Brief Genetics Report

Linkage and Linkage Disequilibrium Mapping of Genes Influencing Human Obesity in Chromosome Region 7q22.1–7q35

Wei-Dong Li,[1] Ding Li,[1] Shuang Wang,[2] Shuanglin Zhang,[2,3] Hongyu Zhao,[2] and R. Arlen Price[1]

Linkage results suggest that the region of chromosome 7 containing the leptin gene cosegregates with extreme obesity; however, leptin coding region mutations are rare. To investigate whether the leptin flanking sequence and/or a larger 40-cM region (7q22.1–7q35) contributes to obesity, we genotyped individuals from 200 European American families segregating extreme obesity and normal weight (1,020 subjects) using 21 microsatellite markers and two single nucleotide polymorphisms (SNPs) and conducted nonparametric linkage (NPL) analyses. We also carried out transmission disequilibrium tests for 135 European American triads using 27 markers (including eight SNPs). Both quantitative (MERLIN-regress) and qualitative (GENEHUNTER and MERLIN-npl) analyses provided evidence for linkage for BMI (GENEHUNTER NPL = 2.98, 20 cM centromeric to leptin at the marker D7S692; MERLIN *Z* score = 3.56). Results for several other regions in 7q gave weak linkage. Transmission disequilibrium test (TDT) and quantitative TDT (and quantitative pedigree disequilibrium test) analyses suggest linkage disequilibrium near leptin and other regions of 7q. Our results suggested that there could be two or more genes in chromosome region 7q22.1–7q35 that influence obesity. A new region found by this study (D7S692-D7S523, 7q31.1) has the most consistent linkage results and could harbor obesity-related genes. *Diabetes* 52: 1557–1561, 2003

From the [1]Department of Psychiatry, Center for Neurobiology and Behavior, University of Pennsylvania, Philadelphia, Pennsylvania; the [2]Department of Medicine, Epidemiology and Public Health, Yale University, New Haven, Connecticut; and the [3]Department of Mathematical Sciences, Michigan Technological University, Houghton, Michigan.

Address correspondence and reprint requests to Dr. R. Arlen Price, Department of Psychiatry, Center for Neurobiology and Behavior, University of Pennsylvania, One Clinical Research Building, Room 105, 415 Curie Blvd., Philadelphia, PA 19104. E-mail: arlen@bgl.psycha.upenn.edu.

Received for publication18 September 2002 and accepted in revised form 18 February 2003.

LD, linkage disequilibrium; LOD, logarithm of odds; NPL, nonparametric linkage; SNP, single nucleotide polymorphism; TDT, transmission disequilibrium test; QPDT, quantitative pedigree disequilibrium test; QTDT, quantitative transmission disequilibrium test.

Obesity is a common condition associated with hypertension, diabetes, and cardiovascular diseases. Twin studies, adoption studies, and segregation analyses have shown that genetic factors play an important role in the pathogenesis of obesity (1,2).

Leptin is an important hormone in body weight regulation (3). Leptin mutations in the rodent (*ob/ob* mice) can cause extreme obesity, but human leptin mutations are rare (4). Linkage analyses have found the human chromosome 7q region containing the leptin gene cosegregates with obesity (5,6). Because leptin coding region mutations do not account for 7q linkage, investigators have focused on the leptin flanking region. Several groups (including ours) (7–9) identified 16 polymorphisms in a 2.5-kb 5′ region flanking leptin. Several common polymorphisms are associated with extreme obesity. Yet we do not know whether the polymorphisms in the leptin flanking region are solely responsible for the linkage on 7q or whether other obesity-related gene(s) are located in this region as well.

To seek other genes that potentially influence obesity, we examined the 40-cM region (7q22.1–7q35) flanking but mostly 5′ of the leptin gene. Analyses evaluated both linkage and linkage disequilibrium (LD) in this region.

We sampled 1,425 European American individuals (1,020 members of 200 nuclear families and 405 individuals from 135 triads, Table 1) using 33 polymorphic markers (including markers for linkage analysis and transmission disequilibrium test [TDT]) (Table 2) across a 40-cM region of 7q22.1–7q35, including the leptin gene flanking region.

Using GENEHUNTER, we obtained the highest nonparametric linkage (NPL) score, 2.98 (P = 0.00015), at marker D7S692 (BMI ≥35 kg/m^2; Fig. 1). There were several regions other than D7S692 that gave weak linkage: D7S685 (NPL = 1.32, 6.8 cM 5′ to leptin) and D7S1804–D7S2452 (NPL = 1.41, 6.7 cM 3′ to leptin); the only empirically significant result for the leptin region was found in families with BMI ≥40 kg/m^2 (NPL = 1.39, P = 0.013).

The strongest result from the nonparametric analyses (NPL) by MERLIN was Z = 3.56 (logarithm of odds [LOD] = 2.75, P = 0.0002) for BMI ≥35 kg/m^2 at marker D7S692 (Table 3, Fig. 1).

TABLE 1
Clinical Characteristics of nuclear family and triads samples data set

	n	Minimum	Maximum	Mean	SD	Skewness	Kurtosis
Nuclear family							
Sex	1,020	321 (male)	698 (female)				
Age (years)	1018	15	90	47.3	14.9	0.532	−0.559
BMI (kg/m^2)	1020	17.3	82.9	35.1	11.1	0.881	0.743
% Fat	820	4.1	62.4	38.3	11.8	−0.358	−0.574
Waist-to-hip ratio	922	0.63	1.66	0.87	0.09	1.095	5.726
Fasting glucose (mg/dl)	896	32.0	500.0	94.07	35.7	4.420	30.387
Triglycerides (mg/dl)	906	28.0	1,482.0	180.9	141.8	3.766	24.097
Leptin (ng/ml)	868	1.0	191.2	32.9	28.6	1.393	2.840
Triads samples							
Sex	405	141 (male)	264 (female)				
Age (years)	405	16	95	55.6	16.0	−0.284	−0.752
BMI (kg/m^2)	405	15.9	97	36.8	13.3	0.892	0.634
% Fat	305	12.8	70.7	40.2	12.5	−0.367	−0.758
Waist-to-hip ratio	345	0.34	1.28	0.86	0.1	0.108	1.704
Fasting glucose (mg/dl)	374	35.0	500.0	101.7	45.6	5.120	34.485
Triglycerides (mg/dl)	374	49.0	1,131.0	187.7	119.4	3.057	16.263
Leptin (ng/ml)	376	1.6	192.7	35.5	30.5	1.077	1.460

We obtained significant results at D7S692 (LOD = 2.26, $P = 0.006$) for BMI and D7S523 (LOD = 2.11, $P = 0.0009$) for age-adjusted BMI using family regression analysis (Table 3, Fig. 1). Estimates of trait mean, variance, and heritability had little effect on results from the family regression analyses (not shown). Sham et al. (24) evaluated their family regression method for non–normally distributed trait via simulation and found that this method should be insensitive to family ascertainment through extreme phenotypes.

TABLE 2
Marker selection and physical and genetic map distances of fine-mapping and TDT analyses

Marker	Nuclear families	Triads	Location (Mb)
D7S796	1		101.98
SNP849370		1	105.005
SNP849403		2	105.014
SNP1476878		3	105.175
SNP257376		4	105.28
D7S2420		5	105.37
D7S2459	2	6	105.82
D7S2456		7	106.17
D7S692	3	8	106.823
D7S2425		9	106.83
D7S525		10	108.13
D7S523	4	11	110.19
D7S643	5	12	119.21
D7S685	6	13	119.77
D7S2529	7	14	120.89
D7S514	8		125.51
D7S2501	9	15	125.99
D7S504	10	16	126.09
D7S1875	11	17	126.23
D7S1529	12	18	126.25
−2548	13	19	126.3668
−633		20	126.3687
−188		21	126.3691
+19	14	22	126.3693
Shintani*	15	23	126.375
D7S530	16	24	127.68
D7S649	17		129.2
D7S1804	18	25	130.62
D7S2452	19	26	131.77
D7S2438	20	27	132.25
D7S1837	21		135.03
D7S2202	22		138.05
D7S794	23		141.97

*A microsatellite marker in leptin 3′-untranslated region reported by Shintani et al.

Using haplotype TDT, we found LD at three locations near the leptin flanking region and a region near marker D7S692. Because of the small triad sample size, results did not reach statistical significance after correction for multiple comparisons and should therefore be considered only suggestive. In triads, we obtained the most significant result for quantitative traits by quantitative TDT (QTDT) (10) at a marker only 130 Kb upstream of leptin (D7S1875, $P = 0.00006$ for BMI, overall Bonferroni significance level = 0.0536, empirical $P < 0.001$ for 1,000 replicates by QTDT). In the nuclear family dataset, QPDT (11) gave LD on the leptin 5′ flanking region (D7S2501 and D7S1875, nominal $P = 0.01$ and $P = 0.002$, respectively).

We also performed quantitative pedigree disequilibrium test (QPDT) analyses of the nuclear family dataset, from which we found LD for markers D7S2459 (close to the first linkage peak at D7S692) and D7S685 (nominal $P = 0.005$ and 0.05, respectively).

In the present study, we found evidence of linkage for obesity phenotypes in the D7S692–D7S523 region. The results are consistent across different (but related) obesity phenotypes and in both discrete and quantitative analyses. The quantitative analysis results (MERLIN) matched the discrete results (GENEHUNTER and MERLIN-npl) very well, the most significant results came from the D7S692–D7S523 region.

Our results suggest the possibility of an obesity-related gene(s) located in a region 20 cM upstream of leptin (D7S692-D7S523). So far, no linkage has been reported in this region for BMI other than by our study. Cheng et al. (12) found linkage for fasting insulin and systolic blood pressure in the D7S523-D7S3061 region. Arya et al. (13) also found linkage in D7S479 and D7S471 (94–106 Mb)

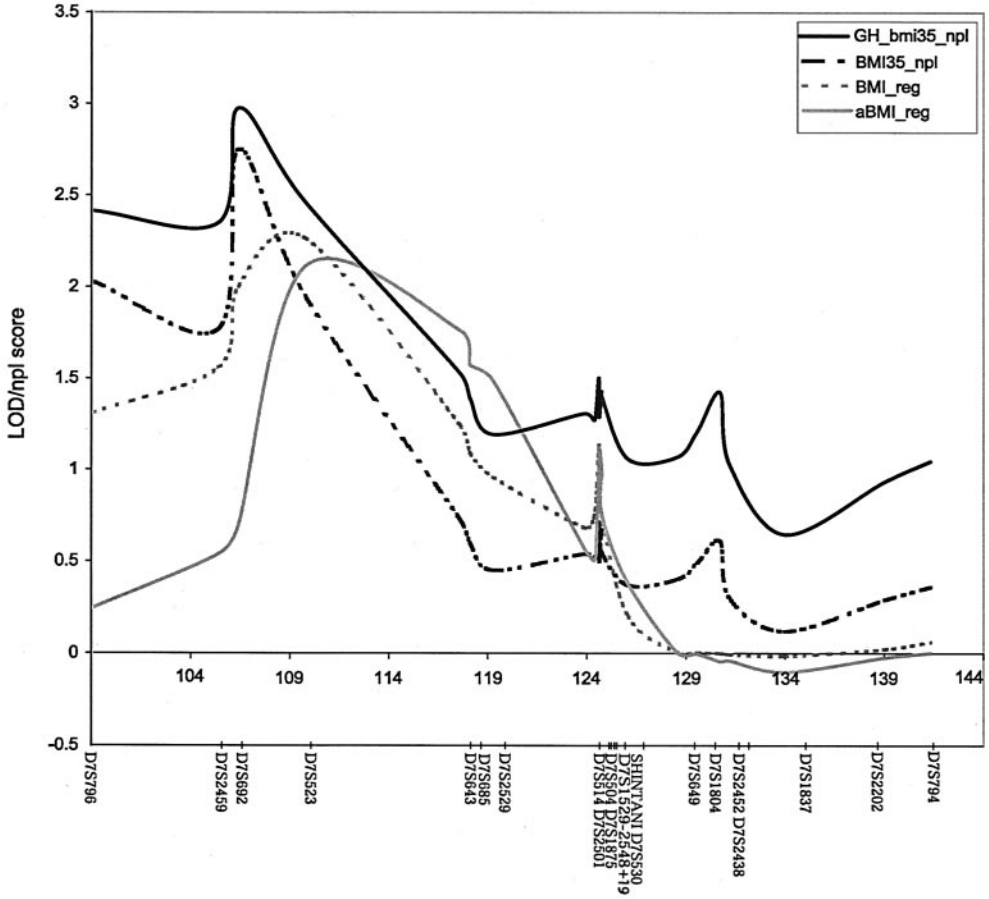

FIG. 1. Results of MERLIN and GENEHUNTER analyses for chromosome 7 markers (7q22.1–7q35): the figure key (upper right) denotes, in order, nonparametric linkage (npl) scores for linkage of BMI ≥35 kg/m^2 by GENEHUNTER (GH_bmi35_npl); NPL analysis for BMI ≥35 kg/m^2 by MERLIN (bmi35_npl); Family Regression analysis for BMI by MERLIN (BMI_reg); and Family Regression analysis for adjusted BMI by MERLIN (aBMI_reg). Distances between markers are given by centimorgans (cM) and *y*-axis refers to LOD score and npl score (GENEHUNTER).

region for HDL and triglyceride in Mexican Americans; however, we have not found the same tendency in our data for triglycerides. Because we do not have enough quantitative data for insulin, blood pressure, and HDL, it is hard to compare with Cheng's and Arya's results. Several candidate genes are close to D7S692-D7S523, including PIK3CG (phosphoinositide-3-kinase catalytic, γ-polypeptide), PRKAR2B (cAMP-dependent protein kinase regulatory subunit type IIβ) and PPP1R3A (protein phosphatase 1, regulatory subunit 3A). Further analysis is needed in this 5-cM region, including candidate gene screening.

Compared with the D7S692-D7S523 region, the present results for the leptin region gave weak linkage only for extreme obesity (BMI ≥40 kg/m^2). The stronger results for extreme obesity are consistent with previously reported linkage to the region (6), including the observed limitation to our extreme phenotype.

We also observed a weak linkage peak close to D7S1804-D7S2452. Because the D7S1804-D7S2452 peak is 5 cM away from leptin (based on the physical mapping), this peak may (or may not) be independent of the result of leptin variants, since the current resolution is poor. Feitosa et al. (14) found an LOD score of 4.9 at D7S1804 in a combined-sample study. Until now, we have not found obvious candidate genes in this region.

Because high-density single neucleotide polymorphism (SNP) mapping has become available, TDT (15) has become more and more popular for complex traits. However, extended pedigrees are not very useful for traditional TDT methods because triads created from extended pedigrees are not totally independent. On the other hand, using the triad samples alone may have less power, particularly in our case where small sample size limitations were compounded by multiple comparisons. During the past several years, there are several articles focused on TDT and extended pedigrees (10,11,16). We used QTDTs and QPDTs to analyze transmission disequilibrium in both nuclear family and triad samples. In triad samples, the leptin 5′ flanking region (0.1–0.6 Mb upstream of leptin) still gave the most significant results and may indicate a "regulatory element" (or another obesity-related gene) located in that region. However, functional analyses are

TABLE 3
Significant results from MERLIN analyses (NPL and family regression) for obesity families

Traits/phenotypes	Marker	*Z* mean	*P*	LOD score	*P*	Empirical *P*
NPL						
BMI ≥35 kg/m^2	D7S692	3.56	0.0002	2.75	0.0002	<0.002
Regression						
BMI	D7S692			2.26	0.0006	0.004
Adjusted BMI	D7S523			2.11	0.0009	0.012

needed. The D7S692-D7S523 region was positive in all our linkage analyses as well as in TDT analyses (GENEHUNTER 2.0) in extreme obesity triads. Usually the 95% CI of QTL mapping is ~10 cM (17) for a study with sample size similar to ours (200 families). However, we saturated the leptin gene region with highly informative markers and achieved a marker density much greater than a usual genome scan. Although we genotyped nine markers in a 0.4-Mb leptin flanking region, D7S692 still showed much more significant linkage. On the other hand, D7S692 is 20 Mb away from leptin, larger than the 10-cM theoretic CI. We found the same tendency for marker D7S2459 (1 cM upstream of D7S692) by QPDT in nuclear families. While only suggestive, these results also support the possibility of an obesity-related gene near D7S692.

In summary, we found suggestive linkage for obesity and related phenotypes in a 40-cM region on 7q22.1–7q35. The most significant result was from the D7S692-D7S523 interval. Positive linkage disequilibrium results suggest a locus is very close to D7S692. It is suggested that at least two obesity-related genes lie in chromosome region 7q22.1–35.

RESEARCH DESIGN AND METHODS

Subjects. A total of 200 European American families (1,020 subjects) were chosen as previously described (18,19). Briefly, all family probands (BMI ≥40 kg/m^2) had at least one obese sibling (BMI ≥30 kg/m^2) and at least one parent and one sibling who were of normal weight (BMI <27 kg/m^2). Average BMI (± SD) for probands, fathers, mothers, sisters, and brothers were 48.8 ± 9.4, 29.4 ± 6.4, 32.3 ± 8.8, 33.3 ± 10.0, and 31.0 ± 7.4 kg/m^2, respectively. Altogether 98 families had both parents' DNA available, 100 families had one parent, and two families had no parent DNA available. Sibship size ranged from 2–9, and most families (195) have 2–6 sibs (1–15 sibpairs) with a median sibship size of 3.

A partially overlapping set of 135 extreme obesity European American triads (405 subjects, 63 of 135 triads were selected from nuclear families) were studied using the TDT analyses. Probands in triads had a BMI ≥40 kg/m^2 and at least one parent with normal weight (BMI <27 kg/m^2). All subjects gave informed consent, and the protocol was approved by the Committee on Studies Involving Human Beings at the University of Pennsylvania.

Phenotypes. BMI was calculated based on measured height and weight: BMI = weight (in killograms)/height (in meters) (2). Percent fat was measured by bioelectric impedance (Tanita TBF310 Pro Body Composition Analyzer; Tanita, Arlington Heights, IL). Fasting serum glucose and triglycerides were assayed by Quest Diagnostic. Plasma leptin levels were measured using the radioimmunoassay method with the Human Leptin RIA Kit (Linco Research). Waist and hip circumferences were measured when blood samples were collected. All quantitative variables were adjusted for linear effects of age, within generation and sex using SPSS 6.1.

Clinical characteristics of nuclear families and triads, including trait-specific coefficients of skewness and kurtosis, are shown in Table 1.

DNA preparation and genotyping. DNA was extracted using a high-salt method (20). A total of 23 polymorphic markers (including two SNPs) for nuclear families and 27 markers (including 8 SNPs) were genotyped (Table 2). Map distances were taken from the Human Genome Working Draft (Human Genome Browser, http://genome.ucsc.edu). Genotyping of microsatellite markers was performed as previously described (18). The –2,548 and –633 polymorphisms were detected by PCR-SSCP; selected samples with shifted bands were reamplified, and PCR products were purified using QIAquick PCR Purification Kit (Qiagen). The Genetics Core Facility at the University of Pennsylvania sequenced this DNA with an ABI 377 automatic sequencer. PCR for the –2,548 polymorphism was performed using 1.6 mmol/l $MgCl_2$. We used special conditions for –188 and +19 SNPs. For –188, we used α-dCTP body labeling, 0.8 units *Taq*/10 μlPCR, and Bss H II 50°C digestion for >3 h. For +19 polymorphism, we used α-^{33}P-dATP body labeling, 1.2 mmol/l $MgCl_2$, and Msp AI1 digestion (37°C) for >3 h. PCR products were separated using 10% PAGE sequencing gel, 1× TBE, 60 W for 1 h. Six samples from each 96-well set were chosen and put on a common plate, which was scored first as a key for all plate genotyping. All band patterns had two independent scorers who were blind to phenotype.

Discrete data analysis. We performed nonparametric multipoint linkage analyses using GENEHUNTER version 1.3 (21) for BMI. We used dichotomous obesity affection status: BMI ≥27, ≥30, ≥35, and ≥40 kg/m^2. Gene frequencies were estimated by allele counting using all individuals who provided DNA. This approach gives asymptotically unbiased estimates of the allele frequencies (22). All Mendelian inheritance and haplotype errors were checked and resolved using GENEHUNTER 1.3. Any genetically unrelated parents and siblings were excluded, as were all half siblings. Using the same affection status (BMI 27, 30, 35, and 40), we performed multipoint NPL analyses by MERLIN (23).

Pedigree regression analysis. We carried out the pedigree-wide regression analyses (24) using MERLIN for quantitative BMI. The trait mean, variance, and heritability were estimated according to epidemiology studies (25). We set trait mean (BMI) = 25.5, variance = 4, and heritability = 50%. Varying these parameter values had little effect on results.

Simulation. Using MERLIN, we simulated 500 replicates of our datasets for NPL and family regression analyses. Empirical *P* values (Table 3) were computed by dividing the number of replicates that exceeded the observed *Z* score or LOD score by the number of replicates (500).

TDT. We performed TDT analysis by GENEHUNTER 2.0. Up to four marker haplotypes were constructed and analyzed using GENEHUNTER 2.0. We used QTDT (10) to analyze quantitative traits in triads. Empirical *P* values were computed by QTDT based on a 1,000-replicates simulation. We also used the QPDT (11) to analyze BMI values in the nuclear families.

ACKNOWLEDGMENTS

This work was supported by the National Institutes of Health (NIH R01DK44073, R01DK48095, and R01DK56210 to R.A.P. and R01 GM 59507 to H.Z.).

We acknowledge the cooperation of our subjects. We thank Guang Ming Yuan, Richard Joseph, Dr. Balasahib Shinde, Elizabeth Joe, Jan Merideth, Amy Morrison, Melissa Rosato, Alison Bernstein, Kenneth Luguya, and Kye Yun for technical assistance. We also thank Dr. Chuanhui Dong for suggestions.

REFERENCES

1. Price RA, Stunkard AJ: Commingling analysis of obesity in twins. *Hum Hered* 39:121–135, 1989
2. Price RA, Ness R, Laskarzewski P: Common major gene inheritance of extreme overweight. *Hum Biol* 62:747–765, 1990
3. Zhang Y, Proenca R, Maffei M, Barone M, Leopold L, Friedman JM: Positional cloning of the mouse obese gene and its human homologue. *Nature* 372:425–432, 1994
4. Considine RV, Considine EL, Williams CJ, Nyce MR, Magosin SA, Bauer TL, Rosato EL, Colberg J, Caro JF: Evidence against either a premature stop codon or the absence of obese gene mRNA in human obesity. *J Clin Invest* 95:2986–2988, 1995
5. 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
6. Reed DR, Ding Y, Xu W, Cather C, Green ED, Price RA: Extreme obesity may be linked to markers flanking the human OB gene. *Diabetes* 45:691–694, 1996
7. Oksanen L, Kainulainen K, Heiman M, Mustajoki P, Kauppinen-Makelin R, Kontula K: Novel polymorphism of the human ob gene promoter in lean and morbidly obese subjects. *Int J Obes Relat Metab Disord* 21:489–494, 1997
8. Mammes O, Betoulle D, Aubert R, Giraud V, Tuzet S, Petiet A, Colas-Linhart N, Fumeron F: Novel polymorphisms in the 5′ region of the LEP gene: association with leptin levels and response to low-calorie diet in human obesity. *Diabetes* 47:487–489, 1998
9. Li WD, Reed DR, Lee JH, Xu W, Kilker RL, Sodam BR, Price RA: Sequence variants in the 5′ flanking region of the leptin gene are associated with obesity in women. *Ann Hum Genet* 63:227–234, 1999
10. 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
11. Zhang S, Zhang K, Li J, Sun F, Zhao H: Test of association for quantitative traits in general pedigrees: the quantitative pedigree disequilibrium test. *Genet Epidemiol* 21:S370–S375, 2001
12. Cheng LS, Davis RC, Raffel LJ, Xiang AH, Wang N, Quinones M, Wen PZ, Toscano E, Diaz J, Pressman S, Henderson PC, Azen SP, Hsueh WA,

Buchanan TA, Rotter JI: Coincident linkage of fasting plasma insulin and blood pressure to chromosome 7q in hypertensive hispanic families. *Circulation* 104:1255–1260, 2001
13. Arya R, Blangero J, Williams K, Almasy L, Dyer TD, Leach RJ, O'Connell P, Stern MP, Duggirala R: Factors of insulin resistance syndrome–related phenotypes are linked to genetic locations on chromosomes 6 and 7 in nondiabetic Mexican Americans. *Diabetes* 51:841–847, 2002
14. Feitosa MF, Borecki IB, Rich SS, Arnett DK, Sholinsky P, Myers RH, Leppert M, Province MA: Quantitative-trait loci influencing body-mass index reside on chromosomes 7 and 13: the National Heart, Lung, and Blood Institute Family Heart Study. *Am J Hum Genet* 70:72–82,2002
15. Spielman RS, McGinnis RE, Ewens WJ: Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). *Am J Hum Genet* 52:506–516, 1993
16. Sinsheimer JS, Blangero J, Lange K: Gamete-competition models. *Am J Hum Genet* 66:1168–1172, 2000
17. Roberts SB, MacLean CJ, Neale MC, Eaves LJ, Kendler KS: Replication of linkage studies of complex traits: an examination of variation in location estimates. *Am J Hum Genet* 65:876–884, 1999
18. Lee JH, Reed DR, Li WD, et al: Genome scan for human obesity and linkage to markers in 20q13 [published erratum appears in *Am J Hum Genet* 66:1472, 2000]. *Am J Hum Genet* 64:196–209, 1999
19. Price RA, Reed DR, Lee JH: Obesity related phenotypes in families selected for extreme obesity and leanness. *Int J Obes Relat Metab Disord* 22:406–413, 1998
20. Lahiri DK, Nurnberger JI Jr: A rapid non-enzymatic method for the preparation of HMW DNA from blood for RFLP studies. *Nucleic Acids Res* 19:5444, 1991
21. Kruglyak L, Daly MJ, Reeve-Daly MP, Lander ES: Parametric and nonparametric linkage analysis: a unified multipoint approach. *Am J Hum Genet* 58:1347–1363, 1996
22. Ott J: Strategies for characterizing highly polymorphic markers in human gene mapping. *Am J Hum Genet* 51:283–290, 1992
23. Abecasis GR, Cherny SS, Cookson WO, Cardon LR: Merlin–rapid analysis of dense genetic maps using sparse gene flow trees. *Nat Genet* 30:97–101, 2002
24. Sham PC, Purcell S, Cherny SS, Abecasis GR: Powerful regression-based quantitative-trait linkage analysis of general pedigrees. *Am J Hum Genet* 71:238–253, 2002
25. Kuczmarski RJ, Carroll MD, Flegal KM, Troiano RP: Varying body mass index cutoff points to describe overweight prevalence among U.S. adults: NHANES III (1988 to 1994). *Obes Res* 5:542–548, 1997