Brief Genetics Report

Fasting Insulin and Obesity-Related Phenotypes Are Linked to Chromosome 2p

The Strong Heart Family Study

Vincent P. Diego,[1] Harald H.H. Göring,[1] Shelley A. Cole,[1] Laura Almasy,[1] Thomas D. Dyer,[1] John Blangero,[1] Ravindranath Duggirala,[1] Sandra Laston,[1] Charlotte Wenger,[1] Teresa Cantu,[1] Bennett Dyke,[1] Kari North,[2] Theodore Schurr,[3] Lyle G. Best,[4] Richard B. Devereux,[5] Richard R. Fabsitz,[6] Barbara V. Howard,[7] and Jean W. MacCluer[1]

To localize quantitative trait loci for insulin metabolism and obesity, genome scans/linkage analyses were performed on >900 members of 32 extended families participating in phase 3 of the Strong Heart Study, an investigation of the genetic and environmental determinants of cardiovascular disease in American-Indian populations from Arizona, Oklahoma, and North and South Dakota. Linkage analyses of fasting insulin and two obesity-related phenotypes, BMI and percent fat mass, were performed independently in each of the three populations. For log fasting insulin, we found a genome-wide maximum, robust logarithm of odds (LOD) score of 3.42 at 51 cM on chromosome 2p in the Dakotas. Bivariate linkage analyses of log fasting insulin with both BMI and fat mass indicate a situation of incomplete pleiotropy, as well as several significant bivariate LOD scores in the Dakotas. *Diabetes* 55: 1874–1878, 2006

American-Indian populations have high incidence rates of type 2 diabetes (1). To better understand the determinants of cardiovascular disease, type 2 diabetes, and related diseases in American Indians, the Strong Heart Study (SHS) was begun in 1989 (2). In phase 3 of the SHS, a Family Study (SHFS) was initiated (3). Here we report linkage results from a genome-wide scan in American-Indian families from Arizona, Oklahoma, and North and South Dakota who are participating in the SHFS (2,3). Despite the low genetic diversity in American Indians as a whole, they still exhibit a substantial range of genetic variation across contemporary populations (4). There are a number of population-genetic and historical reasons that would suggest that looking at Arizona, Oklahoma, and the Dakotas as separate populations in a genome-wide scan study is a worthwhile enterprise. (Mulligan et al. [4] have made a more general argument along these lines.) Indeed, the results reported herein will be seen to support this argument.

The SHFS protocols were approved by the Indian Health Service Institutional Review Board, the institutional review boards of the participating institutions, and the 13 American-Indian tribes participating in these studies. Demographic, pedigree, phenotype, and genotype data are reported in Table 1.

We report results for fasting insulin levels (in microunits per milliliter) and two obesity-related traits, percent fat mass and BMI (weight in kilograms divided by the square of height in meters). Sample sizes for each of the linkage analyses are determined by the number of individuals measured for the trait under analysis (that is, the number phenotyped). These values are given in Table 1. Fasting insulin was assayed by radioimmunoassay, and diabetic subjects on medication were excluded from the analyses. However, diabetic subjects not on medication were included in the analyses. It should be mentioned here that analyses of homeostasis model assessment of insulin resistance, which is highly correlated with fasting insulin and computed after excluding all diabetic subjects, gave rise to essentially identical results compared with the fasting insulin analyses. Therefore, the observed similarity between the two trait analyses would seem to indicate that the inclusion of diabetic subjects did not have a distorting effect on the fasting insulin analyses. Having made this point, only the results from the fasting insulin analyses will be reported here. Fat mass was determined using a Quantum II bioelectric impedance analyzer (RJL Systems, Clinton Township, MI) and a formula validated for American Indians (2). Across populations, fasting insulin data were log transformed to induce normality. For the Arizona population, BMI and fat mass data were exactly normalized using the inverse Gaussian transformation. These normalized traits will be denoted as BMI_n and fat mass_n, respectively. For the Dakotas population, BMI and fat mass data were both sufficiently normally distributed. For the Oklahoma population, fat mass data were

From the [1]Department of Genetics, Southwest Foundation for Biomedical Research, San Antonio, Texas; the [2]Department of Epidemiology, University of North Carolina, Chapel Hill, North Carolina; the [3]Department of Anthropology, University of Pennsylvania, Philadelphia, Pennsylvania; [4]Missouri Breaks Industries Research, Timber Lake, South Dakota; the [5]Cornell University Medical Center, New York, New York; the [6]National Heart, Lung, and Blood Institute, Bethesda, Maryland; and the [7]MedStar Research Institute, Hyattsville, Maryland.

Address correspondence and reprint requests to Vincent P. Diego, PhD, Southwest Foundation for Biomedical Research, Department of Genetics, PO Box 760549, San Antonio, TX 78245-0549. E-mail: vdiego@darwin.sfbr.org.

Received for publication 25 May 2005 and accepted in revised form 27 February 2006.

LOD, logarithm of odds; QTL, quantitative trait locus; SHFS, Strong Heart Family Study; SHS, Strong Heart Study.

DOI: 10.2337/db05-0668

TABLE 1
Descriptive characteristics of SHFS participants

	Arizona	Dakotas	Oklahoma
Demographics			
Age (years)	42 (18–88)	41 (18–88)	44 (18–88)
n (M/F)	341 (142/199)	327 (129/198)	307 (138/169)
Pedigrees			
No.	13	6	11
Pedigree size (*n*)*	53 (3–95)	138 (19–208)	40 (24–55)
No. of generations	5	5	5
Phenotypes			
Fasting insulin (μU/ml)	27.8 ± 23.3 (263)	20.8 ± 19.9 (296)	19.4 ± 22.8 (275)
BMI (kg/m^2)	34.9 ± 8.7 (348)	30.2 ± 5.9 (323)	30.5 ± 6.5 (309)
Percent fat mass	40.3 ± 9.5 (336)	34.6 ± 9.4 (312)	34.3 ± 9.3 (306)
n genotyped	341	327	307

Data are mean (range) or mean ± SD (*n*), unless otherwise indicated. *Not all individuals were studied.

sufficiently normally distributed, whereas BMI data were exactly normalized using the inverse Gaussian transformation. For all linkage analyses, we used age, sex, age^2, age × sex, and age^2 × sex as covariates.

Fasting blood samples were collected at the SHS regional clinics. Buffy coats were stored at the clinics at −80°C and transported on dry ice to the Southwest Foundation for Biomedical Research (San Antonio, TX) for DNA extraction. To genotype SHS participants, we used the ABI PRISM Linkage Mapping Set-MD10 (version 2.5; Applied Biosystems, Foster City, CA), which consists of PCR primers that amplify dinucleotide repeat microsatellite loci (short tandem repeats) selected from the Genethon human linkage map (5). DNA from study participants was arrayed on 384-well PCR plates using the Robbins Hydra-96 Microdispenser (Sunnyvale, CA). PCRs used the True Allele PCR Premix (Applied Biosystems), and amplification was performed in Applied Biosystems 9700 thermocyclers. The products of separate PCRs, for each individual, were pooled using the Robbins Hydra-96 Microdispenser, and a labeled size standard was added to each pool. The pooled PCR products were loaded into an ABI PRISM 377 or 3100 Genetic Analyzer for laser-based automated genotyping. The short tandem repeats and standards were detected and quantified, and genotypes were scored using the Genotyper software package (Applied Biosystems).

We finalized on the pedigrees that were most supported by the genotype data, as inferred from PREST (6) statistics. Using Simwalk II (7), we eliminated mistyping errors at a blanking rate of <2% of the total number of genotypes for the Arizona, Dakotas, and Oklahoma centers. Using Loki (8), we computed matrices of empirical estimates of identity-by-descent allele sharing at points throughout the genome for every relative pair. The identity-by-descent matrices are required for our linkage analyses. PREST, Simwalk II, and Loki require chromosomal maps of genotype markers for all 22 autosomes. Across populations, we used the same sex-averaged chromosomal maps, which were obtained from the Marshfield Center for Medical Genetics (http://research.marshfieldclinic.org/genetics).

For univariate linkage analyses, we used variance components methods as implemented in SOLAR (9). Under a bivariate linkage analysis approach (10), previous studies have established quantitative trait locus (QTL) pleiotropy with respect to insulin and obesity-related traits at putative QTLs on chromosomes 6q and 7q (11,12). QTL pleiotropy may be defined as the condition whereby one QTL influences two or more traits (10–12). To address the QTL pleiotropy hypothesis in American Indians in the SHFS, we performed bivariate linkage analyses of two-trait combinations of log fasting insulin with BMI and fat mass. We distinguish between complete and incomplete pleiotropy. The QTL correlation coefficient is unity (or −1, depending on the sign of the polygenic correlation coefficient) under complete pleiotropy, whereas nonzero departures from unity (or −1) define incomplete pleiotropy. Incomplete pleiotropy may be due to the differential influence of epistasis or genotype-by-environment interaction on one trait relative to the other (10) or to variants in a QTL that have differential effects on the traits (12). In contradistinction, complete pleiotropy is an ideal case whereby one QTL affects two or more traits with no interactions or allelic heterogeneity involved. The only other possible alternative under a bivariate model is coincident linkage, which is defined as the condition whereby two completely uncorrelated loci exist in close proximity under a linkage signal (10–12). To test for pleiotropy, we carried out likelihood ratio tests of the significance of the obesity-related phenotypes under the bivariate model. The likelihood ratio test statistic is distributed as a one-half–to–one-half mixture of χ^2 random variables with 1 and 2 df. To test for complete pleiotropy, we performed similar tests of the null hypothesis, which states that the QTL correlation coefficient between the two traits analyzed is equal to 1 (10,11). In this case, the likelihood ratio test statistic is distributed as a one-half–to–one-half mixture of χ^2 random variables with point mass at zero and 1 df (10,11). Under coincident linkage, the null hypothesis is that the QTL correlation coefficient equals 0. The likelihood ratio test statistic in this third case is distributed as a χ^2 random variable with 1 df (10,11). All linkage analyses were performed independently within field centers (AZ, OK, and ND/SD). We report robust logarithm of odds (LOD) scores according to the correction method analytically derived in Blangero et al. (13) for the Dakotas and Oklahoma univariate linkage analyses. These robust LOD scores are effectively free from type I error due to a misspecified model. For Arizona, the correction method slightly increases the observed LOD scores. To be conservative, we report the lower scores. The correction method is not yet supported in SOLAR for bivariate linkage analyses.

For the Dakotas center, the genome-wide maximum, robust LOD scores for log fasting insulin, BMI, and fat mass are 3.42 at 51 cM on chromosome 2 near marker D2S165 (Fig. 1), 1.12 at 45 cM on chromosome 2, and 2.23

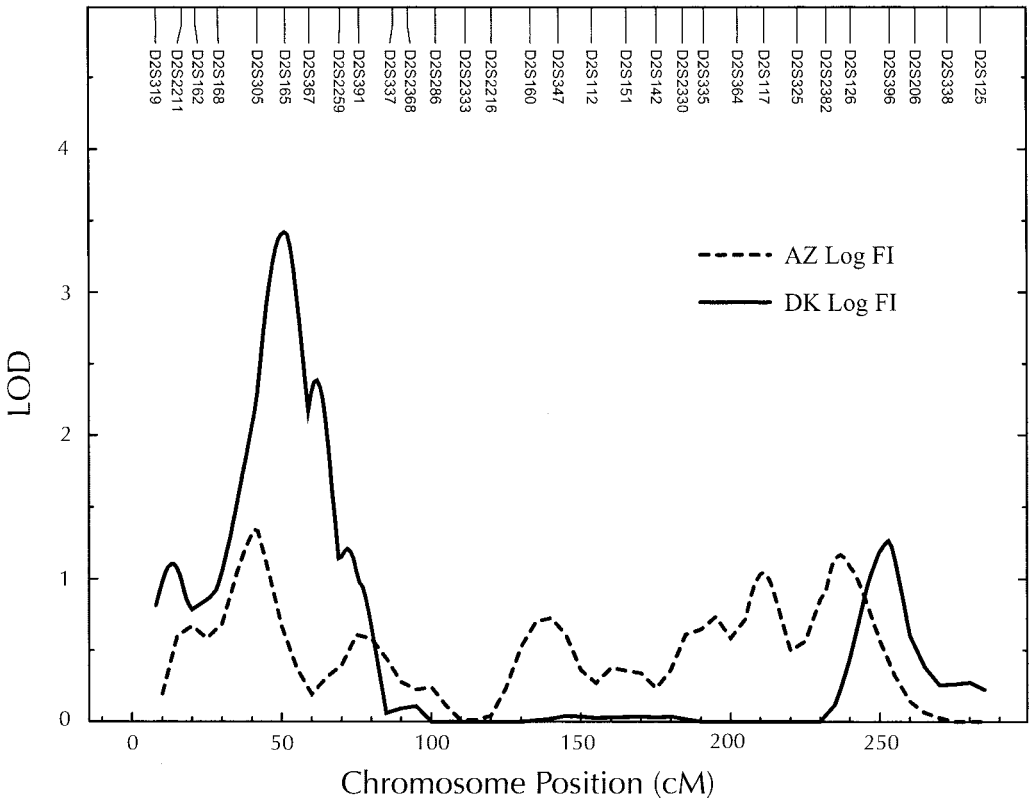

FIG. 1. Univariate linkage signals for log fasting insulin from the Dakotas (solid line) and Arizona (dotted line) centers on chromosome 2.

at 66 cM on chromosome 11, respectively. It should be pointed out that the robust LOD score for log fasting insulin decreases to 2.8 at the same general region on chromosome 2 after excluding all diabetic subjects. Roughly speaking, an LOD score >3.0 is significant. A suggestive, robust LOD score (LOD >2.0) of 2.50 was also found for log fasting insulin (but not for the other two traits) at 69 cM on chromosome 13 for the Dakotas center. For the Arizona center, the genome-wide maximum LOD scores for log fasting insulin, BMI_n, and fat mass_n are 2.01 at 88 cM on chromosome 15, 2.59 at 208 cM on chromosome 4, and 1.94 at 107 cM on chromosome 7, respectively. Another suggestive LOD score for BMI_n is 2.3 at 76 cM on chromosome 2. For the Oklahoma center, the genome-wide maximum LOD scores for log fasting insulin, BMI_n, and fat mass are 2.61 at 165 cM on chromosome 3, 1.18 at 205 cM on chromosome 3, and 1.65 at 22 cM on chromosome 1, respectively. There were no other suggestive LOD scores for Oklahoma. The significant linkage signal for log fasting insulin on chromosome 2p is reported in Fig. 1 (along with a weaker Arizona center log fasting insulin signal in the same region). The log fasting insulin signal on chromosome 2p from the Oklahoma center is negligible.

The Arizona and Oklahoma centers gave negligible signals on chromosome 2 under the bivariate analyses. The bivariate results for the Dakotas center are presented in Table 2 and Fig. 2. We offer the following interpretations for the *P* values in Table 2. Although neither of the obesity-related traits seemed to be underlain by QTLs on chromosome 2 under the univariate analyses, they each made a significant contribution to the bivariate linkage signal, thus implying pleiotropy (Table 2). Moreover, coincident linkage was rejected for these cases (Table 2). Interestingly, complete pleiotropy was also rejected (Table 2), thus indicating incomplete pleiotropy (10,12). Bivariate linkage signals on chromosome 2 for the Dakotas center are reported in Fig. 2. We found LOD scores (adjusted for 3 df) of 3.43 at 48 cM for log fasting insulin plus fat mass and 2.91 at 52 cM for log fasting insulin plus BMI. It is important to note that, while the polygenic genetic correlation between BMI and fat mass was estimated to be ~0.83 ($P = 3.74 \times 10^{-9}$), there was no evidence of QTLs on chromosome 2. This would seem to indicate that the bivariate linkage signals on chromosome 2 are not an artifact of redundancy in the obesity-related traits and that the bivariate linkage signals are perhaps better explained by some general relationship of fasting insulin with obesity.

Because BMI is often used as a covariate in linkage

TABLE 2
QTL pleiotropy specifically on chromosome 2 for fasting insulin in relation to BMI and fat mass

Trait pair	QTL correlation coefficient**	No pleiotropy	Complete pleiotropy	Coincident linkage
Log fasting insulin + BMI	0.92 ± 0.11	$P = 6.24 \times 10^{-8}$	$P = 2.80 \times 10^{-5}$	$P = 1.57 \times 10^{-8}$
Log fasting insulin + fat mass	0.69 ± 0.21	$P = 5.18 \times 10^{-8}$	$P = 2.43 \times 10^{-6}$	$P = 4.08 \times 10^{-8}$

Data are mean ±SD unless otherwise indicated. These results are for the genome-wide maximum under the bivariate linkage analyses for each trait pair (see text).

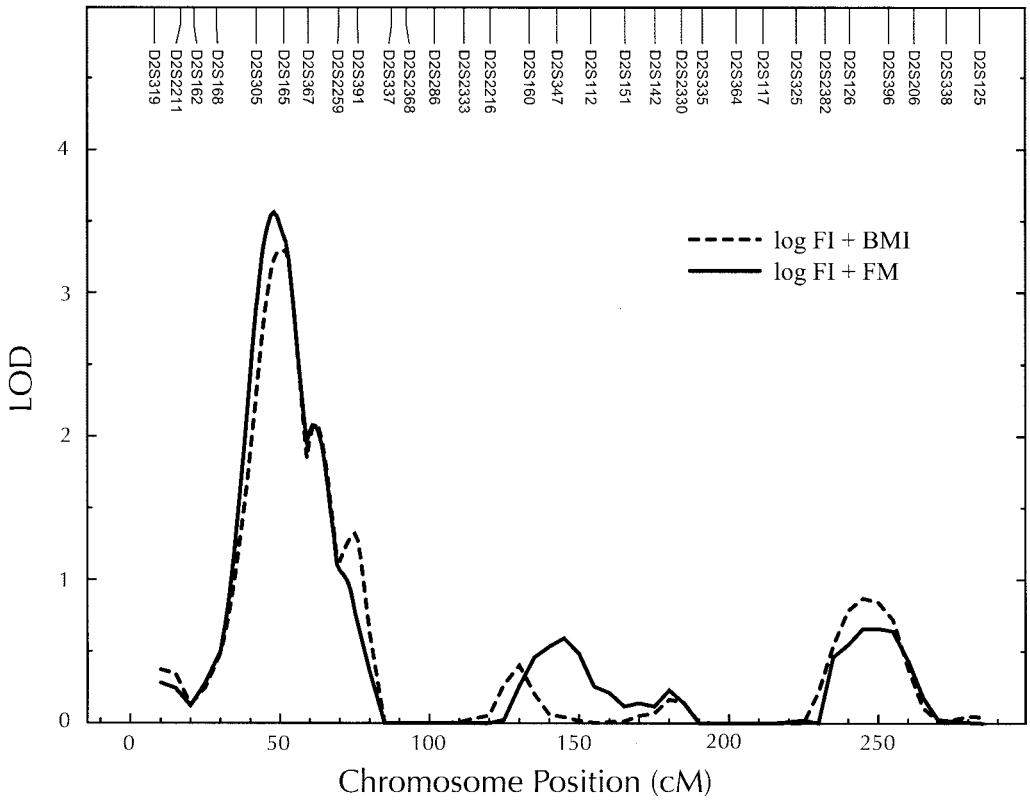

FIG. 2. Bivariate linkage signals for log fasting insulin in relation to obesity-related phenotypes from the Dakotas center. Dashed line, log fasting insulin plus BMI; solid line, log fasting insulin plus fat mass.

analyses of insulin-metabolism phenotypes, we performed our linkage analysis of log fasting insulin in the Dakotas center with BMI as an additional covariate. We found that the original linkage signal became nonsignificant. Duggirala et al. (14) reported a similar result, in which their original genome-wide maximum LOD scores of 4.2 and 3.2 for 32,33-split proinsulin and intact proinsulin, respectively, declined to nonsignificance after including BMI as a covariate. They interpreted this result to indicate strong pleiotropic effects affecting insulin and obesity-related phenotypes. This conclusion is consistent with our bivariate results.

Linkage to chromosome 2p was originally reported for serum leptin levels in Mexican Americans (15) and was replicated in French Caucasians (16) and African Americans (17). The proopiomelanocortin gene (*POMC*) (15–17), which is located at ~44 cM, is one possible candidate gene that may be underlying our linkage signal. That *POMC* variation is important in energy metabolism is suggested by loss-of-function mutations in the *POMC* gene associated with hyperphagia and obesity in humans (18) and by a *POMC*-deletion knockout mouse model of hyperphagia and obesity (19). The point-wise, robust LOD score in the Dakotas center at 44 cM is 2.73, and the point-wise LOD score in the Arizona center at 44 cM is 1.18 for log fasting insulin. One limitation of the present study is that serum leptin could not be tested because leptin was unavailable for analysis. Given the above evidence from univariate linkage analyses of fasting insulin and from bivariate linkage analyses of fasting insulin in relation to obesity-related phenotypes, however, we feel that *POMC* is a likely candidate gene underlying our univariate and bivariate signals.

It is perhaps not too surprising that only the Dakotas center showed evidence of possible linkage to chromosome 2p, whereas the Arizona and Oklahoma centers showed very little to no evidence in this regard. We may have identified a gene that segregates in a few families and/or one that is most readily detectable in the very large extended families from the Dakotas. We found that multiple families contributed to the genome-wide maximum LOD score of 3.42 for log fasting insulin. To clarify the situation up to this point, we carried out formal tests of linkage heterogeneity (rev. in 20) using the program HOMO, which was written by Dr. H.H.H. Göring and is available upon request. We found no evidence of linkage heterogeneity.

In summary, we have detected a QTL on chromosome 2p that influences fasting insulin in American Indians in the Dakotas center of the Strong Heart Family Study; the QTL appears to be present primarily in a few very large families. Additionally, bivariate analyses indicate that this QTL may affect measures of obesity. Future studies will focus on identifying the functional gene in these families.

ACKNOWLEDGMENTS

This research was funded by grants U01 HL65520, U01 HL41642, U01 HL41652, U01 HL41654, U01 HL65521, and MH59490.

Foremost, we thank the American-Indian tribes participating in the SHFS. We also thank the directors and respective staff members of the SHFS clinics in Arizona, Oklahoma, and the Dakotas. The views expressed in this article are those of the authors and do not necessarily reflect those of the Indian Health Service.

REFERENCES

1. Lee ET, Welty TK, Cowan LD, Wang W, Rhoades DA, Devereux R, Go O, Fabsitz R, Howard BV: Incidence of diabetes in American Indians of three geographic areas: the Strong Heart Study. *Diabetes Care* 25:49–54, 2002
2. Lee ET, Welty TK, Fabsitz R, Cowan LD, Le N-A, Oopik AJ, Cucchiara AJ, Savage PJ, Howard BV: The Strong Heart Study: a study of cardiovascular disease in American Indians: design and methods. *Am J Epidemiol* 132:1141–1155, 1990
3. North KE, Howard BV, Welty TK, Best LG, Lee ET, Yeh JL, Fabsitz RR, Roman MJ, MacCluer JW: Genetic and environmental contributions to cardiovascular disease risk in American Indians: the Strong Heart Family Study. *Am J Epidemiol* 157:303–314, 2003
4. Mulligan CJ, Hunley K, Cole S, Long JC: Population genetics, history, and health patterns in Native Americans. *Annu Rev Genomics Hum Genet* 5:295–315, 2004
5. Dib C, Faure S, Fizames C, Samson D, Drouot N, Vignal A, Millasseau P, Marc S, Hazan J, Seboun E, Lathrop M, Gyapay G, Morissette J, Weissenbach J: A comprehensive genetic map of the human genome based on 5,264 microsatellites. *Nature* 380:152–154, 1996
6. Sun L, Wilder K, McPeek MS: Enhanced pedigree error detection. *Hum Hered* 54:99–110, 2002
7. Sobel E, Papp JC, Lange K: Detection and integration of genotyping errors in statistical genetics. *Am J Hum Genet* 70:496–508, 2002
8. Heath SM: Markov chain Monte Carlo segregation and linkage analysis for oligogenic models. *Am J Hum Genet* 61:748–760, 1997
9. Almasy L, Blangero J: Multipoint quantitative-trait linkage analysis in general pedigrees. *Am J Hum Genet* 62:1198–1211, 1998
10. Almasy L, Dyer TD, Blangero J: Bivariate quantitative trait linkage analysis: pleiotropy versus co-incident linkages. *Genet Epidemiol* 14:953–958, 1997
11. Duggirala R, Blangero J, Almasy L, Arya R, Dyer TD, Williams KL, Leach RJ, O'Connell P, Stern MP: A major locus for fasting insulin concentrations and insulin resistance on chromosome 6q with strong pleiotropic effects on obesity-related phenotypes in nondiabetic Mexican Americans. *Am J Hum Genet* 68:1149–1164, 2001
12. Lehman DM, Arya R, Blangero J, Almasy L, Dyer T, Leach RJ, O'Connell P, Stern MP, Duggirala R: Bivariate linkage analysis of the insulin resistance syndrome phenotypes on chromosome 7q. *Hum Biol* 77:231–246, 2005
13. Blangero J, Williams JT, Almasy L: Robust LOD scores for variance component-based linkage analysis. *Genet Epidemiol* 19 (Suppl. 1):S8–S14, 2000
14. Duggirala R, Stern MP, Mitchell BD, Reinhart LJ, Shipman PA, Uresandi OC, Chung WK, Leibel RL, Hales CN, O'Connell P, Blangero J: Quantitative variation in obesity-related traits and insulin precursors linked to the *OB* gene region on human chromosome 7. *Am J Hum Genet* 59:694–703, 1996
15. Comuzzie AG, Hixson JE, Almasy L, Mitchell BD, Mahaney MC, Dyer TD, Stern MP, MacCluer JW, Blangero J: A major quantitative locus determining serum leptin levels and fat mass is located on human chromosome 2. *Nat Genet* 15:273–275, 1997
16. Hager J, Dina C, Francke 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
17. 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* 48:643–644, 1999
18. Krude H, Bieberman H, Grüters A: Mutations in the human proopiomelanocortin gene. *Ann N Y Acad Sci* 994:233–239, 2003
19. Yaswen L, Diehl N, Brennan MB, Hochgeschwender U: Obesity in the mouse model of pro-opiomelanocortin deficiency responds to peripheral melanocortin. *Nat Med* 5:1066–1070, 1999
20. Ott J: *Analysis of Human Genetic Linkage*. 3rd ed. Baltimore, MD, Johns Hopkins University Press, 1999