Published in final edited form as:
Hum Biol. 2008 April ; 80(2): 99–123.

Applying Novel Genome-Wide Linkage Strategies to Search for Loci Influencing Type 2 Diabetes and Adult Height in American Samoa

KAROLINA ÅBERG[1], **GUANGYUN SUN**[2], **DIANE SMELSER**[2], **SUBBA RAO INDUGULA**[2], **HUI-JU TSAI**[3], **MATTHEW S. STEELE**[4], **JOHN TUITELE**[5], **RANJAN DEKA**[2], **STEPHEN T. MCGARVEY**[6], and **DANIEL E. WEEKS**[1,7]

[1]Department of Human Genetics, Graduate School of Public Health, University of Pittsburgh, A300 Crabtree Hall, 130 Desoto St., Pittsburgh, PA 15261

[2]Center for Genome Information, Department of Environmental Health, University of Cincinnati, Cincinnati, OH

[3]Mary Ann and J. Milburn Smith Child Health Research Program, Children's Memorial Hospital and Children's Memorial Research Center, Department of Pediatrics, Feinberg School of Medicine, Northwestern University, Chicago, IL

[4]Program for Appropriate Technology in Health, Seattle, WA

[5]Tafuna Family Health Center, Pago Pago, American Samoa

[6]International Health Institute, Brown University, Providence, RI

[7]Department of Biostatistics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA

Abstract

Type 2 diabetes mellitus (T2DM) is a common complex phenotype that by the year 2010 is predicted to affect 221 million people globally. In the present study we performed a genome-wide linkage scan using the allele-sharing statistic S_{all} implemented in Allegro and a novel two-dimensional genome-wide strategy implemented in Merloc that searches for pairwise interaction between genetic markers located on different chromosomes linked to T2DM. In addition, we used a robust score statistic from the newly developed QTL-ALL software to search for linkage to variation in adult height. The strategies were applied to a study sample consisting of 238 sib-pairs affected with T2DM from American Samoa. We did not detect any genome-wide significant susceptibility loci for T2DM. However, our two-dimensional linkage investigation detected several loci pairs of interest, including 11q22 and 21q21, 9q21 and 11q22, 1p22–p21 and 4p15, and 4p15 and 15q11–q14, with a two-loci maximum LOD score (MLS) greater than 2.00. Most detected individual loci have previously been identified as susceptibility loci for diabetes-related traits. Our two-dimensional linkage results may facilitate the selection of potential candidate genes and molecular pathways for further diabetes studies because these results, besides providing candidate loci, also demonstrate that polygenic effects may play an important role in T2DM. Linkage was detected (p value of 0.005) for variation in adult height on chromosome 9q31, which was reported previously in other populations. Our finding suggests that the 9q31 region may be a strong quantitative trait locus for adult height, which is likely to be of importance across populations.

Keywords

DIABETES; STATURE; LINKAGE ANALYSIS; TWO-DIMENSIONAL LINKAGE; QUANTITATIVE TRAIT LOCUS; QTL-ALL SOFTWARE

Type 2 diabetes mellitus (T2DM) is a multifactorial metabolic disorder characterized by insulin resistance and dysfunction of the pancreatic beta cells. The disease results from a complex interaction of genetic and environmental factors and is strongly associated with obesity and modernized lifestyle. The disease has emerged to become a major public health problem. In 2000 T2DM affected 151 million people globally. This number is predicted to increase to 221 million in 2010, and the greatest proportional increase is expected to occur in regions adopting a modernized lifestyle, such as Africa and Asia with a 50% and 57% increase, respectively (Zimmet et al. 2001).

A genetic contribution to the development of T2DM is supported by family aggregation with increased risk of developing the disease for first-degree relatives of probands and a greater concordance between affected monozygotic twins compared to dizygotic twins (Barroso 2005). In addition, disease prevalence varies substantially between ethnic groups living in the same environment, suggesting racial disparity for T2DM (Carulli et al. 2005).

In this study we search for linkage to T2DM in affected sib-pairs from American Samoa. The population on the Samoan islands was likely established by a relatively small number of founders about 3,000 years ago and had been fairly isolated until the 20th century (Tsai et al. 2004). Since the 1950s American Samoa has received substantial financial support from the U.S. federal government, which has resulted in rapid economic modernization, nutrition transition, and a dramatic increase in the prevalence of obesity and T2DM (McGarvey 2001; Keighley et al. 2007). In 2002 the prevalence of T2DM was 21.6% for men and 18% for women in American Samoa. In contrast, the prevalence of T2DM in the independent nation of Samoa in 2003 was 9.3% for men and 12.6% for women (Keighley et al. 2007). Most inhabitants of Samoa have not been exposed to the same rapid modernization as American Samoans have, but members of both polities share an almost identical genetic background (Deka et al. 1994; Tsai et al. 2001).

Proteinuria is strongly associated with T2DM, causing thickening of the glomerulus and excessive albumin excretion in urine, and has been associated with short stature in T2DM. It has been suggested that a common genetic factor predisposes individuals to both short stature and diabetic renal disease (Fava et al. 2001). Because some evidence shows familial clustering of renal disease in T2DM (Pettitt et al. 1990; Fava et al. 2000), we hypothesize that short stature also could cluster in T2DM families, and therefore the collected data make an interesting study sample for investigations of adult height.

In the present study we performed a traditional genome-wide scan using the nonparametric allele-sharing statistic S_{all}, as implemented in Allegro (Gudbjartsson et al. 2000, 2005) (available at http://www.decode.com/software/), and the more novel nonparametric two-dimensional genome-wide scan in which we search for interaction between any two distant loci linked to T2DM using Merloc (Bell et al. 2006) (available at http://www.well.ox.ac.uk/~jordana/merloc/). In addition, we searched for linkage to adult height within the population of American Samoa using a score statistic implemented in the newly developed QTL-ALL software (Bhattacharjee et al. 2008; Mukhopadhyay et al. 2008) (available at http://watson.hgen.pitt.edu/register/).

Materials and Methods

Sample and Phenotype Collection

In 1998 we ascertained affected sib-pairs diagnosed with T2DM and collected blood samples and phenotypes from 386 adults with T2DM from 174 families residing in American Samoa. All participants reported that all four grandparents originated from the Samoan islands. We obtained T2DM status from the diabetes registry at the American Samoa Department of Health with the diagnosis of T2DM confirmed by medical records. The criterion for inclusion of a sib-pair in the study was that both sibs had T2DM diagnoses stated in their medical records. It should be noted that the registry also includes individuals diagnosed with type 1 diabetes mellitus (T1DM); these individuals were excluded from this study. An extensive medical history was collected with a focus on information about how and when the sibs' T2DM was diagnosed. Eighty-four percent of the study sample was using prescribed medication for their T2DM at the time of collection. In this study we are investigating two T2DM phenotypes: (1) a broad phenotype including all patients diagnosed with T2DM and (2) a narrow phenotype that we call T2DM-strict, which includes patients diagnosed with T2DM and treated with prescribed medication (Table 1).

The average time from diagnosis of T2DM to study participation was 8.4 years for men and 7.5 years for women. The average age of the study individuals was 56.5 ± 10.55 years for men and 53.2 ± 12.5 years for women. General characteristics of the study samples are shown in Table 2.

Height was measured with a portable anthropometer using standard methods according to the *Anthropometric Standardization Reference Manual* (Lohman et al. 1988). One female research assistant measured all female participants, and male participants were measured by one of the two male research assistants. There was no formal estimation of error of measurement of adult height. The Institutional Review Board (IRB) of the Miriam Hospital, Providence, Rhode Island, and the Department of Health, American Samoa, approved study protocols, and written informed consent was obtained from all participants.

Marker Genotyping

We successfully genotyped 386 microsatellite markers from the ABI prism linkage mapping set MD10, v. 2.5 (Applied Biosystems Inc., Foster City, California) using an ABI 377 automated DNA sequencer (Applied Biosystems) as described previously (Tsai et al. 2001). The microsatellites in this marker set have an approximate spacing of 10 cM. Allele calls were generated using GeneScan, v. 3.1, and Genotyper, v. 2.5 (Applied Biosystems), software. In addition, nine forensic markers, from the Amp-Fl STR Profiler Plus kit (Applied Biosystems), were genotyped using the same strategy as was used for the ABI prism linkage mapping set.

Statistical Methods: Data Handling and Error Checking

R statistical software (R Project for Statistical Computing; available at http://www.r-project.org/) and Mega2 (Mukhopadhyay et al. 2005) (available at http://watson.hgen.pitt.edu/register/softdoc.html) were used to set up files for all software used. To check the accuracy of self-reported pedigree structure, we used the maximum log-likelihood ratio statistic as implemented in the relationship testing program PREST (McPeek and Sun 2000; Sun et al. 2002) (available at http://fisher.utstat.toronto.edu/sun/Software/Prest). Because this test is performed between all genotyped relative pairs within each pedigree, it results in a large number of tests that have to be accounted for when the results are interpreted. A p value of 0.00004 from PREST corresponds to an alpha level of 0.01 that has been Bonferroni-corrected for the number of independent tests performed. Because each sib-

pair is not fully independent from the other sib-pairs within its sibship, we have adjusted for the number of pair equivalents of information, where each sibship contributes with $[2n - 3 + (1/2)^{n-1}]/1.5$ (where n is the size of the sibship) (Hodge 1984). Thus we applied a p value greater than or equal to 0.00004 as a cutoff value to reject the reported pedigree structure for a specific relative pair. To resolve relationship errors, we carefully removed the minimum number of problematic individuals so that no more relationship errors remained; for the included individuals the reported pedigree structure agreed with the output from PREST, and this formed the "conservative" sample set investigated in this study.

To try to include as many of the collected genotyped and phenotyped samples as possible, we used RelPair (Boehnke and Cox 1997; Epstein et al. 2000) (available at http://csg.sph.umich.edu/index.php) to estimate the likelihood that the relative pairs were half-sibs or first cousins rather than full sibs. All relative pairs that were detected by PREST as possible relationship errors and that were called by RelPair half-sibs or first cousins instead of full sibs, together with the full sibs that were in accordance with the self-reported pedigree structure, formed the "alternative" sample set studied. That is, we altered the pedigree structure to redefine relationships to agree with the most likely relationship as reported by RelPair.

To further ensure correct pedigree structure, we applied the genotypes from nine hypervariable autosomal forensic markers to the updated conservative and alternative pedigree structures and used RelPair to check that no other structures were more likely. In addition, we used the forensic X chromosome marker and the microsatellite markers located on the X chromosome to ensure that the reported sex was correct.

After relationship tests were performed and pedigree structures were updated accordingly, we checked all pedigrees for Mendelian-inconsistent genotypes using the software PedCheck (O'Connell and Weeks 1998) (available at http://watson.hgen.pitt.edu/register/softdoc.html). When inconsistent genotypes were detected, we removed the genotypes at that marker from all individuals within the inconsistent family.

To further ensure good-quality genotype data, we used the alignment strategy described previously (Aberg et al. 2008) to align the allele-frequency patterns from this data set with the pattern observed in our previous investigation of the Samoan population (Aberg et al. 2008; Dai et al. 2008) for 357 autosomal markers that were genotyped in common. The alignment itself does not detect any specific genotyping errors. However, if the overall allele-frequency pattern from a specific marker agrees with frequency patterns from other data sets from the same population, the alignment decreases the risk that artifacts from the genotyping process and allele-calling process have been included in the investigated data set as true alleles.

Allele-Frequency Estimation

Allele frequencies at each marker were obtained by counting the alleles in all genotyped individuals within the data set, ignoring their relationships. Using the genotype data from all individuals in a sample set to estimate allele frequencies is unbiased and provides improved information, compared with using only the data of one individual from each family. However, properly accounting for relationships between individuals might improve the estimates of allele frequencies but is usually of minor importance for the outcome of linkage results (Broman 2001). In addition, to account for any biases resulting from nonrandomly ascertained samples, we used allele frequencies obtained in the same way from an external data set of 1,164 individuals who were ascertained without selection for any phenotype from the population on the Samoan islands (Aberg et al. 2008). Unless otherwise stated, the allele frequency obtained from the sample set studied in the present study is used.

Genetic Map

In this genome-wide study and in three other genome-wide studies previously performed in the population of the Samoan islands (Aberg et al. 2008; Dai et al. 2007, 2008), we have used a genetic map based on Kosambi centimorgans (cM) taken from the Rutgers Combined Linkage-Physical Map (Kong et al. 2004). Linear interpolation was used to include polymorphic markers not included in the Rutgers map.

Linkage Analyses for Diabetes-Related Traits

To investigate the two qualitative diabetes-related phenotypes, T2DM and T2DM-strict, we performed single-and multipoint nonparametric linkage analyses using Allegro (Gudbjartsson et al. 2000, 2005). Based on the results from Sengul et al. (2001), which suggest that the commonly used S_{all} statistic performs well in most situations, we report the S_{all} statistics under the linear model, which uses identity-by-descent (IBD) sharing among all affected members within a family (Whittemore and Halpern 1994). Allegro returns a nonparametric linkage score as well as an allele-sharing LOD score. Here we have chosen to report the LOD score multiplied by the sign of the $\hat{\delta}$ parameter, which allows the LOD score to indicate both increased (positive values) and decreased allele sharing (negative values). Allegro conducts the analysis for autosomes as well as for the X chromosome. Because we were not able to determine the shared parent for the half-sibs included in the alternative data set, we performed X-chromosome-linked analysis for only the conservative sample set.

Empirical Genome-Wide Level of Significance

In this study we evaluated the results from the one-dimensional genome-wide linkage analysis of T2DM using locus counting (Wiltshire et al. 2002). In short, we estimated the null distribution of regions showing evidence for linkage by simulating 1,000 replicates of the autosomal genome using the precise marker map, allele frequencies, family structure, and missing genotype pattern as observed in our sample through the use of Simulate (Laboratory of Statistical Genetics, Rockefeller University; available at ftp://linkage.rockefeller.edu/software/simulate). The simulated genomes were analyzed for linkage using Allegro (Gudbjartsson et al. 2000, 2005). For the single-point linkage analysis we counted the number of markers that showed linkage, and for the multipoint linkage analysis we counted the numbers of independent regions showing linkage. Regions were considered independent of each other if their maxima were separated by 55 cM (Haldane) (Wiltshire et al. 2002).

Power Simulations for T2DM

We used Allegro (Gudbjartsson et al. 2000, 2005) to simulate multipoint data given a single-locus disease model. The pedigree structure and information on who was affected and on who was genotyped were kept identical with the original data set. The simulations were performed under three additive models, where the relative risks were 1.2, 2.1, and 3.2, as previously described by Sengul et al. (2007). We performed 900 simulated data sets for each model, calculated nonparametric linkage as described, and calculated the percentage of how often a genome-wide significant linkage peak was detected in the location of the true signal.

Two-Dimensional Linkage Analyses for Diabetes-Related Traits

To investigate pairwise interaction between genetic markers located on different chromosomes, we used the software Merloc (Bell et al. 2006) to test for two-locus nonparametric linkage to the diabetes-related traits. For any combination of chromosome regions located on two different chromosomes, the program generates a two-dimensional grid and calculates the two-loci maximum LOD score (MLS) for each marker pair, using Merlin (Abecasis et al. 2002) as a multipoint likelihood-calculating engine. Because the

Merloc software is currently limited to analyzing nuclear families, we did not perform two-dimensional linkage analyses for the alternative sample set, as this data set contains relative pairs that exceed this limit.

Autosomal Linkage Analyses for Adult Height

To investigate the quantitative trait adult height in the studied sample from American Samoa, we used a score statistic. Score statistics for quantitative trait locus (QTL) mapping have been discussed by, for example, Tang and Siegmund (2001), Putter et al. (2002), and Wang and Huang (2002). The version we use is similar to the version discussed by these investigators, but it uses a maximized empirical variance estimate in the denominator that makes it both robust and powerful in the presence of selected sampling. This statistic is implemented under the name Score.Max in the newly developed software QTL-ALL (Bhattacharjee et al. 2007; Mukhopadhyay et al. 2008). The QTL-ALL software is limited to handling nuclear families, so adult height was not studied in the alternative data set. We calculated multipoint statistics, and we estimated all parameters needed to calculate the score statistics from the data set using QTL-ALL. Before performing the linkage analysis, we checked the data set for extreme outliers, and we removed systematic variation caused by age by fit-ting a linear model to the height observations for males and females, respectively. We then used the differences between the observed and the fitted values (i.e., the residuals) as the trait in the linkage analysis. The combined residuals appeared to be normally distributed, and therefore no further transformation was performed. Because this strategy does not account for variations of the age effect associated with different developmental stages, such as normal growth in young adults and shrinking as a result of, for example, osteoporosis in elderly, we restricted our sample set to individuals between 22 and 55 years of age. By including individuals from only this limited age span, we aimed to minimize differences in adult height caused by a potential cohort effect, with the younger generations being taller than the older generations, which has been observed previously in a different population (Mukhopadhyay et al. 2003).

Power Simulations for Adult Height

We used SLINK (Ott 1989; Weeks et al. 1990; Cottingham et al. 1993) to simulate two microsatellite markers located 5 cM on each side of a true QTL. The pedigree structure and information on who was phenotyped and on who was genotyped were kept identical with the original data set. The simulations were performed under three additive genetic models, without any dominant effect, where the locus-specific heritability was set to 0.7, 0.35, and 0.175, the allele frequencies were set to 0.5, and the simulated trait had an overall mean of 0 and a total variance of 1. We performed 1,000 simulated data sets for each model, calculated the score statistics as described earlier with QTL-ALL (Bhattacharjee et al. 2008; Mukhopadhyay et al. 2008), and calculated the percentage of how often a significant linkage peak (genome-wide $\alpha = 0.05$) was detected in the location of the true signal.

Linkage Analyses for Adult Height on the X Chromosome

The current version of QTL-ALL does not carry out analyses for the X chromosome, and to our knowledge no other suitable score statistic is available for studying quantitative traits on the X chromosome. We have therefore applied a variance component method, as recently implemented in Mendel (Lange et al. 2001) (available at http://www.genetics.ucla.edu/software/), to search for linkage to adult height on the X chromosome. We standardized the trait by sex and applied three different models in which X-chromosome-linked QTL and random environmental variance components were always included, whereas the autosomal additive polygenic and X-chromosome-linked additive polygenic variance components were included simultaneously as well as one at a time.

Results

Relationship and Sex Checking

Sixty-one relative pairs were not supported by statistical testing using PREST on the self-reported pedigrees. Genotype and phenotype data from these 61 relative pairs were removed, and the remaining samples formed the conservative data set investigated in this study. Eleven of the relative pair discrepancies were included in the alternative data set as half-sibs or first cousins, as suggested by RelPair. The total number of phenotyped and genotyped relative pairs included in this investigation is reported in Table 1. No additional discrepancies in the pedigree structures were detected with the nine hypervariable autosomal forensic markers.

Genotype Error Checking

The genotyped relative pairs included one monozygotic twin pair that was eliminated from all statistical calculations but was used for calculating genotype error rates. Out of 634 autosomal allele calls for markers in common and 26 allele calls for markers in common on the X chromosome, we found 12 and 3 discrepancies, respectively. This suggests an approximate genome-wide error rate for the data set of 2.3%. However, because of the limited size of the families included in this study, it is difficult to identify the precise individual carrying a genotype error (Mukhopadhyay et al. 2004). Therefore this fairly low genotype error rate results in a higher percentage of missing genotypes resulting from genotype errors because we removed all genotypes for a specific maker from all individuals within each family that contained a Mendelian inconsistency.

When aligning the allele-frequency patterns from each marker with the patterns from another data set from the Samoan islands, the overall alignments matched well and no marker had an allele with a frequency greater than 10% that was not detected in the other data set.

Genome-Wide Linkage Analysis for T2DM Phenotypes

The results from our simulation study suggest that in this particular investigation a multipoint and a single-point LOD score of 3.7 corresponds to a genome-wide empirical p value of 0.01 and a LOD score of 2.9 corresponds to a p value of 0.05.

We detected a LOD score of 1.26 with single-point statistics at marker *D4S403* for the T2DM phenotype in the conservative data set. When the more stringent phenotype, T2DM-strict, was used, the LOD scores slightly increased to 1.44 at this marker. With the multipoint statistics we detected a LOD score of 1.12 within this region for the T2DM-strict phenotype. No LOD score of interest was detected for the broader T2DM phenotype. In addition, potential linkage with a LOD score of 1.71 to the T2DM-strict phenotype was detected at marker *D7S484* when using single-point statistics (Figure 1). The one-dimensional linkage results for the alternative sample set were similar overall to the results detected in the conservative data set. For instance, the peaks reported were only marginally increased (0.03–0.17) in the alternative data set.

Two-Loci Linkage Analysis for T2DM Phenotypes

To search for two loci that simultaneously affect T2DM or T2DM-strict, we performed a two-dimensional genome-scan using Merloc (Bell et al. 2006). In this investigation we found five chromosome combinations with a MLS greater than 1.6 for the broad T2DM phenotype and eight chromosome combinations for the more stringent, T2DM-strict phenotype. One of the chromosome combinations detected with the broad phenotype was also detected with the T2DM-strict phenotype (Table 3).

QTL Analysis for Height

We have performed a QTL analysis for height, using the multipoint Score.Max statistics as implemented in QTL-ALL (Bhattacharjee et al. 2008; Mukhopadhyay et al. 2008). We excluded subjects whose age was less than 22 or greater than 55 and adjusted for age within each sex in the remaining samples before performing the QTL analysis. No extreme outliers were observed in this study sample. The most promising QTL for stature was found on chromosome 9q31. Within this region we detected a Score.Max *p* value of 0.005 with marker *D9S1690* (Table 4 and Figure 2). We detected three additional chromosome regions with *p* values less than or equal to 0.01. We detected a Score.Max *p* value of 0.009 with marker *D1S2836* on chromosome 1q44, a *p* value of 0.010 with *D10S249* on chromosome 10p15, and a *p* value of 0.008 with *D22S274* on chromosome 22q13 (Table 4 and Figure 2). Furthermore, we detected eight chromosome regions with *p* values ranging from 0.05 to 0.01 on chromosomes 2p25, 2q32–q33, 10q26, 12q21,12q23, 12q24, and 15q11 (Table 4 and Figure 2).

To investigate linkage between adult height and the X chromosome, we used the variance component approach as implemented in Mendel (Lange et al. 2001). With this method and the three tested models described earlier, we did not detect any LOD score greater than 0.75 on the X chromosome.

Discussion

Flat One-Dimensional Linkage Peak Pattern

In any genome-wide linkage study one would expect to see a moderate number of modest peaks across the genome just by chance. Our simulations predict that the average simulated genome-wide scan for the studied data set should have 2.2 multipoint linkage peaks greater than 1.175 and 4.0 single-point LOD scores greater than 1.175. In the present study, when investigating the diabetes-related traits, we observed a linkage peak pattern that is slightly flatter than what our prediction suggests. We found no multipoint results greater than 1.175 and only one marker for the broad phenotype and two markers for the narrow phenotype with single-point results greater than 1.175. However, according to our simulations, it is a fairly high probability (about 8–18%) to observe these slightly flatter genome-wide linkage peak patterns by chance. Because the expected number of peaks is derived under the assumption that the markers have nothing to do with the disease and because in our study we observed a slightly lower number of peaks than expected on average, we have no support for one-dimensional linkage to the studied T2DM-related phenotypes in this data set. This lack of linkage to the studied phenotypes could be due to many various factors.

First, T2DM is a complex pathophysiologic disease state with variation in its natural history and multiple putative molecular scenarios. Thus a common T2DM phenotype could be due to different underlying genetic factors, inheritance patterns, and environmental influences (Barroso 2005). The T2DM diagnoses for this study sample may have high heterogeneity because they were obtained from a diabetes registry and medical records, with no selection for body mass index and sex and a substantial interval of about 8 years between diagnosis and study participation. In addition, the level of T2DM knowledge and awareness was low at the time of the study in American Samoa, suggesting that many participants may have had T2DM for many years before being diagnosed. In an attempt to investigate a more homogeneous phenotype, T2DM-strict, we were actually able to slightly increase the variation of the linkage peak pattern, which supports the theory that the studied individuals diagnosed with T2DM could be a quite heterogeneous group. In addition, several studies have reported sex effects on the underlying genetic architecture of diabetes-related traits (Stone et al. 2002; Lewis et al. 2005; North et al. 2007). In this study we have investigated

sib-pairs without making any distinctions between their sexes, which might further increase the heterogeneity of the phenotypes studied.

Second, despite careful relationship testing, the study sample might contain undetected relationship errors. If the true relationship is more distant than the reported relationship, the observed sharing between relative pairs would be decreased relative to the expected sharing, which in turn would decrease the power of our investigation. Furthermore, because the studied sample set originates from a small population founded by a small numbers of individuals that have been fairly isolated for a large number of generations (Tsai et al. 2004), it is likely that inbreeding exists within the study population. However, not accounting for inbreeding would lead to increased allele sharing and therefore to an increased number of false-positives (Genin and Clerget-Darpoux 1996; Leutenegger et al. 2002) rather than a flat linkage peak pattern.

Third, although we put strong efforts into detecting genotype errors, it is still possible that the sample set contains undetected Mendelian genotype errors, particularly because no parents were genotyped. However, the comparison of the genotyped monozygotic twin pair and the allele-frequency patterns between different data sets suggests that the genotype error rate is low (≈ 2.3%) in this data set.

Fourth, if a marker is in strong linkage disequilibrium with a disease susceptibility gene, the marker allele frequency estimates obtained from affected relative pairs might be biased. Therefore, in addition to estimating allele frequencies from the collected sample set, we also used an external set of allele frequencies for the linkage investigation of the conservative data set. No notable differences between the linkage studies performed with the allele frequencies obtained from the studied sample set and the external allele frequencies were observed. We therefore argue that the flat linkage peak pattern observed when studying the diabetes phenotypes is not an effect of incorrect allele frequencies.

Another factor that could cause a flat linkage peak pattern is an incorrect marker map. The microsatellite markers and the applied marker map used in this study are commonly used by genome-wide scans of various phenotypes in many populations. The markers have been mapped both physically and genetically and therefore are unlikely to be dramatically misplaced. Furthermore, in the present study we observe flat linkage patterns for the multipoint study as well as for the single-point study. Because the single-point study is not affected by the marker map but still shows a flat pattern, we do not believe that the lack of peaks is due to problems with the marker map.

Finally, it is possible that our flat linkage peaks reflect a true negative result, meaning that no susceptibility loci for the studied heterogeneous T2DM phenotypes segregate within the American Samoan study sample. It is possible that the development of the T2DM phenotype is mainly mediated by physical and environmental factors, such as blood pressure, obesity, food intake, and lack of exercise. However, because multiple studies of various populations have identified susceptibility loci for T2DM, it is more likely that such loci also exist in the Samoan population but that the present study does not have adequate statistical power to detect such loci. Simulations show that our study sample has 78% power to detect significant multipoint linkage ($\alpha = 0.05$) when the relative risk is fairly high (3.2). When the relative risk is reduced to 2.1, the power to detect genome-wide significant linkage is reduced to 47%. However, when the relative risk is reduced to 1.2, the present study sample does not have adequate power to detect linkage. Indeed, recent genome-wide association studies for T2DM susceptibility loci (Saxena et al. 2007; Scott et al. 2007; Sladek et al. 2007; Steinthorsdottir et al. 2007; Zeggini et al. 2007) found genes with only relatively

small relative risks (maximum $\lambda_S \approx 1.07$) (Zeggini et al. 2007); such loci are not of large enough effect to be detectable in our study.

Susceptibility Loci for Diabetes-Related Traits

Neither the one-dimensional nor the two-dimensional linkage studies of the broad and narrow phenotypes revealed any genome-wide significant susceptibility loci for diabetes. However, the regions on chromosome 4p15 and 7q31 that were detected by the one-dimensional scan and most of the loci with a MLS above 1.6 detected in the two-dimensional scan have been reported previously as susceptibility loci for diabetes or diabetes-related traits in other populations. Chromosome 4p15 has been suggested as a susceptibility region for T2DM (Wiltshire et al. 2004), and 7q31 has been suggested to carry a susceptibility locus for metabolic syndrome (Loos et al. 2003) and for fasting insulin levels (Lakka et al. 2003).

The largest two-dimensional MLS for the stringent phenotype was detected in region 11q22, which previously has been implicated as conferring susceptibility to T2DM (Traurig et al. 2006) and T1DM (Dong et al. 2007), and in region 21q21, which has been suggested as a potential locus for T1DM (Bergholdt et al. 2005). Similarly, the largest two-dimensional MLS for the broad phenotype was detected on chromosomes 12q24 and 20q13. The 12q24 region previously has been reported as linked to both T2DM (Lindgren et al. 2002; Rotimi et al. 2004; Wiltshire et al. 2004) and T1DM (Tessier et al. 2006) and to diabetic risk factors such as high body mass index (Cornes et al. 2005) and abdominal subcutaneous fat (Perusse et al. 2001). The chromosome 20q13 region has been linked to T2DM in several studies (Klupa et al. 2000; Bento et al. 2004; Damcott et al. 2004; Rotimi et al. 2004; Silander et al. 2004; Cheyssac et al. 2006).

Furthermore, both the narrow and the broad phenotype detected a two-dimensional MLS with 11q22 (discussed earlier) and 9q21. The 9q21 region has been linked to T2DM (Lindgren et al. 2002) and to other diabetes-related traits (Loos et al. 2003). Interestingly, the locus on chromosome 4p15 detected for the narrow phenotype in the one-dimensional analyses was also detected in the two-dimensional linkage study in pairs with five different loci. Among these five loci were 11q22 and 20q13 as well as 15q11–q14, which partly overlaps with the 15q13–q21 region [which has been reported as linked to T2DM in the Japanese population (Mori et al. 2002)], and 19q13, which has been linked to glucose, insulin, and insulin resistance in a meta-analysis of multiethnic populations (An et al. 2005) and has been linked to T2DM in independent studies (Majer et al. 1996; Loos et al. 2003; van Tilburg et al. 2003). Both chromosome 4p15 and 11q22 were seen in pairs with markers located in 1p22–p21. To our knowledge, no previously reported diabetes susceptibility locus is located in this region on chromosome 1p. Nor have we found any previous reports showing susceptibility to chromosome 2p15, which in this study was detected in association with the commonly reported diabetes locus on 12q24, described earlier. Furthermore, 5q14 was detected for the broad phenotype in pair with 14q21. Chromosome 5q14 has previously been reported as linked to diabetes-related phenotypes such as body mass index, fat mass, and percent body fat (Chen et al. 2005), and 14q21 has previously been suggested as linked to T2DM (Wiltshire et al. 2004).

Recently, a number of genome-wide association studies have been completed (Saxena et al. 2007; Scott et al. 2007; Sladek et al. 2007; Steinthorsdottir et al. 2007; Zeggini et al. 2007). These studies have partly overlapping results and report nine replicated susceptibility loci, all with relatively common risk alleles, with potential candidate genes for T2DM. However, according to Zeggini et al. (2007), these nine loci explain only a minor proportion of the observed familiarity. The fact that none of these regions overlap with any of the chromosome regions discussed here further supports the complexity of diabetes-related

traits. The linkage strategy used in the present study is heavily underpowered, so we could not detect such small genetic effects for such complex traits; the linkage strategy has high power to detect only genes of relatively large effect with relatively rare susceptibility alleles.

Susceptibility Loci for Adult Height

Human adult height has been suggested to be highly heritable, with heritability estimates (h^2) ranging from 0.5 to 0.9 (Mukhopadhyay and Weeks 2003; Willemsen et al. 2004) in various populations and with an h^2 of 0.7 in the Samoan islands (unpublished data). Our simulation study shows that, if the heritability that is explained by a single trait locus ranges from 0.7 to 0.175, the power to detect a linkage signal (genome-wide $\alpha = 0.05$) to adult height in the current study sample ranges from 31% to 14%. These fairly low power estimates suggest that additional chromosome regions not detected in the current study might be of great importance for variation in adult height in the studied population. Since the first genome-wide linkage scans for height were performed in 2001 (Hirschhorn et al. 2001; Perola et al. 2001), several susceptibility loci for height have been reported [for a recent overview see Perola et al. (2007)], but, as for most complex phenotypes, only some have been successfully replicated by further studies (Liu et al. 2004, 2006; Willemsen et al. 2004; Sale et al. 2005; Shmulewitz et al. 2006). The most promising susceptibility loci for height (Score.Max $p = 0.005$) detected in our current study of sib-pairs from American Samoa was located on chromosome 9q31 at marker *D9S1690*. This region and its flanking regions have previously been reported as linked to height in different populations. Perola et al. (2001) reported a QTL for height with a maximum LOD score of 2.61 located telomeric of our findings at 9q34, and Liu et al. (2004) reported three loci of interest on the long arm of chromosome 9. Their highest linkage peak (LOD score 2.74, $p = 0.0003$) was detected on 9q22. Furthermore, they detected a second peak (LOD score 2.22, $p = 0.001$) on 9q33 and a third peak (LOD score 2.66, $p = 0.0004$) on 9q34. In addition, Liu et al. (2006) confirmed the QTL for height on chromosome 9q22 in an extended data set. Based on our results and on those of the studies previously discussed, we suggest that chromosome region 9q22–qtel hosts at least one, possibly more, QTLs that are of importance for variation in adult height in populations of various ethnicities.

The region on chromosome 22q13 that in this study shows linkage (Score. Max $p = 0.008$) to adult height has previously been reported as linked to adult height in men originating from Framingham, Massachusetts (Mukhopadhyay et al. 2003) and in a combined sample of men and women in a Finnish study (Sammalisto et al. 2005). The additional two regions on chromosomes 1q44 and 10p15 that showed significant linkage with a p value less than 0.01 (Score.Max $p = 0.009$ and 0.010, respectively) in the present study have to our knowledge not been suggested as QTLs for variation in adult height previously. These loci might be more important for variation in adult height in our fairly isolated population from American Samoa than for more outbred cosmopolitan populations. Furthermore, we detected eight chromosome regions with p values ranging from 0.05 to 0.01 on chromosomes 2p25, 2p23, 2q32–q33, 10q26, 12q21, 12q23, 12q24, and 15q11. No previous genome-wide linkage studies have reported significant linkage for adult height to these loci. However, the region on chromosome 15q11 is suggested as a susceptibility region for Prader-Willi syndrome, a complex multifactorial syndrome with short stature as one of its major clinical features (Gunay-Aygun et al. 2001), which supports the finding that a locus regulating stature might be located in this chromosome region.

Conclusions

In conclusion, we did not detect any genome-wide significant susceptibility loci for T2DM within our study sample from American Samoa. However, our two-dimensional linkage investigation detected several loci pairs that are of potential interest. Interestingly, most

individual loci detected in the two-dimensional study have previously been suggested as susceptibility loci for diabetes-related traits. Thus the present study supports the theory that T2DM is a complex disease that is influenced by multiple genes or gene clusters that most likely one by one contribute with a minor effect but together create the genetic susceptibility for the T2DM phenotype. We believe that our two-dimensional linkage results can be helpful when selecting potential candidate genes and molecular pathways for further studies, because these results, in contrast to one-dimensional linkage results, also contribute information regarding multigenic effects.

We did not find any overlap of interest between the loci detected for the diabetes-related traits and the loci detected for adult height. We therefore do not believe that there is any detectable locus with major pleiotropic effect acting simultaneously on diabetes-related traits and adult height in our sample. Interestingly, the most promising chromosome region detected for variation of adult height in this Samoan sample is located on chromosome 9q31, which has been reported previously as linked to adult height in other populations. We therefore argue that the region on chromosome 9q31 is a strong susceptibility locus for variation in adult height, which is likely to be of importance across populations.

Acknowledgments

This work was supported by the National Institutes of Health through grants AG09375, HL52611, DK55406, and DK59642. K. Åberg is supported by the Swedish Research Council (through grant 2006–4274). We would like to thank Eleanor Feingold for describing the Score.Max statistic. We are very thankful to the Department of Health, Government of American Samoa, for their support, to the local political officials for their permission to work in the villages, and to the study participants for their contribution to this research.

Literature Cited

Abecasis GR, Cherny SS, Cookson WO, et al. Merlin: Rapid analysis of dense genetic maps using sparse gene flow trees. Nat Genet. 2002; 30:97–101. [PubMed: 11731797]

Aberg, K.; Dai, F.; Sun, G., et al. Susceptibility loci for adiposity-related phenotypes on 8p, 9p, and 16q in a combined sample set from American Samoa and Samoa. 2008. Unpublished

An P, Freedman BI, Hanis CL, et al. Genome-wide linkage scans for fasting glucose, insulin, and insulin resistance in the National Heart, Lung, and Blood Institute Family Blood Pressure Program: Evidence of linkages to chromosome 7q36 and 19q13 from meta-analysis. Diabetes. 2005; 54:909–914. [PubMed: 15734873]

Barroso I. Genetics of type 2 diabetes. Diabet Med. 2005; 22:517–535. [PubMed: 15842505]

Bell JT, Wallace C, Dobson R, et al. Two-dimensional genome-scan identifies novel epistatic loci for essential hypertension. Hum Mol Genet. 2006; 15:1365–1374. [PubMed: 16543358]

Bento JL, Palmer ND, Mychaleckyj JC, et al. Association of protein tyrosine phosphatase 1B gene polymorphisms with type 2 diabetes. Diabetes. 2004; 53:3007–3012. [PubMed: 15504984]

Bergholdt R, Nerup J, Pociot F. Fine mapping of a region on chromosome 21q21.11–q22.3 showing linkage to type 1 diabetes. J Med Genet. 2005; 42:17–25. [PubMed: 15635070]

Bhattacharjee S, Kuo CL, Mukhopadhyay N, et al. Robust score statistics for QTL linkage analysis. Am J Hum Genet. 2008; 82:567–582. [PubMed: 18304491]

Boehnke M, Cox NJ. Accurate inference of relationships in sib-pair linkage studies. Am J Hum Genet. 1997; 61:423–429. [PubMed: 9311748]

Broman KW. Estimation of allele frequencies with data on sibships. Genet Epidemiol. 2001; 20:307–315. [PubMed: 11255240]

Carulli L, Rondinella S, Lombardini S, et al. Review article: Diabetes, genetics, and ethnicity. Aliment Pharmacol Ther. 2005; 22(suppl 2):16–19. [PubMed: 16225465]

Chen G, Adeyemo AA, Johnson T, et al. A genome-wide scan for quantitative trait loci linked to obesity phenotypes among West Africans. Int J Obes. 2005; 29:255–259.

Cheyssac C, Lecoeur C, Dechaume A, et al. Analysis of common *PTPN1* gene variants in type 2 diabetes, obesity, and associated phenotypes in the French population. BMC Med Genet. 2006; 7:44. [PubMed: 16677372]

Cornes BK, Medland SE, Ferreira MA, et al. Sex-limited genome-wide linkage scan for body mass index in an unselected sample of 933 Australian twin families. Twin Res Hum Genet. 2005; 8:616–632. [PubMed: 16363087]

Cottingham RW Jr, Idury RM, Schaffer AA. Faster sequential genetic linkage computations. Am J Hum Genet. 1993; 53:252–263. [PubMed: 8317490]

Dai F, Keighley ED, Sun G, et al. Genome-wide scan for adiposity-related phenotypes in adults from American Samoa. Int J Obes. 2007; 31:1832–1842.

Dai, F.; Sun, GY.; Aberg, K., et al. A whole genome linkage scan identifies multiple chromosome regions influencing adiposity-related traits among Samoans. 2008. Unpublished

Damcott CM, Hoppman N, Ott SH, et al. Polymorphisms in both promoters of hepatocyte nuclear factor 4-alpha are associated with type 2 diabetes in the Amish. Diabetes. 2004; 53:3337–3341. [PubMed: 15561969]

Deka R, McGarvey ST, Ferrell RE, et al. Genetic characterization of American and Western Samoans. Hum Biol. 1994; 66:805–822. [PubMed: 8001911]

Dong GP, Yu ZS, Liang L, et al. IL-18 gene promoter –137C/G and –607C/A polymorphisms in Chinese Han children with type 1 diabetes mellitus. Int J Immunogenet. 2007; 34:75–79. [PubMed: 17373930]

Epstein MP, Duren WL, Boehnke M. Improved inference of relationship for pairs of individuals. Am J Hum Genet. 2000; 67:1219–1231. [PubMed: 11032786]

Fava S, Azzopardi J, Hattersley AT, et al. Increased prevalence of proteinuria in diabetic sibs of proteinuric type 2 diabetic subjects. Am J Kidney Dis. 2000; 35:708–712. [PubMed: 10739793]

Fava S, Azzopardi J, Watkins PJ, et al. Adult height and proteinuria in type 2 diabetes. Nephrol Dial Transplant. 2001; 16:525–528. [PubMed: 11239026]

Genin E, Clerget-Darpoux F. Consanguinity and the sib-pair method: An approach using identity by descent between and within individuals. Am J Hum Genet. 1996; 59:1149–1162. [PubMed: 8900245]

Gudbjartsson DF, Jonasson K, Frigge ML, et al. Allegro, a new computer program for multipoint linkage analysis. Nat Genet. 2000; 25:12–13. [PubMed: 10802644]

Gudbjartsson DF, Thorvaldsson T, Kong A, et al. Allegro, version 2. Nat Genet. 2005; 37:1015–1016. [PubMed: 16195711]

Gunay-Aygun M, Schwartz S, Heeger S, et al. The changing purpose of Prader-Willi syndrome clinical diagnostic criteria and proposed revised criteria. Pediatrics. 2001; 108:E92. [PubMed: 11694676]

Hirschhorn JN, Lindgren CM, Daly MJ, et al. Genomewide linkage analysis of stature in multiple populations reveals several regions with evidence of linkage to adult height. Am J Hum Genet. 2001; 69:106–116. [PubMed: 11410839]

Hodge SE. The information contained in multiple sibling pairs. Genet Epidemiol. 1984; 1:109–122. [PubMed: 6599401]

Keighley, ED.; McGarvey, ST.; Quested, C., et al. Nutrition and health in modernizing Samoans: Temporal trends and adaptive perspectives. In: Ohtsuka, R.; Ulijaszek, SJ., editors. Health Changes in the Asia-Pacific Region: Bio-cultural and Epidemiological Approaches. Cambridge, U.K: Cambridge University Press; 2007. p. 147-199.

Klupa T, Malecki MT, Pezzolesi M, et al. Further evidence for a susceptibility locus for type 2 diabetes on chromosome 20q13.1–q13.2. Diabetes. 2000; 49:2212–2216. [PubMed: 11118028]

Kong X, Murphy K, Raj T, et al. A combined linkage-physical map of the human genome. Am J Hum Genet. 2004; 75:1143–1148. [PubMed: 15486828]

Lakka TA, Rankinen T, Weisnagel SJ, et al. A quantitative trait locus on 7q31 for the changes in plasma insulin in response to exercise training: The HERITAGE Family Study. Diabetes. 2003; 52:1583–1587. [PubMed: 12765975]

Lange K, Cantor R, Horvath S, et al. Mendel version 4.0: A complete package for the exact genetic analysis of discrete traits in pedigree and population data sets. Am J Hum Genet. 2001; 69:A1886.

Leutenegger AL, Genin E, Thompson EA, et al. Impact of parental relationships in maximum lod score affected sib-pair method. Genet Epidemiol. 2002; 23:413–425. [PubMed: 12432507]

Lewis CE, North KE, Arnett D, et al. 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. 2005; 29:639–649.

Lindgren CM, Mahtani MM, Widen E, et al. Genomewide search for type 2 diabetes mellitus susceptibility loci in Finnish families: The Botnia study. Am J Hum Genet. 2002; 70:509–516. [PubMed: 11791216]

Liu YZ, Xiao P, Guo YF, et al. Genetic linkage of human height is confirmed to 9q22 and Xq24. Hum Genet. 2006; 119:295–304. [PubMed: 16446976]

Liu YZ, Xu FH, Shen H, et al. Genetic dissection of human stature in a large sample of multiplex pedigrees. Ann Hum Genet. 2004; 68:472–488. [PubMed: 15469424]

Lohman, TG.; Roche, AF.; Martorell, R., editors. Anthropometric Standardization Reference Manual. Champaign, IL: Human Kinetics Books; 1988.

Loos RJ, Katzmarzyk PT, Rao DC, et al. Genome-wide linkage scan for the metabolic syndrome in the HERITAGE Family Study. J Clin Endocrinol Metab. 2003; 88:5935–5943. [PubMed: 14671193]

Majer M, Mott DM, Mochizuki H, et al. Association of the glycogen synthase locus on 19q13 with NIDDM in Pima Indians. Diabetologia. 1996; 39:314–321. [PubMed: 8721777]

McGarvey ST. Cardiovascular disease (CVD) risk factors in Samoa and American Samoa, 1990–95. Pac Health Dialog. 2001; 8:157–162. [PubMed: 12017817]

McPeek MS, Sun L. Statistical tests for detection of misspecified relationships by use of genome-screen data. Am J Hum Genet. 2000; 66:1076–1094. [PubMed: 10712219]

Mori Y, Otabe S, Dina C, et al. Genome-wide search for type 2 diabetes in Japanese affected sib-pairs confirms susceptibility genes on 3q, 15q, and 20q and identifies two new candidate loci on 7p and 11p. Diabetes. 2002; 51:1247–1255. [PubMed: 11916952]

Mukhopadhyay N, Almasy L, Schroeder M, et al. Mega2: Data-handling for facilitating genetic linkage and association analyses. Bioinformatics. 2005; 21:2556–2557. [PubMed: 15746282]

Mukhopadhyay, N.; Bhattacharjee, S.; Kuo, C-L., et al. QTL-ALL (Quantitative Trait Locus—Analysis and Linkage Library): Software for human QTL linkage analysis. 2008. Unpublished

Mukhopadhyay N, Buxbaum SG, Weeks DE. Comparative study of multipoint methods for genotype error detection. Hum Hered. 2004; 58:175–189. [PubMed: 15812175]

Mukhopadhyay N, Finegold DN, Larson MG, et al. A genome-wide scan for loci affecting normal adult height in the Framingham Heart Study. Hum Hered. 2003; 55:191–201. [PubMed: 14566097]

Mukhopadhyay N, Weeks DE. Linkage analysis of adult height with parent-of-origin effects in the Framingham Heart Study. BMC Genet. 2003; 4(suppl 1):S76. [PubMed: 14975144]

North KE, Franceschini N, Borecki IB, et al. Genotype-by-sex interaction on fasting insulin concentration: The HyperGEN study. Diabetes. 2007; 56:137–142. [PubMed: 17192475]

O'Connell JR, Weeks DE. PedCheck: A program for identification of genotype incompatibilities in linkage analysis. Am J Hum Genet. 1998; 63:259–266. [PubMed: 9634505]

Ott J. Computer-simulation methods in human linkage analysis. Proc Natl Acad Sci USA. 1989; 86:4175–4178. [PubMed: 2726769]

Perola M, Ohman M, Hiekkalinna T, et al. Quantitative trait locus analysis of body-mass index and of stature, by combined analysis of genome scans of five Finnish study groups. Am J Hum Genet. 2001; 69:117–123. [PubMed: 11410840]

Perola M, Sammalisto S, Hiekkalinna T, et al. Combined genome scans for body stature in 6,602 European twins: Evidence for common Caucasian loci. PLoS Genet. 2007; 3:e97. [PubMed: 17559308]

Perusse L, Rice T, Chagnon YC, et al. A genome-wide scan for abdominal fat assessed by computed tomography in the Quebec Family Study. Diabetes. 2001; 50:614–621. [PubMed: 11246882]

Pettitt DJ, Saad MF, Bennett PH, et al. Familial predisposition to renal disease in two generations of Pima Indians with type 2 (non-insulin-dependent) diabetes mellitus. Diabetologia. 1990; 33:438–443. [PubMed: 2401399]

Putter H, Sandkuijl LA, van Houwelingen JC. Score test for detecting linkage to quantitative traits. Genet Epidemiol. 2002; 22:345–355. [PubMed: 11984866]

Rotimi CN, Chen G, Adeyemo AA, et al. A genome-wide search for type 2 diabetes susceptibility genes in West Africans: The Africa America Diabetes Mellitus (AADM) Study. Diabetes. 2004; 53:838–841. [PubMed: 14988271]

Sale MM, Freedman BI, Hicks PJ, et al. Loci contributing to adult height and body mass index in African American families ascertained for type 2 diabetes. Ann Hum Genet. 2005; 69:517–527. [PubMed: 16138910]

Sammalisto S, Hiekkalinna T, Suviolahti E, et al. A male-specific quantitative trait locus on 1p21 controlling human stature. J Med Genet. 2005; 42:932–939. [PubMed: 15827092]

Saxena R, Voight BF, Lyssenko V, et al. Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. Science. 2007; 316(5829):1331–1336. [PubMed: 17463246]

Scott LJ, Mohlke KL, Bonnycastle LL, et al. A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. Science. 2007; 316(5829):1341–1345. [PubMed: 17463248]

Sengul H, Bhattacharjee S, Feingold E, et al. The elusive goal of pedigree weights. Genet Epidemiol. 2007; 31:51–65. [PubMed: 17086515]

Sengul H, Weeks DE, Feingold E. A survey of affected-sibship statistics for nonparametric linkage analysis. Am J Hum Genet. 2001; 69:179–190. [PubMed: 11404816]

Shmulewitz D, Heath SC, Blundell ML, et al. Linkage analysis of quantitative traits for obesity, diabetes, hypertension, and dyslipidemia on the island of Kosrae, Federated States of Micronesia. Proc Natl Acad Sci USA. 2006; 103:3502–3509. [PubMed: 16537441]

Silander K, Mohlke KL, Scott LJ, et al. Genetic variation near the hepatocyte nuclear factor-4 alpha gene predicts susceptibility to type 2 diabetes. Diabetes. 2004; 53:1141–1149. [PubMed: 15047633]

Sladek R, Rocheleau G, Rung J, et al. A genome-wide association study identifies novel risk loci for type 2 diabetes. Nature. 2007; 445:881–885. [PubMed: 17293876]

Steinthorsdottir V, Thorleifsson G, Reynisdottir I, et al. A variant in *CDKAL1* influences insulin response and risk of type 2 diabetes. Nat Genet. 2007; 39(6):770–775. [PubMed: 17460697]

Stone S, Abkevich V, Hunt SC, et al. A major predisposition locus for severe obesity, at 4p15–p14. Am J Hum Genet. 2002; 70:1459–1468. [PubMed: 11957135]

Sun L, Wilder K, McPeek MS. Enhanced pedigree error detection. Hum Hered. 2002; 54:99–110. [PubMed: 12566741]

Tang HK, Siegmund D. Mapping quantitative trait loci in oligogenic models. Biostatistics. 2001; 2:147–162. [PubMed: 12933546]

Tessier MC, Qu HQ, Frechette R, et al. Type 1 diabetes and the OAS gene cluster: Association with splicing polymorphism or haplotype? J Med Genet. 2006; 43:129–132. [PubMed: 16014697]

Traurig MT, Permana PA, Nair S, et al. Differential expression of matrix metalloproteinase 3 (MMP3) in preadipocytes/stromal vascular cells from nonobese nondiabetic versus obese nondiabetic Pima Indians. Diabetes. 2006; 55:3160–3165. [PubMed: 17065356]

Tsai HJ, Sun G, Smelser D, et al. Distribution of genome-wide linkage disequilibrium based on microsatellite loci in the Samoan population. Hum Genomics. 2004; 1:327–334. [PubMed: 15588493]

Tsai HJ, Sun G, Weeks DE, et al. Type 2 diabetes and three calpain-10 gene polymorphisms in Samoans: No evidence of association. Am J Hum Genet. 2001; 69:1236–1244. [PubMed: 11704924]

van Tilburg JH, Sandkuijl LA, Strengman E, et al. A genome-wide scan in type 2 diabetes mellitus provides independent replication of a susceptibility locus on 18p11 and suggests the existence of novel loci on 2q12 and 19q13. J Clin Endocrinol Metab. 2003; 88:2223–2230. [PubMed: 12727978]

Wang K, Huang J. A score-statistic approach for the mapping of quantitative-trait loci with sibships of arbitrary size. Am J Hum Genet. 2002; 70:412–424. [PubMed: 11791211]

Weeks DE, Ott J, Lathrop GM. SLINK: A general simulation program for linkage analysis. Am J Hum Genet. 1990; 47:A204.

Whittemore AS, Halpern J. A class of tests for linkage using affected pedigree members. Biometrics. 1994; 50:118–127. [PubMed: 8086596]

Willemsen G, Boomsma DI, Beem AL, et al. QTLs for height: Results of a full genome scan in Dutch sibling pairs. Eur J Hum Genet. 2004; 12:820–828. [PubMed: 15305175]

Wiltshire S, Cardon LR, McCarthy MI. Evaluating the results of genomewide linkage scans of complex traits by locus counting. Am J Hum Genet. 2002; 71:1175–1182. [PubMed: 12355401]

Wiltshire S, Frayling TM, Groves CJ, et al. Evidence from a large U.K. family collection that genes influencing age of onset of type 2 diabetes map to chromosome 12p and to the *MODY3* NIDDM2 locus on 12q24. Diabetes. 2004; 53:855–860. [PubMed: 14988275]

Zeggini E, Weedon MN, Lindgren CM, et al. Replication of genome-wide association signals in U.K. samples reveals risk loci for type 2 diabetes. Science. 2007; 316(5829):1336–1341. [PubMed: 17463249]

Zimmet P, Alberti KG, Shaw J. Global and societal implications of the diabetes epidemic. Nature. 2001; 414:782–787. [PubMed: 11742409]

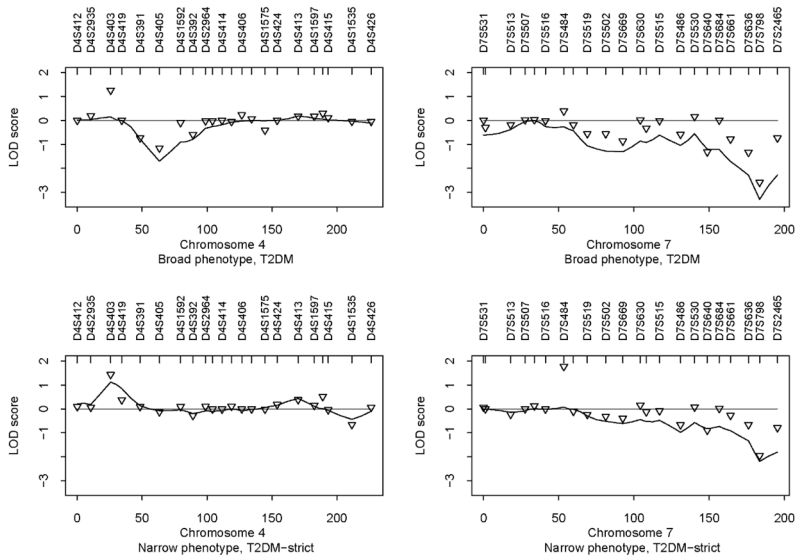
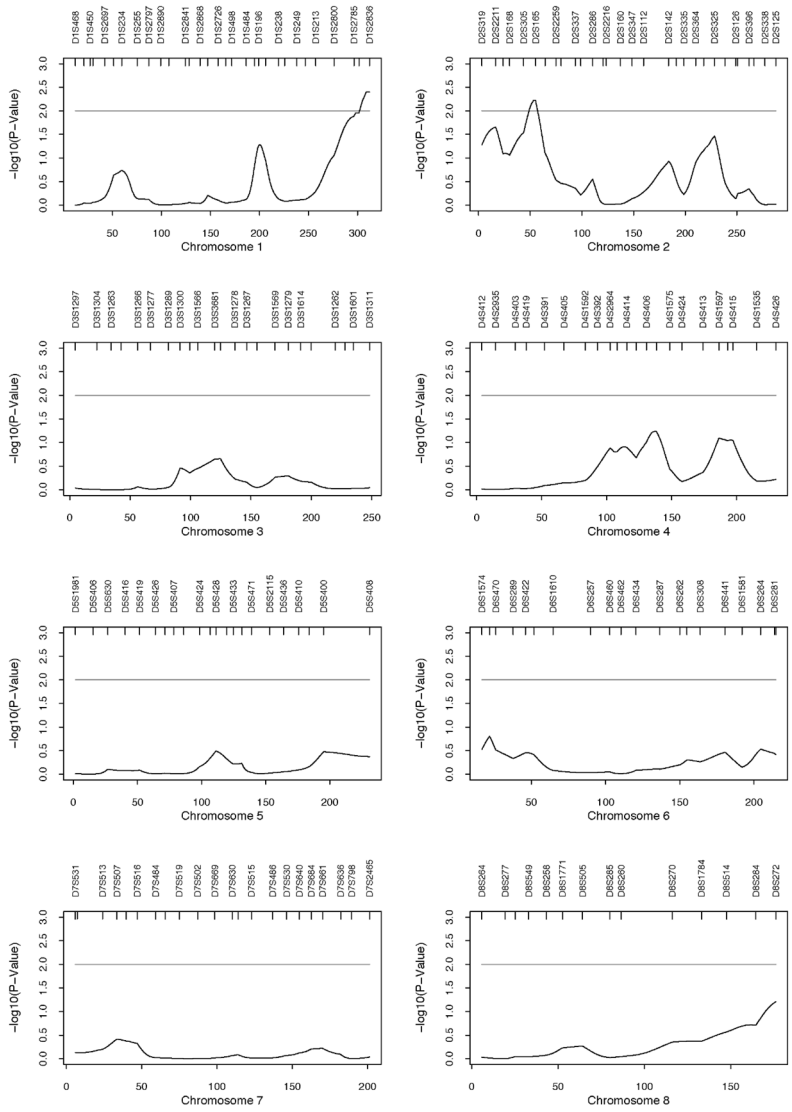
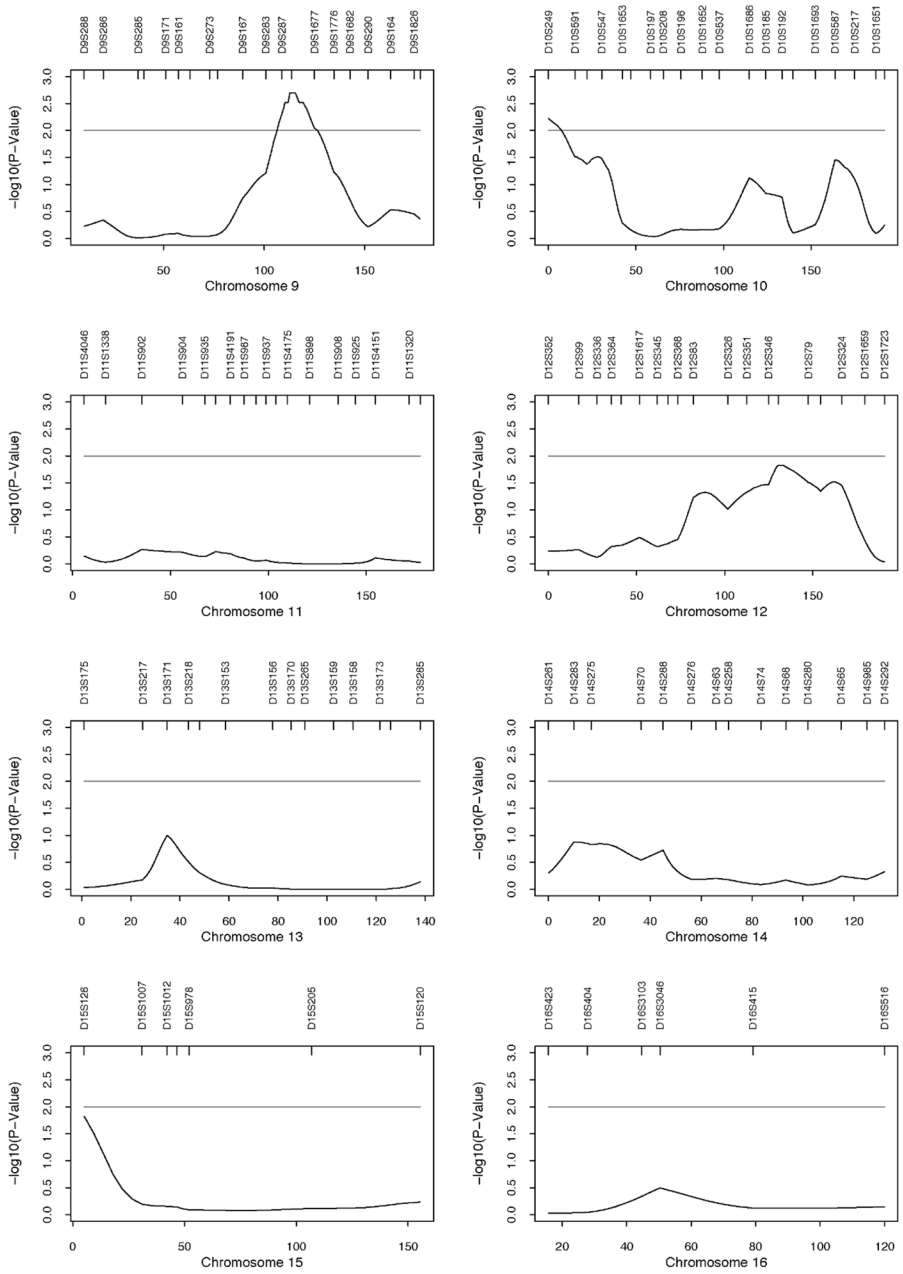


Figure 1.

One-dimensional linkage results for chromosomes 4 and 7 for the conservative data set. Multipoint results are illustrated by a solid line, and single-point results are indicated by triangles.





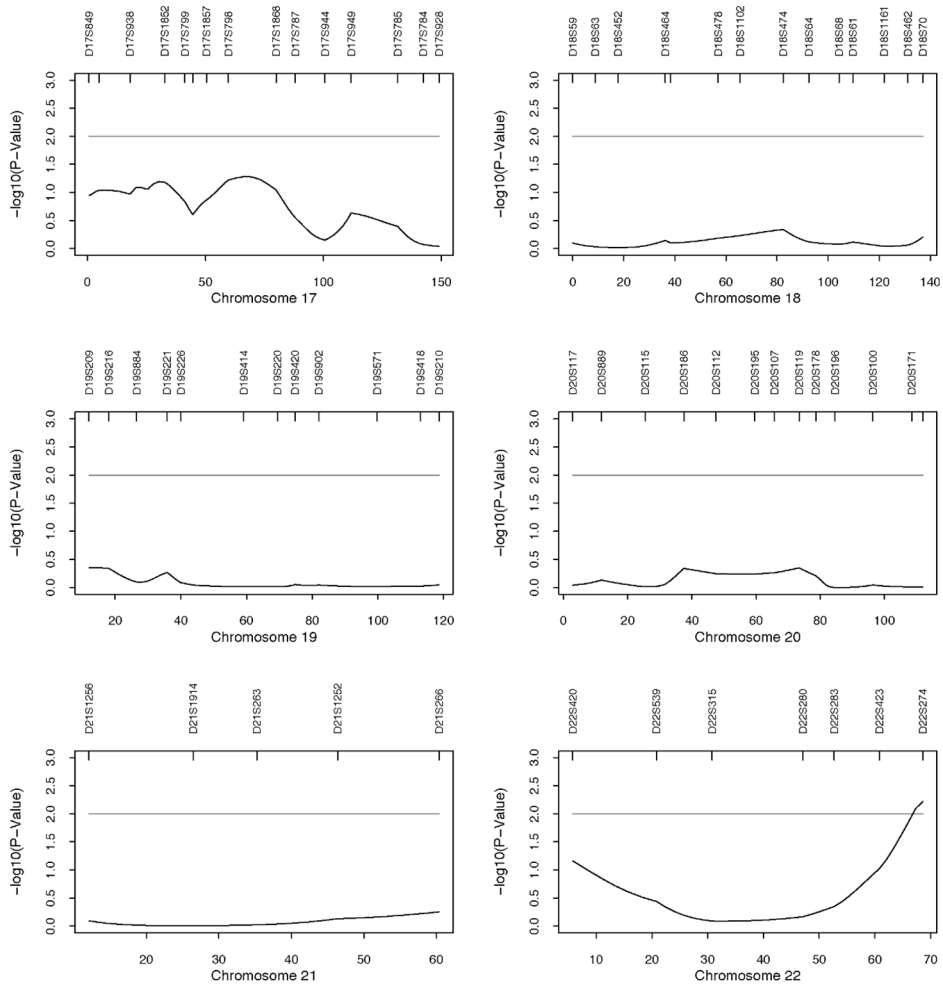


Figure 2.

Genome-wide linkage results, −log10(*p* value) from the Score.Max statistic, for adult height in full sib pairs (22–55 years old) from the conservative data set. Horizontal line indicates $p = 0.01$.

Table 1

Number of Genotyped Samples per Data Set

Phenotype	Data Set	Number of Genotyped Families[a]	Number of Genotyped Individuals	Number of Genotyped Full Sib Pairs	Number of Genotyped Half-Sib Pairs	Number of Genotyped First-Cousin Pairs
T2DM	Collected	174	386	299	1	0
T2DM	Alternative	131	311	238	4	8
T2DM	Conservative	124	296	238	1	0
T2DM-strict	Collected	168	314	196	1	0
T2DM-strict	Alternative	128	250	151	3	8
T2DM-strict	Conservative	121	236	151	1	0
Stature	Conservative[b]	72	122	63	0	0

[a] Families with two or more individuals genotyped.

[b] Counts for individuals from 22 to 55 years old.

Table 2

General Characteristics of Phenotyped and Genotyped Participants

		Alternative Data Set			Conservative Data Set		
Phenotype	**Characteristic**	**All**	**Males**	**Females**	**All**	**Males**	**Females**
T2DM	Number of individuals	311	154	157	296	145	151
	Males/females (%)	49.5/50.5			49.0/51.0		
	Mean age ± SD	54.4 ± 11.7	56.2 ± 11.0	52.2 ± 12.2	54.7 ± 11.7	56.3 ± 11.1	52.4 ± 12.1
T2DM-strict	Number of individuals	250	123	127	236	115	121
	Males/females (%)	49.2/50.8			48.7/51.3		
	Mean age ± SD	55.2 ± 11.7	56.8 ± 10.9	53.7 ± 12.3	55.7 ± 11.6	56.8 ± 11.1	53.5 ± 12.1
Stature	Number of individuals	Data set not investigated			122	51	71
	Males/females (%)	Data set not investigated			41.8/58.2		
	Mean age ± SD	Data set not investigated			45.0 ± 7.2	46.6 ± 6.6	44.0 ± 7.3
	Mean stature ± SD	Data set not investigated			168 ± 8.3	174.6 ± 6.5	164.1 ± 6.9

Table 3

Locus Combinations with Two-Dimensional MLSs Greater Than or Equal to 1.6

	Locus 1		Locus 2		
Phenotype	**Cytogenetic Position**	**Marker**	**Cytogenetic Position**	**Marker**	**MLS**
T2DM-strict	1p22	D1S2868	11q22	D11S898	1.69
T2DM-strict	1p22–p21	D1S2868, D1S206[a]	4p15	D4S403,[a] D4S419	2.05
T2DM	2p15	D2S337	12q24	D2S337	1.64
T2DM-strict	4p15	D4S403	11q22	D11S898	1.60
T2DM-strict	4p15	D4S403	15q11–q14	D15S128,[a] D15S1007	2.01
T2DM-strict	4p15	D4S403,[a] D4S419	19q13	D19S902	1.80
T2DM-strict	4p15	D4S403,[a] D4S419	20q13	D20S100, D20S171,[a] D20S173	1.87
T2DM	5q14	D5S428	14q21	D14S288	1.89
T2DM	9q21	D9S273, D9S175[a]	11q22	D11S898	1.81
T2DM-strict	9q21	D9S273, D9S175[a]	11q22	D11S898	2.07
T2DM	11q22	D11S898	20q13	D20S171	1.67
T2DM-strict	11q22	D11S898	21q21	D21S1256	2.27
T2DM	12q24	D12S1723	20q13	D20S171,[a] D20S173	1.94

[a]Marker that generated the largest MLS.

Hum Biol. Author manuscript; available in PMC 2013 July 04.

Table 4

Susceptibility Regions for Adult Height with Score.Max *p* Values Less Than or Equal to 0.05[a]

Cytogenetic Position[b]	Closest Marker	Score.Max p Value
1q44	**D1S2836**	**0.009**
2p25	D2S2211	0.028
2p23	D2S165	0.014
2q32–q33	D2S117–D2S325	0.042
9q31	**D9S1690**	**0.005**
10p15	**D10S249**	**0.010**
10q26	D10S587–D10S217	0.044
12q21	D12S351	0.026
12q23	D12S78	0.019
12q24	D12S324	0.043
15q11	D15S128	0.018
22q13	**D22S274**	**0.008**

[a] Regions with a *p* value less than or equal to 0.01 are marked in bold.

[b] Cytogenetic position for closest marker(s).