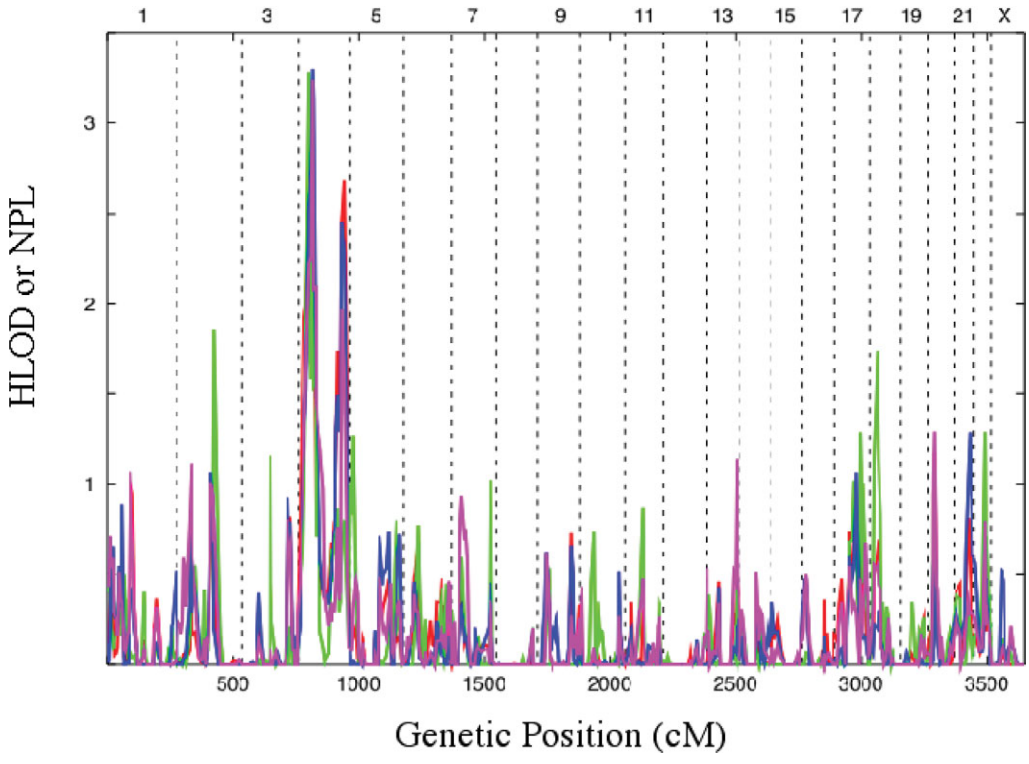

Figure 4. Results of the genome search for obesity loci using only the six 4p15–14-linked families (LOD ≥ 1) that also segregated R125W with obesity. The colored lines correspond to the results from four different linkage analyses: purple, NPL; red, codominant model; green, dominant model; blue, recessive model. The phenotype is females-only BMI (14). HLOD or NPL score is plotted on the *y*-axis and marker positions (by cM position) are plotted on the *x*-axis. Vertical dashed lines delimit the chromosomes. Chromosome numbers (odd only) are indicated across the top of the figure.

studied, including using site-directed mutagenesis to investigate structure–function relationships within the domain (19–22). Some of these studies have shown that altering an R at a position corresponding to R125 in TBC1D1 affected the binding properties of the PTB domain (19,21). In addition, on the basis of sequence in the public domain, R125 has been conserved as either R or Q (glutamine) throughout mammalian evolution (data not shown).

Analysis of genetic disequilibrium between R125W and surrounding SNPs

Although there are strong arguments to suggest that R125W may affect gene function and cause disease, it remained possible that R125W was not the causal variant. Rather, R125W may be in linkage disequilibrium (LD) with another variant that actually causes obesity.

If R125W was not causal, then the bona fide disease variant should be shared by a majority of the 4p15–14-linked pedigrees that contained R125W. To identify candidate variants, we resequenced the genomic region around R125W that was shared in common by the six R125W-containing families from Table 1. The shared SNP haplotype spanned 9.7 kb of genomic DNA and contained 17 variable positions (Fig. 7). On the basis of the genetic criteria described earlier, we did not find any SNPs that were more likely than R125W to cause obesity. We did find four variants that were in complete disequilibrium with R125W. But in contrast to R125W, these four variants are unlikely to affect gene function. They are in intronic regions that were not conserved through evolution, and they are far removed from intron/exon junctions.

Of course, the possibility remains that R125W is in long-range disequilibrium with a variant elsewhere in genome. However, on the basis of the most recent version of the HapMap (23), there is no evidence of extended LD in the region that contains R125W, and in fact, there is little evidence for extensive LD in the entire TBC1D1 region (data not shown).

Analysis of the TBC1D1 protein sequence

TBC1D1 has been characterized as being differentially expressed during mast cell differentiation (24). Nothing else is known about the biological function of TBC1D1. However, functional clues can be derived from an analysis of primary protein sequence of TBC1D1. TBC1D1 contains two conserved protein domains (Fig. 6). As mentioned earlier, the protein contains a PTB domain at its N-terminus. PTB domains are typically involved in protein/protein interactions that are often, but not always, dependent on the phosphorylation status of tyrosine residues on the binding partner (20). At the C-terminus, the protein contains a TBC (Tre-2/BUB2/Cdc16) domain. TBC domains are assumed to have Rab GAP activity and therefore possibly regulate Rab family small molecular weight GTPases (25,26). The protein also contains smaller conserved sequence elements.

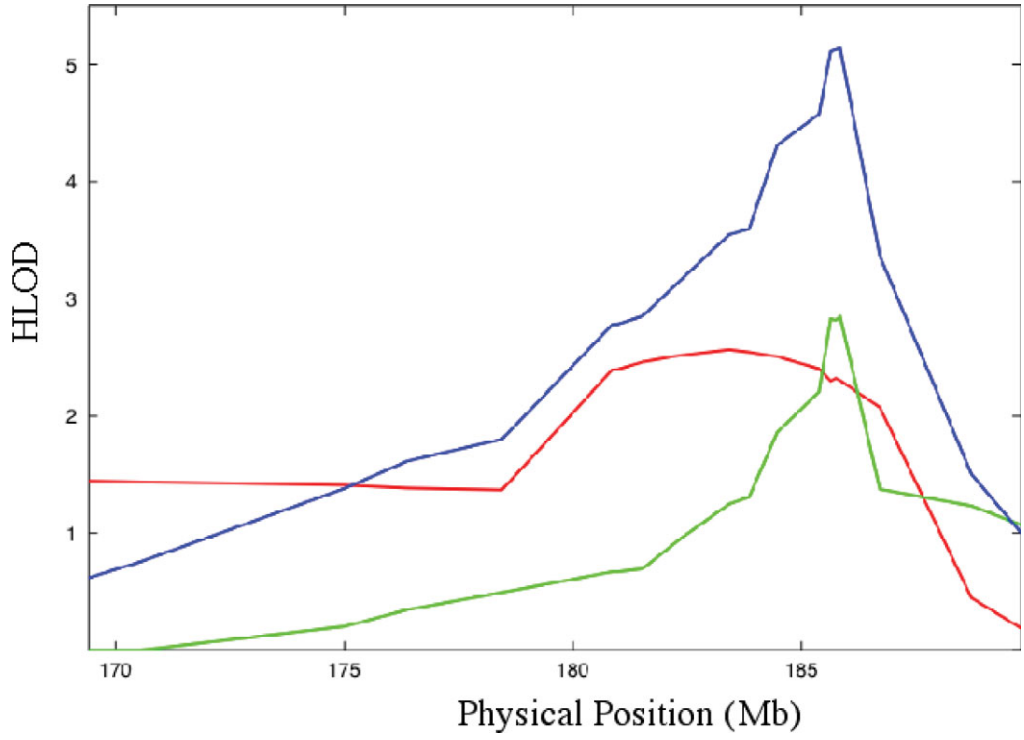

Figure 5. Results of adding additional markers and families to the linkage analysis of the 4q34–35 interval. The analysis employed a female-only BMI phenotype and codominant inheritance model as previously described (14). The colored lines correspond to different pedigree sets: red, six 4p15–14-linked pedigrees (LOD $\geq$ 1) that also segregated R125W with obesity; green, 18 additional pedigrees that were selected because they segregated R125W with obesity; blue, all 24 families analyzed together. HLOD is plotted on the *y*-axis and physical position (in Mb) is plotted on the *x*-axis.

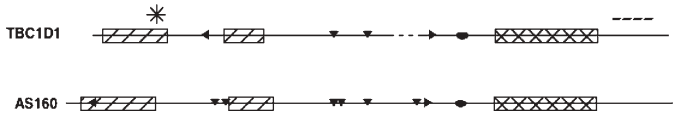

Figure 6. Schematic alignment of TBCD1 and AS160. The highlighted features include: PTB domains (▨, at 21–159 and 281–387); R125W (✳); a potential AMPK phosphorylation site (◀, at 232–241); potential AKT phosphorylation sites (▼, at 502–507 and 591–506); a potential PKA phosphorylation site (▶, at 657–661); a calmodulin-binding site (⬬, at 715–738); a TBC domain (▨, at 797–1017). All positions are the amino acid coordinates for TBCD1. The broken line in TBC1D1 represents a gap that was introduced to improve the alignment. The dashed line above TBC1D1 indicates the approximate position of the probe used for northern analysis.

Specifically, we have identified potential phosphorylation sites for AKT, 5′AMP-activated kinase (AMPK) and protein kinase A (PKA). Finally, the peptide contains a potential calmodulin-binding site (27). The order and number of conserved protein domains in TBC1D1 have been conserved in AS160 (28), a closely related paralog to TBC1D1 on chromosome 13 (Fig. 6). Overall, the AS160 and TBC1D1 proteins are 50% identical.

Tissue-specific expression of TBC1D1

We examined the tissue distribution of the TBC1D1 transcript by northern analysis. Northern blots containing RNA from various tissues were probed with a portion of the 3′ coding sequence of the TBC1D1 mRNA (Fig. 6). TBC1D1 is expressed predominantly in muscular tissues (e.g. heart, skeletal muscle and uterus) (Fig. 8). We also investigated TBC1D1 expression by qualitative RT–PCR (TaqMan® assay). In general, these results supported the data from northern blots, but also indicated relatively high levels of TBC1D1 mRNA in fat and lung tissue (data not shown). Some of the tissues with high levels of TBC1D1 expression (specifically muscle and fat) are very important targets for insulin signaling, and the tissue specificity closely matches the expression pattern of AS160 (29).

DISCUSSION

The genetic evidence presented in this paper implicates the R125W variant of TBC1D1 in obesity predisposition. In summary, we found that the W allele of TBC1D1: (i) segregates with affected females in a majority of the 4p15–14-linked obesity-prone pedigrees; (ii) is enriched on disease-associated haplotypes ($P = 0.000007$) and (iii) accounts for the majority of the linkage evidence at 4p15–14.

In addition, we found that linkage analysis conditioned on R125W generated significant evidence for an obesity predisposition locus at 4q34–35. This result suggests that R125W interacts with a gene at 4q34–35 in determining the risk for obesity and therefore leads to three important results. First, it identifies the location of a second obesity predisposition gene. Second, it indicates that gene/gene interactions are an important determinant of obesity predisposition. This is not

Kindred	rs10010366*	rs2995922	C to A	C to T	A to G	G to T	rs2279028	rs2279027*	rs2279026*	C toT(R125W)	rs2995921	C to T	rs13108308	rs2995917*	rs2925847*	rs2925848*	rs11726078*
436(1)	G	G	A	T	G	T	G	C	C	T	G	C	C	T	A	G	T
436(2)	G	C	A	T	G	T	G	C	C	T	G	T	G	C	G	A	T
7380	G	G	A	T	G	T	G	C	C	T	G	C	C	T	A	G	C
708201	G	G	A	T	G	T	G	C	C	T	G	C	C	T	A	G	T
722801	G	G	A	T	G	T	G	C	C	T	G	C	C	T	A	G	T
72561	A	G	A	T	G	T	G	C	C	T	G	C	C	T	A	G	T
7135101	G	G	A	T	G	T	G	C	C	T	G	C	C	T	A	G	T
position (bp)	37,723,388	37,724,523	37,724,594	37,725,231	37,725,304	37,725,641	37,726,235	37,726,322	37,726,648	37,726,655	37,726,868	37,729,122	37,730,546	37,730,777	37,732,061	37,732,407	37,733,072

Figure 7. SNP-based haplotype shared by the 4p15–14-linked families that segregated R125W with obesity. The SNP alleles shaded in yellow are on the shared haplotype. The boundaries of the sequenced region were defined by identifying two ancient recombination break-points on each end of the interval. Family 43601 contained two R125W haplotypes and both are indicated. The SNPs colored red were in complete disequilibrium with R125W. SNPs marked with an asterisk were analyzed by the HapMap consortium and those without rs numbers are not in the public domain.

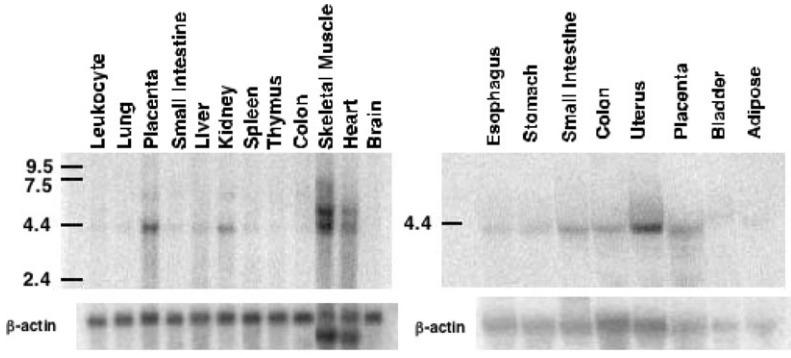

Figure 8. Expression of TBC1D1 mRNA in various human tissues. The northern blots were probed with [α-^{32}P]-labeled cDNA fragment corresponding to the 3′ end of the TBC1D1 coding region. The blots were also hybridized with a probe for β-actin to control for amount of mRNA in each lane (lower panel). Muscle and heart tissue contain a smaller isoform of β-actin that is not used for tissue-to-tissue comparison.

an unexpected result. In fact, many investigators have argued that gene/gene interactions are likely to be common in complex genetic diseases including obesity (9–11,30,31). Our results provide evidence that gene/gene interactions can play a role in disease predisposition. Third, the conditional linkage analysis dramatically strengthens the case that TBC1D1 is an obesity predisposition gene. The other genetic arguments presented in this paper (segregation, enrichment, etc.) are based on observations of R125W in linked pedigrees, and generated the hypothesis that TBC1D1 was an obesity predisposition gene. In contrast, the conditioned linkage was performed after the disease candidacy of TBC1D1 was established. The conditioned linkage is independent of our original observations and therefore strongly supports a role for TBC1D1 in obesity predisposition. Of course, a replication of this specific hypothesis (i.e. the epistatic relationship between R125W and the gene at 4q34–35) in a different population would significantly strengthen our conclusions.

Association studies are difficult to interpret in the context of complex genetic interactions (32). Our evidence that TBC1D1 interacts with a locus at 4q34–35 exemplifies the complexity underlying obesity predisposition. In this case, R125W confers risk for obesity only in the context of a specific gene/gene interaction. As a result, our association study had limited power to detect a disease association. In fact, we were unable to show that the W allele of TBC1D1 is enriched in random severely obese cases. In contrast, the families in our original linkage analysis were selected for a strong family history of severe obesity. This selection may have enriched for cases that simultaneously segregated multiple obesity susceptibility variants. Also, the lack of disease association in random cases may indicate that R125W only accounts for a small fraction of the population risk for severe obesity. Although the allele is relatively common in the general population (allele frequency of 0.09), the frequency of the genetic background that includes the gene/gene interaction(s) that is necessary to increase disease risk is unknown and could be much lower.

Our data suggest that both TBC1D1 and 4q34–35 may interact with additional genes to confer disease susceptibility. To identify the interaction, families segregating R125W were divided into two-generation nuclear families. As a result of our selection criteria, every two-generation family segregated the W allele with obesity, but they did not all support an obesity linkage at 4q34–35. Furthermore, in the families that did support 4q34–35, not every affected female carried the W allele of TBC1D1. These data suggest that the genetics underlying obesity predisposition are very complex. TBC1D1 and a gene at 4q34–35 may interact with each other to confer disease risk. But, it also appears that they can confer obesity risk independently of each other, perhaps by interacting with other, as yet unidentified obesity-related genes.

The genetic data presented here provide a strong evidence that links TBC1D1 with severe obesity. However, there are alternative interpretations of the data. For example, there may be a gene at 4p15–14 (but not TBC1D1) that interacts with a gene at 4q34–35 to confer obesity risk. In this scenario, selecting families that are linked to 4p15–14 (regardless of whether they carry R125W) will enrich for families that generate linkage evidence to 4q34–35. However, two results oppose this view. First, 4p15–14 linked families (LOD $\geq$ 1) that do not segregate R125W with obesity also do not generate linkage evidence at 4q34–35 (data not shown). And second, solely on the basis of segregation of R125W with obesity, we selected an independent family set that supported the linkage at 4q34–35.

The function of TBC1D1 is unknown, but it may be a regulator of either glucose or fatty acid transporters. Briefly, these transporters are sequestered on vesicles inside the cell and transported to the cell surface in response to certain extracellular signals (e.g. insulin) (33,34). This functional assignment is based on three observations. First, TBC1D1 is expressed preferentially in tissues that are important depots for both glucose and fatty acids. Second, TBC1D1 is a member of the protein family that contains TBC domains. TBC domains are known regulators of Rab GTPases, which in turn are known regulators of vesicle trafficking (35,36). This includes the regulation of vesicles that contain the glucose transporter, GLUT4. Third, TBC1D1 is closely related to AS160, and AS160 has been shown to be a regulator of GLUT4 (28,37). The proteins are 50% identical, share the number and relative order of all functional domains and have largely overlapping tissue-specific expression patterns. However, there is at least one important difference between AS160 and TBC1D1. The region including R125W is not conserved between these two proteins, which suggests that despite their similarities, TBC1D1 and AS160 have at least some different functions.

The proposed function for TBC1D1 is reasonable for a gene involved in obesity predisposition. Systemic energy balance is normally achieved, at least in part, by regulating the transport of calories into cells (38–40). Perhaps, sequence variation in TBC1D1 alters caloric disposal in insulin sensitive tissues. Over time, and in combination with interacting variants elsewhere in the genome, this caloric shift could lead to obesity and perhaps other metabolic disorders. Of course, given incomplete information, it is difficult to predict the exact molecular etiology of TBC1D1-induced obesity. For example, we cannot currently predict how R125W affects the TBC1D1 function. As the variation occurs in a protein interaction domain, we assume that R125W alters the normal regulation of TBC1D1. However, we cannot predict whether the alteration results in more or less enzymatic activity (or perhaps it results in a novel activity). Clearly, there remains much to learn about how TBC1D1 confers obesity risk. The identification of the gene underlying the obesity locus at 4q34–35 should be very instructive in this regard.

MATERIALS AND METHODS

Pedigrees and subjects

Our pedigree collection has been described previously (14,41). Briefly, we collected pedigrees from two sources. The majority of families were selected from the Utah population-based Health Family Tree Program (42). Other families were collected from a consecutive series of over 8000 gastric bypass patients in a local registry begun in 1980 by a group of surgeons (43). Population controls were derived from two sources. First, we used 90 CEPH (Centre d'Etude du Polymorphisme Humain) EBV immortalized cell lines from unrelated Utah-born subjects (ATCC). Second, we collected and used 333 unrelated subjects from the Utah population. All subjects ascertained specifically for this study signed a consent form and the University of Utah Institutional Review Board approved this investigation.

Pedigrees were selected for the 4q34–35 linkage analysis on the basis of whether or not the R125W variant in TBC1D1 segregated with obesity. To enable high throughput genotyping of R125W, the SNP was developed into a TaqMan assay (Applied BioSystems). After all the subjects in obesity pedigrees were genotyped, the pedigrees were manually inspected for segregation of R125W with obesity. The selection rule was as follows: selected pedigrees had at least two affected sisters (BMI $\geq$ 35) that carried R125W. Pedigree founders were selected to eliminate portions of the pedigree that did not contain R125W. The final pedigree was formed by unilineal descent from the selected founders, to include all of their ascertained posterity. The resultant HLOD scores were biased at 4p15–14, as genotype at TBC1D1 was included in the selection criteria, but unbiased in the rest of the genome.

Phenotypes

Height was measured by a stadiometer in bare feet with the head in the Frankfort plane and with examiner prompts to maintain a straight back and posture. Weight was measured by an electronic scale that weighs to a maximum of 800 lbs. (Scaletronic, Sharp, Inc.). BMI was calculated in kg/m^2. For subjects with gastric bypass surgery, greatest measured or reported BMI was used instead of current BMI. Subjects less than 15 years of age and deceased pedigree members were defined as having an unknown phenotype. For the female-specific phenotype used in this study, males were defined as having an unknown phenotype.

Genotyping and STR markers

For the genome search data, genotyping and marker development have been previously described (14). After identification of linkage at 4q34–35, 19 additional markers (di-, tri- and tetra-nucleotide repeats) were developed within a 32 cM interval centered on the peak of LOD (average marker spacing of 1.7 cM). The relative order of the markers within this interval was determined by using the UCSC Draft Human Genome Browser at http://genome.ucsc.edu/ (International Human Genome Sequencing Consortium 2001, 2003 and 2004), and the genetic positions of the markers were defined according to the default settings on the May 2004 version of the USCS genome browser sex-averaged genetic map (http://genome.ucsc.edu/cgi-bin/hgTrackUi?hgsid=64361850&c=chr4&g=recombRate). Inheritance of alleles was verified using the PedCheck program (44). Samples with incompatible calls

were generally regenotyped. The marker data for a family were set to missing if the incompatibility could not be resolved. The average completeness of genotyping was 95% after incompatibilities were set to zero.

Statistical analysis

The linkage analysis employed in this paper has been described previously. Briefly, haplotype solutions were determined by MCLINK (45) and LOD-score calculations employed affecteds-only versions of dominant, codominant and recessive genetic models. BMI values were used as a quantitative trait. Specifically, a logistic distribution was used to calculate the BMI-dependent ratio of the sporadic rate to the penetrance (14). We also used a non-parametric analysis (NPL) to generate a linkage statistic (46). All *P*-values reported for comparisons between cases and controls were one-sided and calculated using Fisher's exact test. This test generates exact *P*-values for both large and small sample sizes.

Selection of samples for resequencing

As mentioned earlier, haplotypes were generated and segregation patterns determined by the program MCLINK (45). Each haplotype was based on STR markers and was distinctly different from haplotypes in other families. We selected nine pedigrees with the strongest linkage evidence to 4p15–14 (LOD score $\geq$1.0) when analyzed with a dominant genetic model. We focused on the dominant model to insure that there was a simple relationship between haplotype that co-segregated with obesity (i.e. the disease-associated haploytpe) and LOD score. As we focused on families that generated LOD scores under a dominant model, the assignment of disease-associated haplotypes was straightforward (for an example, see Fig. 2). In contrast, the relationship between haplotype and LOD score under other parametric or non-parametric analyses can be very complicated.

Most of nine selected families with LOD score $\geq$1.0 were described in Stone *et al*. (14) (Table 4 in that publication lists all families with LOD score over 1.0 at D4S3350). However, 72561 and 11135 were not included in that table. They generated LOD scores over 1.0 in the linkage interval, but their linkage scores were less than 1.0 at D4S3350.

From each selected pedigree, we identified subjects for resequencing so that DNA variants could be assigned to a particular haplotype (Fig. 2). In general, the subjects shared the disease-associated haplotype, and this allowed us to make a direct but putative determination of whether the variants found during resequencing co-segregated with the disease. The assignment to the disease-associated haplotype could be confirmed later by resequencing additional pedigree members. In total, 18 individuals that carried 10 disease-associated haplotypes were selected for resequencing. In general, there was one selected haplotype for each family. The exception was pedigree 43601, which contained two haplotypes that segregated with obesity. Genomic DNA was isolated from the selected subjects and used as template for the resequencing reactions.

Gene identification and assembly

We defined the region of interest (linkage interval) from 4-MYR0264 to D4S1581. Using UCSC's assemblies of NCBI's draft human genome sequence from April 2001 to August 2001 and then NCBI's assemblies from December 2001 to July 2003 as represented on the UCSC Genome Browser, 16 genes were identified in this interval using NCBI's Reference Sequence, Human mRNA and Human EST databases. These genes were identified by searching the NCBI databases for hits using the BLAST program and by visualizing the annotations on the UCSC Genome Browser Interface (http://genome.ucsc.edu). The linkage interval in the most recent assembly (NCBI Build 35, May 2004) spans 7.45 Mb and includes two gaps. Several genes were extended using 5′- and 3′-RACE.

Resequencing

PCR amplification was used to generate products to screen for segregating variants in all identified genes located within the linkage interval. To conserve DNA, a primary amplicon was created using genomic DNA (10–20 ng) and a 25-cycle amplification. The primary amplicon was then diluted 45-fold and used as template to amplify several secondary amplicons using nested M13-tailed primers for an additional 25 cycles. If a particular amplicon proved difficult to sequence, then it was resequenced after only a single round of amplification with M13-tailed primers using 10 ng genomic DNA and 35 cycles. All samples were amplified with *Taq* Platinum (Invitrogen) DNA polymerase. PCR cycles included an initial denaturation at 96°C (12 s), annealing at 57°C (15 s) and extension at 72°C (30–60 s). Excess primers and deoxynucleotide triphosphates from M13-tailed PCR products were digested with exonuclease I (United States Biochemicals) and shrimp alkaline phosphatase (Amersham). PCR products were sequenced with M13 forward and reverse fluorescent (Big Dye, ABI) dye-labeled primers on ABI 3730 sequencers.

The R125W variant identified in TBC1D1 resulted from a change of C to T at nucleotide position 716 in reference sequence NM_015173. As yet, no rs number has been assigned to this SNP.

Northern analysis

Human multiple tissue northern blots, poly A+ RNA membranes (Clontech and Research Genetics), were hybridized overnight at 42°C with [α-^{32}P]-labeled cDNA probe (prepared by Random Primed StripAble-DNA Probe Kit, Strip-EZ DNA, Ambion). The cDNA probe was generated by PCR and contained nucleotides 3306–4134 on reference sequence NM_015173. After hybridization, the final wash was performed in 0.1% SDS and 0.1× SSC at 55°C. To measure the total amount of poly A+ RNA in each lane, the blots were rehybridized with [α-^{32}P]-labeled cDNA probes for cellular actin (β-actin) and a probe for the ribosomal protein gene 17 (rp17, data not shown).

ACKNOWLEDGEMENTS

We thank the Utah families that participated in this study. This work was funded by Myriad Genetics Inc. and by NIH grant DK44655 to S.C.H. 'Funding to pay the Open Access publication charges for this article was provided by Myriad Genetics Inc.'.

Conflict of Interest statement. Many of the authors are current employees of Myriad Genetics. Publication of this article may have a positive effect on the company's stock price and therefore financially benefit those authors.

REFERENCES

1. Hedley, A.A., Ogden, C.L., Johnson, C.L., Carroll, M.D., Curtin, L.R. and Flegal, K.M. (2004) Prevalence of overweight and obesity among US children, adolescents, and adults, 1999–2002. *JAMA*, **291**, 2847–2850.
2. Flegal, K.M., Carroll, M.D., Ogden, C.L. and Johnson, C.L. (2002) Prevalence and trends in obesity among US adults, 1999–2000. *JAMA*, **288**, 1723–1727.
3. Brolin, R.E. (2002) Bariatric surgery and long-term control of morbid obesity. *JAMA*, **288**, 2793–2796.
4. Price, R.A., Ness, R. and Laskarzewski, P. (1990) Common major gene inheritance of extreme overweight. *Hum. Biol.*, **62**, 747–765.
5. Hunt, S.C., Hasstedt, S.J., Kuida, H., Stults, B.M., Hopkins, P.H. and Williams, R.R. (1989) Genetic heritability and common environmental components of resting and stressed blood pressures, lipids, and body mass index in Utah pedigrees and twins. *Am. J. Epidemiol.*, **129**, 625–638.
6. Bell, C.G., Walley, A.J. and Froguel, P. (2005) The genetics of human obesity. *Nat. Rev. Genet.*, **6**, 221–234.
7. Price, R.A. and Lee, J.H. (2001) Risk ratios for obesity in families of obese African-American and Caucasian women. *Hum. Hered.*, **51**, 35–40.
8. Borecki, I.B., Higgins, M., Schreiner, P.J., Arnett, D.K., Mayer-Davis, E., Hunt, S.C. and Province, M.A. (1998) Evidence for multiple determinants of the body mass index: the National Heart, Lung, and Blood Institute Family Heart Study. *Obes. Res.*, **6**, 107–114.
9. Dong, C., Wang, S., Li, W.D., Li, D., Zhao, H. and Price, R.A. (2003) Interacting genetic loci on chromosomes 20 and 10 influence extreme human obesity. *Am. J. Hum. Genet.*, **72**, 115–124.
10. Dong, C., Li, W.D., Li, D. and Price, R.A. (2005) Interaction between obesity-susceptibility loci in chromosome regions 2p25–p24 and 13q13–q21. *Eur. J. Hum. Genet.*, **13**, 102–108.
11. Warden, C.H., Yi, N. and Fisler, J. (2004) Epistasis among genes is a universal phenomenon in obesity: evidence from rodent models. *Nutrition*, **20**, 74–77.
12. Lewis, C.E., North, K.E., Arnett, D., Borecki, I.B., Coon, H., Ellison, R.C., Hunt, S.C., Oberman, A., Rich, S.S. and Province, M.A. (2005) Sex-specific findings from a genome-wide linkage analysis of human fatness in non-Hispanic whites and African Americans: the HyperGEN study. *Int. J. Obes. Relat. Metab. Disord.*, **29**, 639–649.
13. Pérusse, L., Rankinen, T., Zuberi, A., Chagnon, Y.C., Weisnagel, S.J., Argyropoulos, G., Walts, B., Snyder, E.E. and Bouchard, C. (2005) The human obesity gene map: the 2004 update. *Obes. Res.*, **13**, 381–490.
14. Stone, S., Abkevich, V., Hunt, S.C., Gutin, A., Russell, D.L., Neff, C.D., Riley, R., Frech, G.C., Hensel, C.H. and Jammulapati, S. (2002) A major predisposition locus for severe obesity, at 4p15–p14. *Am. J. Hum. Genet.*, **70**, 1459–1468.
15. Arya, R., Duggirala, R., Jenkinson, C.P., Almasy, L., Blangero, J., O'Connell, P. and Stern, M.P. (2004) Evidence of a novel quantitative-trait locus for obesity on chromosome 4p in Mexican Americans. *Am. J. Hum. Genet.*, **74**, 272–282.
16. Clerget-Darpoux, F., Bonaiti-Pellie, C. and Hochez, J. (1986) Effects of misspecifying genetic parameters in lod score analysis. *Biometrics*, **42**, 393–399.
17. Bell, C.G., Benzinou, M., Siddiq, A., Lecoeur, C., Dina, C., Lemainque, A., Clement, K., Basdevant, A., Guy-Grand, B. and Mein, C.A. (2004) Genome-wide linkage analysis for severe obesity in French Caucasians finds significant susceptibility locus on chromosome 19q. *Diabetes*, **53**, 1857–1865.
18. Rice, T., Chagnon, Y.C., Perusse, L., Borecki, I.B., Ukkola, O., Rankinen, T., Gagnon, J., Leon, A.S., Skinner, J.S. and Wilmore, J.H. (2002) A genomewide linkage scan for abdominal subcutaneous and visceral fat in black and white families: The HERITAGE Family Study. *Diabetes*, **51**, 848–855.
19. Zhou, M.M., Ravichandran, K.S., Olejniczak, E.F., Petros, A.M., Meadows, R.P., Sattler, M., Harlan, J.E., Wade, W.S., Burakoff, S.J. and Fesik, S.W. (1995) Structure and ligand recognition of the phosphotyrosine binding domain of Shc. *Nature*, **378**, 584–592.
20. Yan, K.S., Kuti, M. and Zhou, M.M. (2002) PTB or not PTB—that is the question. *FEBS Lett.*, **513**, 67–70.
21. Zwahlen, C., Li, S.C., Kay, L.E., Pawson, T. and Forman-Kay, J.D. (2000) Multiple modes of peptide recognition by the PTB domain of the cell fate determinant Numb. *EMBO J.*, **19**, 1505–1515.
22. Yaich, L., Ooi, J., Park, M., Borg, J.P., Landry, C., Bodmer, R. and Margolis, B. (1998) Functional analysis of the Numb phosphotyrosine-binding domain using site-directed mutagenesis. *J. Biol. Chem.*, **273**, 10381–10388.
23. Altshuler, D., Brooks, L.D., Chakravarti, A., Collins, F.S., Daly, M.J. and Donnelly, P. (2005) A haplotype map of the human genome. *Nature*, **437**, 1299–1320.
24. Richardson, P.M. and Zon, L.I. (1995) Molecular cloning of a cDNA with a novel domain present in the tre-2 oncogene and the yeast cell cycle regulators BUB2 and cdc16. *Oncogene*, **11**, 1139–1148.
25. Will, E. and Gallwitz, D. (2001) Biochemical characterization of Gyp6p, a Ypt/Rab-specific GTPase-activating protein from yeast. *J. Biol. Chem.*, **276**, 12135–12139.
26. Letunic, I., Copley, R.R., Schmidt, S., Ciccarelli, F.D., Doerks, T., Schultz, J., Ponting, C.P. and Bork, P. (2004) SMART 4.0: towards genomic data integration. *Nucleic Acids Res.*, **32**, D142–D144.
27. Xu, X.Z., Wes, P.D., Chen, H., Li, H.S., Yu, M., Morgan, S., Liu, Y. and Montell, C. (1998) Retinal targets for calmodulin include proteins implicated in synaptic transmission. *J. Biol. Chem.*, **273**, 31297–31307.
28. Sano, H., Kane, S., Sano, E., Miinea, C.P., Asara, J.M., Lane, W.S., Garner, C.W. and Lienhard, G.E. (2003) Insulin-stimulated phosphorylation of a Rab GTPase-activating protein regulates GLUT4 translocation. *J. Biol. Chem.*, **278**, 14599–14602.
29. Matsumoto, Y., Imai, Y., Lu Yoshida, N., Sugita, Y., Tanaka, T., Tsujimoto, G., Saito, H. and Oshida, T. (2004) Upregulation of the transcript level of GTPase activating protein KIAA0603 in T cells from patients with atopic dermatitis. *FEBS Lett.*, **572**, 135–140.
30. Moore, J.H. (2003) The ubiquitous nature of epistasis in determining susceptibility to common human diseases. *Hum. Hered.*, **56**, 73–82.
31. Cox, N.J., Frigge, M., Nicolae, D.L., Concannon, P., Hanis, C.L., Bell, G.I. and Kong, A. (1999) Loci on chromosomes 2 (NIDDM1) and 15 interact to increase susceptibility to diabetes in Mexican Americans. *Nat. Genet.*, **21**, 213–215.
32. Ott, J. (2004) Issues in association analysis: error control in case–control association studies for disease gene discovery. *Hum. Hered.*, **58**, 171–174.
33. Bonen, A., Luiken, J.J. and Glatz, J.F. (2002) Regulation of fatty acid transport and membrane transporters in health and disease. *Mol. Cell. Biochem.*, **239**, 181–192.
34. Furtado, L.M., Somwar, R., Sweeney, G., Niu, W. and Klip, A. (2002) Activation of the glucose transporter GLUT4 by insulin. *Biochem. Cell Biol.*, **80**, 569–578.
35. Seabra, M.C., Mules, E.H. and Hume, A.N. (2002) Rab GTPases, intracellular traffic and disease. *Trends Mol. Med.*, **8**, 23–30.
36. Cormont, M. and Le Marchand-Brustel, Y. (2001) The role of small G-proteins in the regulation of glucose transport (review). *Mol. Membr. Biol.*, **18**, 213–220.
37. Zeigerer, A., McBrayer, M.K. and McGraw, T.E. (2004) Insulin stimulation of GLUT4 exocytosis, but not its inhibition of endocytosis, is dependent on RabGAP AS160. *Mol. Biol. Cell*, **15**, 4406–4415.
38. Felber, J.P. and Golay, A. (2002) Pathways from obesity to diabetes. *Int. J. Obes. Relat. Metab. Disord.*, **26**(Suppl. 2), S39–S45.
39. Kahn, B.B. and Flier, J.S. (2000) Obesity and insulin resistance. *J. Clin. Invest.*, **106**, 473–481.
40. Raz, I., Eldor, R., Cernea, S. and Shafrir, E. (2005) Diabetes: insulin resistance and derangements in lipid metabolism. Cure through intervention in fat transport and storage. *Diabetes Metab. Res. Rev.*, **21**, 3–14.

41. Hunt, S.C., Abkevich, V., Hensel, C.H., Gutin, A., Neff, C.D., Russell, D.L., Tran, T., Hong, X., Jammulapati, S. and Riley, R. (2001) Linkage of body mass index to chromosome 20 in Utah pedigrees. *Hum. Genet.*, **109**, 279–285.
42. Williams, R.R., Hunt, S.C., Barlow, G.K., Chamberlain, R.M., Weinberg, A.D., Cooper, H.P., Carbonari, J.P. and Gotto, A.M. (1988) Health family trees: a tool for finding and helping young members of coronary and cancer prone pedigrees in Texas and Utah. *Am. J. Public Health*, **78**, 1283–1286.
43. Adams, T.D., Avelar, E., Cloward, T., Crosby, R.D., Farney, R.J., Gress, R., Halverson, R.C., Hopkins, P.N., Kolotkin, R.L. and Lamonte, M.J. (2005) Design and rationale of the Utah obesity study. A study to assess morbidity following gastric bypass surgery. *Contemp. Clin. Trials*, **26**, 534–551.
44. O'Connell, J.R. and Weeks, D.E. (1998) PedCheck: a program for identification of genotype incompatibilities in linkage analysis. *Am. J. Hum. Genet.*, **63**, 259–266.
45. Thomas, A., Gutin, A., Abkevich, V. and Bansal, A. (2000) Multilocus linkage analysis by blocked Gibbs sampling. *Statist. Comput.*, **10**, 259–269.
46. Camp, N.J., Gutin, A., Abkevich, V., Farnham, J.M., Cannon-Albright, L. and Thomas, A. (2001) A new nonparametric linkage statistic for mapping both qualitative and quantitative trait loci. *Genet. Epidemiol.*, **21** (Suppl. 1), S461–S466.