A Genome-Wide Scan for Obesity in African-Americans

Xiaofeng Zhu,[1] Richard S. Cooper,[1] Amy Luke,[1] Guanjie Chen,[1] Xiaodong Wu,[1] Donghui Kan,[1] Aravinda Chakravarti,[2] and Alan Weder[3]

A genome-wide scan using 387 short tandem repeat markers was conducted for obesity among 618 black individuals from 202 families residing in a suburb of Chicago. Evidence for linkage was evaluated with BMI and percent body fat (PBF) using a variance component analysis approach. Suggestive evidence for linkage was found for BMI on chromosome 5 (logarithm of odds [LOD] score = 1.9) and PBF on chromosome 6 (LOD score = 2.7). One additional region on chromosome 3 was linked to these phenotypes at a lower level of significance (LOD score = 1.8 and 0.95 for BMI and PBF, respectively); the linked marker on this chromosome lies in the same region implicated as harboring obesity genes in a previous study of a white population. The replication of linkage evidence using different ethnic groups reinforces the potential significance of this latter candidate region. *Diabetes* 51:541–544, 2002

The technical capacity to examine the genetic underpinnings of obesity has moved forward rapidly in the last several years (1–3). Despite these advances, only a small proportion of the genetic susceptibility to obesity can be currently explained. A major weakness remains the absence of consistent findings from epidemiologic studies. Whereas studies of candidate genes whose functional significance can be identified will ultimately be required, whole-genome search methods are attractive because they allow assessment of all potential regions independently of prior hypotheses. Unfortunately, linkage has not proven to be an efficient method for the study of complex traits like obesity (4). Given the large potential for both type I and type II errors that are inherent in this method, consistency of findings in multiple independent samples must be a primary criterion on which to judge their importance. Genome search methods could potentially play a role in this process of establishing reproducibility and consistency of findings. For example, evidence of linkage of leptin levels to a region on chromosome 2 in Mexican Americans by Comuzzie and colleagues (5,6) was subsequently confirmed using the same marker among African-Americans. In this article, we describe results of a genome-wide scan in a population-based sample of U.S. blacks that are, to some degree, consistent with prior data.

From the [1]Department of Preventive Medicine and Epidemiology, Loyola University Medical School, Maywood, Illinois; [2]McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University, Baltimore, Maryland; and the [3]Division of Hypertension, Department of Medicine, University of Michigan School of Medicine, Ann Arbor, Michigan.

Address correspondence and reprint requests to Xiaofeng Zhu, Department of Preventive Medicine and Epidemiology, Loyola University Medical School, Maywood, IL 60153. E-mail: xzhu1@luc.edu.

Received for publication 26 December 2000 and accepted in revised form 26 October 2001.

IBD, identity-by-descent; LOD, logarithm of odds; MLS, maximum LOD score; PBF, percent body fat; QTL, quantitative trait locus.

RESEARCH DESIGN AND METHODS

Recruitment of participants. Participants were recruited as part of the GenNet network of the Family Blood Pressure Program (7). The field center is based in Maywood, Illinois, a predominantly African-American working class suburb of Chicago, as has been previously described (6,7). Eligibility requirements included age between 25 and 45 years and a blood pressure in the upper 25th percentile of the age-sex–specific distribution for this community. No restrictions were imposed in terms of weight or diet history. A clinical examination was conducted that included a medical history, anthropometric measurements, and blood sampling. Weight was measured in light clothes on an electronic scale, and height was measured with a stadiometer. Body composition was determined using bioelectric impedance analysis (RJL, Township, MI) to estimate percent body fat (PBF). The correlation between BMI and PBF was 0.79. The data set included 232 parent-offspring pairs, 332 siblings, and 197 half-siblings.

Genotyping. A set of 387 short tandem repeat markers (Set 9, www.marshmed.org/genetics) was genotyped by the Mammalian Genotyping Service in Marshfield, Wisconsin (8). The mean heterozygosity for these markers is ~76%, with an average sex-equal distance of 10 cM (8). DNA samples were provided in two batches, and extensive quality checks were carried out to verify consistency of marker genotyping and stated pedigree relationships. First, PedCheck (9) was used to check the mendelian inconsistency. ASPEX/KINSHIP (10) was then used to check the biological relationships among the pedigrees where significant mendelian inconsistencies were found. ASPEX/KINSHIP checks the degree of biological consistency within families using available markers and classifies the biological relationship for each sibpair based on the likelihood method. The relationships were reassigned if they were clearly misclassified. The errors identified in PedCheck not based on misclassification of pedigrees were assumed to have occurred in the genotyping process, and the associated markers were set to missing among the appropriate family members.

Statistical analysis. Genome-wide linkage analysis of the obesity phenotypes was performed using the multipoint variance component program in GENEHUNTER2 (11). The variance component method in GENEHUNTER2 specifies the expected genetic covariances between relatives as a function of their identity-by-descent (IBD) relationships at a marker locus. The IBD probabilities were estimated from all available genotyped marker loci. The likelihood ratio test was applied to test the null hypothesis of no additive genetic variance due to a quantitative trait locus (QTL). Both siblings and half-siblings were used in the variance component method. Sex, age, and age^2 were incorporated as covariates, and their effects were simultaneously estimated by the maximum-likelihood method. In all analyses, allele frequencies were estimated from the marker data. We used the Marshfield map distances in the linkage analysis. The width of the candidate regions was defined by a logarithm of odds (LOD) score of >0.588, which corresponds to a P value of 0.05—a significant nominal level without considering multiple tests.

To assess the empirical significance of our result, we performed a simulation study. We retained the pedigree data and phenotype data and simulated the marker genotype data based on the observed marker allele frequencies. There were 200 replications generated and analyzed by GENEHUNTER2. To determine the power we could achieve from our data structure, we did simulations using a model with an additive quantitative trait locus explaining 20% of the variance, 30% of the polygenic variance, and 50% of the random environmental variance.

We next performed a quantitative transmission/disequilibrium test using the program QTDT (12) for those markers in the regions where linkage evidence was found. QTDT uses variance components to construct a test that

TABLE 1
Descriptive characteristics of GenNet participants

	Men		Women
Individuals (n)	250		368
Families (n)		201	
Sibpairs (n)		332	
Half-sibpairs (n)		197	
Age (years)	38.5 ± 10.2*		41.6 ± 12.4
Height (cm)	176.9 ± 7.0*		164.0 ± 7.3
Weight (kg)	84.1 ± 20.7		87.0 ± 23.8
BMI (kg/m^2)	26.9 ± 6.4*		32.4 ± 8.7
BMI ≥30 (%)	24.0		54.1
PBF (%)	25.0 ± 8.5*		42.5 ± 8.1
Leptin	8.1 ± 10.4		26.3 ± 16.1
Blood pressure			
Systolic	129.0 ± 18.6*		125.8 ± 21.8
Diastolic	78.9 ± 14.5		77.8 ± 13.8

Data are means ± SE unless otherwise indicated. *Represents significant difference compared with women ($P < 0.05$).

uses information from all available offspring and avoids the false-positive result that could be due to the population structure.

RESULTS

The descriptive characteristics of the participants are presented in Table 1. Women were on average slightly older than men and more likely to be obese. The mean diastolic blood pressures of men and women were similar, although systolic pressures were higher for men. The narrow-sense heritabilities for BMI and PBF were 0.54 and 0.56, respectively, after adjusting for sex, age, and age^2 using an additive model in GENEHUNTER2, which uses all the pedigree relationships. Figure 1 presents the multipoint linkage results from GENEHUNTER2 for BMI and PBF across all chromosomes. The regions with linkage evidence are also presented in Table 2. In total, three regions with linkage evidence (maximum LOD score [MLS] ≥1.8) were found for at least one of the obesity phenotypes. The largest LOD score was found for PBF on a region around D6S1959 located on chromosome 6 (LOD score = 2.7, $P = 0.0002$). For BMI, the region at D5S817 on chromosome 5 showed suggestive linkage evidence (MLS = 1.9, $P = 0.0017$). We also found LOD = 1.8 ($P = 0.002$) to be associated with BMI on a region of chromosome 3 around D3S2477, with LOD = 0.95 ($P = 0.018$) for PBF on the same region. A LOD score of 1.0 ($P = 0.016$) for BMI was also found on both chromosome 2 (genomic region –60 to 85) and chromosome 10 (genomic region –69 to 88).

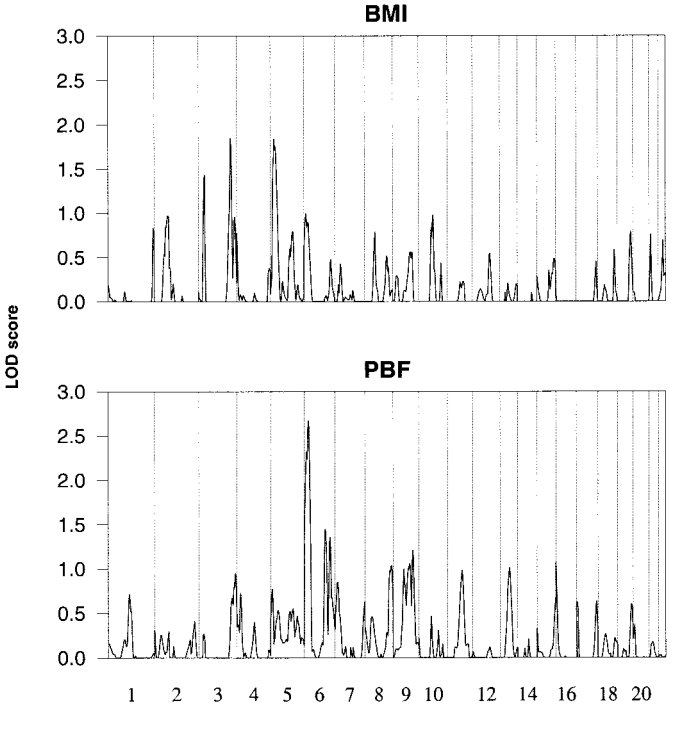

FIG. 1. Results from a genome wide scan of BMI and PBF using GENEHUNTER2. The numbers under the vertical axis represent the chromosomes.

We performed simulations to determine the corresponding empirical genomewide significance levels of P values for our results. In our simulations, we retained the pedigree structures, phenotypes, and covariates and simulated the marker data based on the observed allele frequencies. The genomewide empirical P values were calculated based on 200 replicate data sets. These P values are 0.15, 0.69, and 0.79 for LOD scores 2.7 of PBF on chromosome 6, and 1.9 and 1.8 of BMI on chromosome 3 and 5, respectively. Because the peak regions on chromosome 2, 3, and 5 were reported with significant linkage evidence in previous studies (5,13,14), we looked the significance levels adjusting multiple tests on these regions. Simulation results show the corresponding P values are as follows: 0.08 for chromosome 2 (region 60–85), 0.005 for chromosome 3 (region 172–197), and 0.025 for chromosome 5 (region 10–38).

To determine the empirical power of our sample to detect the presence of a modest QTL effect, we generated 500 data sets with the observed pedigree structure and assumed that the QTL explains 20% of the total variance with the remaining variance consisting of 30% polygenic variance and 50% environmental variance. In the presence of such a QTL, we found that 7% of the scans gave a LOD score >3.3.

We next performed a quantitative transmission/disequilibrium test using QTDT for the five markers in the region with linkage evidence on chromosome 3 for BMI. Table 3 summarizes the results for these alleles, with a P value <0.1 when tested against all the other alleles pooled together for BMI. Several alleles in this region were associated with a P value of <0.05.

DISCUSSION

During the past decade, the genome-wide search for a susceptibility locus has been widely applied for complex human diseases. Lander and Kruglyak (13) proposed criteria for declaring linkage to avoid false-positive evidence imposed by the use of LOD scores. However, the success of the genome-wide search strategy to date has been limited, and the results of the studies focused on the same phenotype are often inconsistent (15). Therefore, consistency across different studies would seem to be an even more important standard against which to judge these results. Accordingly, we draw attention to our finding on chromosome 3 (3q27), where a significant linkage to several obesity-related traits has been previously established (16). Because the region with linkage evidence on

TABLE 2
Results from genome-wide linkage analysis of obesity phenotypes

Phenotype	Chromosome	Genomic marker	Max location*	Region†	*P* value
BMI	3	D3S2427	187	172–197	1.8
PBF		D3S2418	215	201–225	0.95
BMI	5	D5S817	21	10–38	1.9
PBF	6	D6S1959	23	0–41	2.7

*The map position of the maximally linked marker; †the boundaries of the peak where LOD score >0.588 or *P* value <0.05.

chromosome 3 almost completely overlaps with a region found to have linkage in white families, our finding on chromosome 3 can be considered a replication. This region of ~20 cM with a *P* value of 0.01 then qualifies for consideration as a significant replication result after adjusting for multiple comparisons (13). Hence, our finding on chromosome 3 demonstrates significant linkage evidence to obesity. Simulation results provided LOD scores of 1.8 for BMI and 0.95 for PBF, corresponding to *P* values of 0.005 and 0.055 after adjusting for multiple comparisons in this region. The transmission/disequilibrium test also demonstrated evidence of linkage and association between the markers and BMI in this region, providing further evidence that this region could harbor obesity genes. In this region, several candidate genes have been identified that can influence energy expenditure, fat partitioning, and insulin sensitivity (16). In a recent report, this region was also strongly linked to type 2 diabetes (17). Therefore, it is possible that genes located in this region affect both obesity and type 2 diabetes.

Our finding on chromosome 5 also overlaps with the region reported to have suggestive linkage evidence in a French population (14). However, the region found by the study of French population is too wide to conclude that our finding on chromosome 5 qualifies as replication, and further studies will be required before any conclusions can be reached.

Comuzzie et al. (5) performed multipoint variance component analysis adjusting for the covariates of sex, age, and age^2 using the program package SOLAR. Strong linkage (LOD score = 4.95) was demonstrated at 74 cM away from *p*-ter on chromosome 2. Although we failed to find a LOD score at the level usually required to satisfy criteria for replication significance (13), namely >1.17, our study does provide minimal linkage evidence to the region of 42–60 (LOD = 1.0), close to this same region reported by Comuzzie et al. Likewise, in a separate set of African-American families from the Chicago area, we previously found evidence of linkage with the same markers as reported by the San Antonio group (5). With this second replication, the region of 42–60 thus becomes an important candidate region; the POMC gene is the primary candidate lying in this interval. Our results also support linkage with chromosome 10, consistent with the recent article by Hager et al. (14), who found a LOD score of 4.85 at 55 cM from *p*-ter on chromosome 10 by model-free multipoint affected sibpair analysis.

TABLE 3
Results from QTDT for testing association between BMI and the markers in the region of chromosome 3 where linkage evidence is found

Marker and allele	*P* value
D3S3053	
226	0.0710
D3S2427	
225	0.0676
233	0.0186
237	0.0391
D3S1262	
112	0.0098
D3S2394	
282	0.0557
D3S2419	0.0072

Our study showed consistent evidence for linkage for both phenotypes only on chromosome 3. There are several factors that could contribute to this inconsistent finding across the phenotypes. *1*) The genetic determinants of interindividual variation in obesity-related traits are likely to be multiple and interacting, with a variety of effects on the obesity traits. *2*) Our data were ascertained through measurement of blood pressure rather than obesity-related traits; therefore, we may not have enough power to find linkage for some phenotypes in some genomic regions with even moderate effect. *3*) Some positive findings may be due to type I error. *4*) The heterogeneity due to the admixture in African-American population may be another source that could contribute to inconsistency.

In conclusion, this African-American population sample of moderate size provides suggestive evidence in support of linkage on several chromosomes identified in prior research. Further replication of these findings will be required before they should be considered guides to fine mapping efforts. The most important limitation of our data is probably the lack of power provided by this sample to detect linkage evidence, as shown in our simulations. Genotyping of larger samples and improved statistical methods, such as multivariate analysis of the two phenotypes, should increase power and provide challenges for future studies.

ACKNOWLEDGMENTS

This work was supported by grants from the National Heart Lung and Blood Institute in support of the GenNet study and the Family Blood Pressure Program (UO1-HL54485; HL54512, HL54466, HL65702). Genotyping was conducted by the NHLBI Mammalian Genotyping Service, Marshfield, Wisconsin.

REFERENCES

1. Barsh GS, Farooqi IS, O'Rahilly S: Genetics of body-weight regulation. *Nature* 404:644–651, 2000
2. Chagnon YC, Perusse L, Weisnagel SJ, Rankinen T, Bouchard C: The human obesity gene map: The 1999 update. *Obes Res* 8:89–117, 2000
3. Comuzzie AG, Allison AB: The search for human obesity genes. *Science* 280:1374–1377, 1998

4. Risch N: Searching for the genetic determinants in the new millennium. *Nature* 405:847–856, 2000
5. Comuzzie AG, Hixson JE, Almasy L, Mitchell BD, Mahaney MC, Dyer TD, Stern MP, MacCluer JW, Blangero J: A major quantitative trait locus determining serum leptin levels and fat mass is located on human chromosome 2. *Nat Genet* 15:273–276, 1997
6. Rotimi CN, Comuzzie AG, Lowe WL, Luke A, Blangero J, Cooper RS: The quantitative trait locus on chromosome 2 for serum leptin levels is confirmed in African Americans. *Diabetes* 18:643–644, 1999
7. Province MA, Boerwinkle E, Chakravarti A, Cooper R, Fornage M, Leppert M, Risch N, Ranade K: Lack of association of the angiotensinogen-6 polymorphism with blood pressure levels in the comprehensive NHLBI Family Blood Pressure Program. *J Hypertens* 18:867–875, 2000
8. Weber JL, Broman KW: Genotyping for human whole-genome scans: past, present, and future. *Adv Genet* 42:77–96, 2001
9. O'Connell JR, Weeks DE: PedCheck: A program for identifying genotype incompatibilities in linkage analysis. *Am J Hum Genet* 63:259–266, 1998
10. Hinds D, Risch N: The ASPEX package: affected sib-pair mapping. ftp://lahmed.stanford.edu/pub/aspex, 1996
11. Kruglyak L, Daly MJ, Reeve-Daly MP, Lander ES: Parametric and nonparametric linkage analysis: a unified multipoint approach. *Am J Hum Genet* 8:1347–1363, 1996
12. Abecasis GR, Cardon LR, Cookson WO: A general test of association for quantitative traits in nuclear families. *Am J Hum Genet* 66:279–292, 2000
13. Lander ES, Kruglyak L: Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat Genet* 11:241–247, 1995
14. Hager J, Dina C, Franck S, Dubois S, Houari M, Vatin V, Vaillant E, Lorentz N, Basdevant A, Clement K, Guy-Grand B, Froguel P: A genome-wide scan for human obesity genes reveals a major susceptibility locus on chromosome 10. *Nat Genet* 20:304–308, 1998
15. Altmuller J, Palmer LJ, Fischer G, Scherb H, Wjst M: Genomewide scans of complex human diseases: true linkage is hard to find. *Am J Hum Genet* 69:936–950, 2001
16. Kissebah AH, Sonnenberg GE, Myklebust J, Goldstein M, Broman K, James RG, Marks JA, Krakower GR, Jacob HJ, Weber J, Martin L, Blangero J, Comuzzie AG: Quantitative trait loci on chromosomes 3 and 17 influence phenotypes of the metabolic syndrome. *Proc Natl Acad Sci U S A* 97:14478–14483, 2000
17. Vionnet N, Hani EH, Dupont S, Gallina S, Francke S, Dotte S, Matos FD, Durand E, Leprêtre F, Lecoeur C, Gallina P, Zekiri L, Dina C, Froguel P: Genomewide search for type 2 diabetes susceptibility genes in French whites: evidence for a novel susceptibility locus for early-onset diabetes on chromosome 3q27-qter and independent replication of a type 2 diabetes locus on chromosome 1q21q24. *Am J Hum Genet* 67:1470–1480, 2000