Genetic Linkage and Association of the Growth Hormone Secretagogue Receptor (Ghrelin Receptor) Gene in Human Obesity

Andrea Baessler,[1,2] Michael J. Hasinoff,[1] Marcus Fischer,[1,2] Wibke Reinhard,[2] Gabriele E. Sonnenberg,[1] Michael Olivier,[1] Jeanette Erdmann,[3] Heribert Schunkert,[3] Angela Doering,[4] Howard J. Jacob,[1] Anthony G. Comuzzie,[5] Ahmed H. Kissebah,[1] and Anne E. Kwitek[1]

The growth hormone secretagogue receptor (GHSR) (ghrelin receptor) plays an important role in the regulation of food intake and energy homeostasis. The *GHSR* gene lies on human chromosome 3q26 within a quantitative trait locus strongly linked to multiple phenotypes related to obesity and the metabolic syndrome. Because the biological function and location of the *GHSR* gene make it an excellent candidate gene, we tested the relation between common single nucleotide polymorphisms (SNPs) in the *GHSR* gene and human obesity. We performed a comprehensive analysis of SNPs, linkage disequilibrium (LD), and haplotype structure across the entire *GHSR* gene region (99.3 kb) in 178 pedigrees with multiple obese members (DNA of 1,095 Caucasians) and in an independent sample of the general population (MONICA Augsburg left ventricular hypertrophy substudy; DNA of 1,418 Caucasians). The LD analysis revealed a disequilibrium block consisting of five SNPs, consistent in both study cohorts. We found linkage among all five SNPs, their haplotypes, and BMI. Further, we found suggestive evidence for transmission disequilibrium for the minor SNP alleles ($P < 0.05$) and the two most common haplotypes with the obesity affection status ("susceptible" $P = 0.025$, "nonsusceptible" $P = 0.045$) in the family cohort using the family-based association test program. Replication of these findings in the general population resulted in stronger evidence for an association of the SNPs (best $P = 0.00001$) and haplotypes with the disease ("susceptible" $P = 0.002$, "nonsusceptible" $P = 0.002$). To our knowledge, these data are the first to demonstrate linkage and association of SNPs and haplotypes within the *GHSR* gene region and human obesity. This linkage, together with significant transmission disequilibrium in families and replication of this association in an independent population, provides evidence that common SNPs and haplotypes within the *GHSR* region are involved in the pathogenesis of human obesity. *Diabetes* 54:259–267, 2005

Obesity is a common multifactorial disorder of considerable heterogeneity and, as a pivotal component of the metabolic syndrome, a major risk factor for type 2 diabetes, hypertension, and coronary heart disease as well as premature cardiovascular morbidity and death (1). Together with its associated pathologic features, obesity is among the major causes of illness and death worldwide, as its prevalence continues to rise dramatically (2).

The etiology of obesity is complex, determined by the interplay of genetic and environmental factors. Epidemiological studies have demonstrated a substantial heritable component to the risk for obesity; specifically, 50–70% of the variation in BMI may be attributable to genetic factors (3).

To unravel the genetic etiology of obesity, we previously performed a genomewide linkage scan on a large cohort of Caucasian families from which we localized on chromosome 3q26-q29 a major quantitative trait locus (QTL) strongly linked to six phenotypes of obesity and the metabolic syndrome (4). This QTL has been replicated in several studies and represents one of the most stable findings in complex human genetics (5–9).

A comprehensive review of the available genomic information in the QTL region reveals a positional candidate gene of ~4.3 kb in length encoding the growth hormone secretagogue, or ghrelin, receptor (GHSR). GHSR is known to be involved in growth hormone secretion (10,11). Its major physiological role, however, appears to be in regulating food intake and energy homeostasis by partaking in neuronal mechanisms involving neuropeptide Y and agouti-related protein (12–15). The endogenous GHSR ligand, ghrelin, plays a key role as the major orexigenic hormone. It is secreted in the gastrointestinal

From the [1]Medical College of Wisconsin, Milwaukee, Wisconsin; the [2] Clinic for Internal Medicine II, University of Regensburg, Regensburg, Germany; the [3]Clinic for Internal Medicine II, University of Luebeck, Luebeck, Germany; the [4]GSF Institute of Epidemiology, Neuherberg, Germany; and the [5]Department of Genetics, Southwest Foundation for Biomedical Research, San Antonio, Texas.

Address correspondence and reprint requests to Anne Kwitek, PhD, Department of Physiology, Human and Molecular Genetics Center, Medical College of Wisconsin, 8701 Watertown Plank Rd., Milwaukee, WI 53226. E-mail: akwitek@mcw.edu.

Received for publication 19 May 2004 and accepted in revised form 14 September 2004.

dbSNP, single nucleotide polymorphism public database; FBAT, family-based association test; GHSR, growth hormone secretagogue receptor; LD, linkage disequilibrium; LVH, left ventricular hypertrophy; MONICA, Monitoring Trends and Determinants in Cardiovascular Disease; QTDT, quantitative transmission disequilibrium test; QTL, quantitative trait locus; SNP, single nucleotide polymorphism; TDT, transmission disequilibrium test; TOPS, Take Off Pounds Sensibly.

tract and is carried to the hypothalamic areas that govern food intake, thereby counterbalancing the effects of a multitude of anorectic hormones, such as leptin, insulin, and $PYY_{3–36}$ (16).

The importance of ghrelin in the central regulation of feeding has been demonstrated in animals and humans (17,18). Ghrelin administration increases appetite and food intake in normal subjects and patients with decreased appetite, such as those suffering from cancer cachexia (17). It reduces insulin secretion and enhances energy intake by ~30% (19). Moreover, given that plasma ghrelin levels have been shown to be lower in obese subjects (20,21), recent evidence suggests that obesity is associated with an impairment of the entire ghrelin system (22).

Because its biological function and location make the *GHSR* gene a strong candidate gene for obesity, we carried out a family-based linkage disequilibrium (LD) study in 178 pedigrees as well as in an independent case-control study from the general population. We systematically explored the LD and haplotype structure of the genomic region encompassing the *GHSR* gene and comprehensively assessed the role of common sequence variants and haplotypes in obesity. We report evidence for linkage and association of five single nucleotide polymorphisms (SNPs) and the two most common five-marker haplotypes with obesity in our family cohort. In addition, we describe the association of the same SNPs and haplotypes with obesity and, more striking, with the quantitative phenotype BMI in the general population. Thus, the replication of our findings, together with the location and biological function of the *GHSR* gene, indicate that this gene region is involved in the pathogenesis of the complex disease of obesity.

RESEARCH DESIGN AND METHODS

The family subjects of the study, a large group of Caucasian families consisting of obese and nonobese members residing in the midwestern U.S., were ascertained through the TOPS (Take Off Pounds Sensibly) Club membership as part of the Metabolic Risk and Complications of Obesity Genes project at the Medical College of Wisconsin. The ascertainment strategies and exclusion criteria have been previously published (4). Informed consent was obtained from all participants. Data for the analyses presented here were based on results obtained from 1,302 phenotyped individuals distributed among 178 families (average 7.3 members per family; 441 founders; both parents available for genotyping in 148 families; average generations 2.3). These families have been identified as primary contributors to the QTL on chromosome 3. In all, 1,095 DNA samples were available for genotyping in the present study, including samples of 307 (28%) men and 788 (72%) women. Research protocols were approved by the institutional review board of the Medical College of Wisconsin.

For the general population arm of the study, we used data from subjects in the Monitoring Trends and Determinants in Cardiovascular Disease (MONICA) Augsburg left ventricular hypertrophy (LVH) substudy, as part of the Third MONICA Augsburg survey, which now is continued in the framework of KORA (Cooperative Health Research in the Augsburg Area). The study population of the LVH substudy was sampled from the general population of the city of Augsburg, Germany, in 1994/1995, which originated from a sex- and age-stratified cluster sample of all German residents of the city of Augsburg. The Augsburg project was part of the international collaborative World Health Organization MONICA study (23). The study design, sampling frame, and data collection methods have been described in detail elsewhere (23,24). All the participants gave written informed consent. The LVH substudy represents individuals aged 25–74 years, with ~300 subjects for each 10-year increment (n = 1,674) (24). Of these, 1,418 DNA samples were available for genotyping in the present study, including 724 men (51%) and 694 women (49%). The study was approved by the local ethics committee.

BMI was calculated as weight (kg) divided by height (m) squared. In both cohorts, obesity was defined by a BMI >32 kg/m^2. Subjects were classified as "unaffected" if they presented with a BMI <28 kg/m^2. These cutoff values were chosen to ensure clear phenotypes and avoid misclassification regarding affection status. The obesity affection status of subjects with a BMI of 28–32 kg/m^2 was treated as "unknown."

SNPs and genotyping methods

SNPs. To obtain complete coverage of the *GHSR* gene, 10 SNPs covering the *GHSR* gene and its flanking regions were selected from the SNP public databases (dbSNP; available at http://www.ncbi.nlm.nih.gov/SNP) (Fig. 1). We preferred validated SNPs with a minor allele frequency of >5%. Priority was given to SNPs submitted multiple times then to SNPs discovered by The SNP Consortium (25,26). Regarding the intergenic regions, we prioritized SNPs located in highly conserved noncoding regions. Of the 10 selected SNPs, 1 was located in exon 1, 1 was in the intron, 3 were within 41.5 kb past the 3′ end of the gene, and 5 covered a region of 53.5 kb past the 5′ end of the gene. The coding SNP (rs572169) led to a synonymous amino acid substitution. The eight SNPs located beyond the boundaries of the gene were picked to determine the extent of LD and explore the impact of sequence variations in noncoding and intergenic regions on the disease. In total, a 99.3-kb region was covered with SNPs, with an average resolution of one SNP per 10 kb.

Genotyping. In families, SNP genotyping was performed using the Biplex Invader assay (Third Wave Technologies), as previously described (27). In the general population, SNPs were genotyped using the 5′-exonuclease activity (TaqMan) assay on a HT7900 (Applied Biosystems, Darmstadt, Germany). H_2O controls, CEPH DNA samples of known genotype, as well as blind duplicate samples were included in each round of amplification to check for consistency and to ensure intra- and interplate genotype quality control.

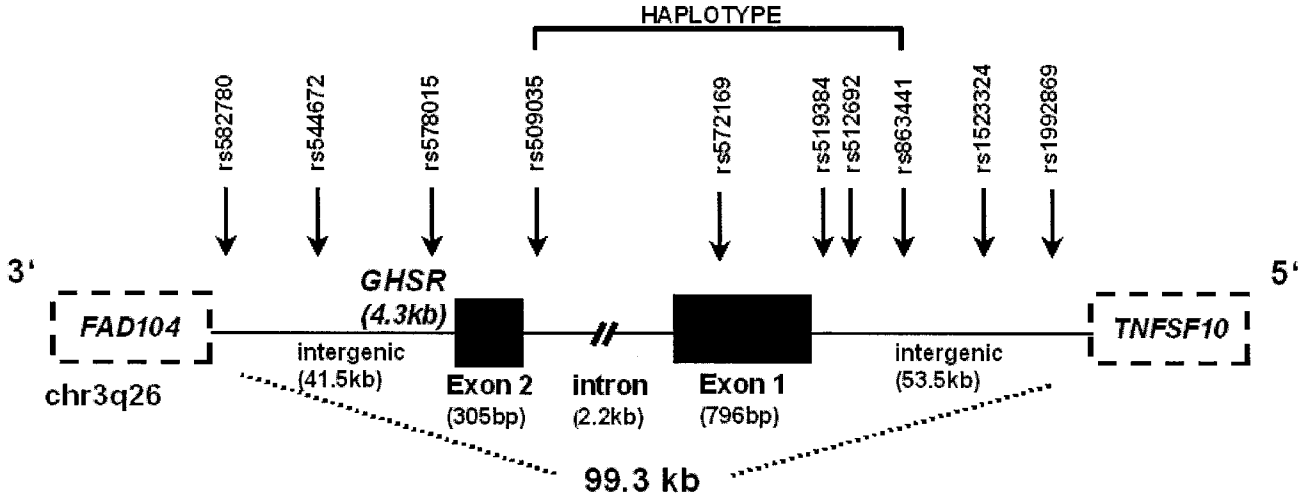
In neither cohort were genotyping discrepancies detected between the repeated samples. The overall misgenotyping rate of <0.5% was due to insufficient PCR amplification.

Statistical analysis. For each of the 10 SNPs, we tested whether the observed allele frequencies departed from the Hardy-Weinberg proportion. No deviations from the expected genotype proportions were detected for any of the SNPs used in the analyses. We also assessed LD between all pairs of SNPs, applying the standard definition of r^2 (28,29).

Families. All genotype data were analyzed using PEDCHECK to check for Mendelian inconsistencies and genotype incompatibilities (30). To test for linkage between these SNPs and the quantitative trait BMI, we used the quantitative transmission disequilibrium test (QTDT) program, which is based on the standard variance component methods and identity by descent among relatives (31). To analyze transmission disequilibrium between the discrete trait obesity and the SNPs, we selected trios with both available parents and one randomly chosen "obesity-affected" offspring and computed the conventional transmission disequilibrium test (TDT) statistic (32). In addition, we applied the software program family-based association test (FBAT) version 1.5 to handle the different types of family structures and use all the information in each family (33). Here, we tested family-based association under the null hypothesis of linkage but no association using the −e flag. This option computes the test statistic through use of the empirical variance (34), which is needed because the markers are in an area of known linkage and the sample contains multiple nuclear families in some pedigrees and multiple affected individuals in these nuclear families.

Structures of haplotypes were analyzed from parental genotypes based on LD pattern using the expectation-maximization algorithm (35). A haplotype block was defined as a region in which all pairwise r^2 values were >0.45 (29). Haplotypes with a frequency >5% were tested for association with obesity using the haplotype FBAT program, which can be used in candidate gene studies with tightly linked markers (36). *P* values were corrected for multiple testing by a method that specifically considers SNPs being in LD with each other and is based on the spectral decomposition (available at http://genepi.qimr.edu.au/general/daleN/SNPSpD/) of matrices of pairwise LD between SNPs (37). Moreover, we tested whether the linkage at this region could be explained by the haplotypes associated with BMI using the QTDT software by modeling linkage and association simultaneously as suggested by Fulker et al. (38). For this purpose, individual haplotypes were reconstructed using HAPLORE (39).

General population. Haplotype frequencies were estimated using the expectation-maximization algorithm (35). By performing an automated, randomized selection of control subjects, all available obese subjects (n = 130 case subjects) were carefully matched by age (±5 years) and sex, whereby one case subject was matched with up to three control subjects (n = 364 control subjects). Frequencies of genotypes in the total study sample as well as in case and control subjects were compared by the Armitage test for trend, and odds ratios with their 95% CIs were reported (40). In addition to single-locus analysis, statistically inferred five-marker haplotypes were tested for association with the discrete trait obesity as well as with the continuous trait BMI



SNP*	Position (bp)	Inter-SNP distance (bp)	Alleles	MAF† (TOPS)	MAF† (MONICA)
rs582780	173442356	12692	A>G	0.41	0.39
rs544672	173455048	25887	C>T	0.15	0.14
rs578015	173480935	3427	A>G	0.29	0.27
rs509035	173484362	2278	G>A	0.29	0.30
rs572169	173486640	2780	G>A	0.29	0.27
rs519384	173489420	2335	T>A	0.27	0.27
rs512692	173491755	4232	A>T	0.26	0.27
rs863441	173495987	22698	G>C	0.27	0.27
rs1523324	173518685	23017	T>C	0.10	0.09
rs1992869	173541702	-----	C>A	0.33	0.33

FIG. 1. Structure of the *GHSR* 1a isoform and position of the 10 selected SNPs covering the *GHSR* gene region including general SNP characteristics as given in the dbSNP database and the Golden Path Genome Browser (July 2003 release). There is no publicly available SNP in exon 2. *SNPs are shown as rs numbers from the dbSNP database. †MAF, minor allele frequency.

using the haplotype trend regression method (41). In this analysis, all individuals were considered, including those with a BMI of 28–32 kg/m^2. Permutation tests (50,000 permutations) were used to test for empirical significance. We also divided the MONICA LVH population into quartiles of BMI distribution, calculated the haplotype frequencies, and tested for significance using Cochran's test for trend.

RESULTS

Phenotypic characteristics. The phenotypic characteristics of TOPS families and the MONICA LVH population are presented in Table 1. In TOPS families, the mean BMI and prevalence of obesity in men and women were markedly different ($P < 0.001$). In the MONICA LVH population, the prevalence of obesity was significantly higher in women than in men ($P < 0.001$).

Linkage disequilibrium evaluation and haplotype structure in families and the general population. Figure 1 depicts the gene structure and all SNPs used in this study, including their position and general characteristics based on the July 2003 release of the Golden Path Genome Browser (available at http://www.genome.ucsc.edu). The pairwise LD block structure defined by the 10 SNPs covering a 99.3-kb region in TOPS families and the MONICA LVH population is shown in Fig. 2. A region of strong LD (r^2 >0.75) was detected between five SNPs (rs509035, rs572169, rs519384, rs512692, and rs863441), and the LD pattern was comparable in both study populations. The pairwise LD between these five SNPs spanned a 11.63-kb region and encompassed most of the intron, exon 1, and 5′ adjacent region of the *GHSR* gene.

According to this high-LD block, five-marker haplotypes were constructed. Only 3 of 2^5 possible haplotypes were estimated to have frequencies >0.01 in obese and/or nonobese subjects. Less frequently occurring haplotypes were not shown, owing to concern over the accuracy of low-frequency alleles in the expectation-maximization algorithm. The two most frequently occurring haplotypes, haplotypes 1 and 2, comprised 94% of total chromosomes in subjects of both study cohorts (Tables 2 and 3). For further linkage and association analyses, we focused on those markers that contributed to this haplotype block. We tested the hypothesis of a relation between these five SNPs and/or five-marker haplotypes and obesity phenotypes. To ensure that the SNPs, which were not included in the high-LD block, were not associated with obesity, we performed association analysis for these SNPs. None of the SNPs showed any evidence for association with the obesity affection status or with BMI (data not shown).

TABLE 1
Phenotypic characteristics of genotyped men and women of the TOPS families and the MONICA Augsburg LVH substudy population

	Men	Women
TOPS		
n	307	788
Age (years)	50.9 ± 17.6	46.0 ± 14.4
Obesity affection status (%)	24.7	50.1
BMI (kg/m^2)		
Total sample	29.1 ± 5.6	33.1 ± 8.4
Affected	37.0 ± 4.5 (75)	39.8 ± 6.2 (393)
Unaffected	24.5 ± 2.4 (141)	24.3 ± 2.4 (244)
Body weight (kg)	93.2 ± 21.1	89.7 ± 24.2
Waist-to-hip ratio	0.95 ± 0.09	0.85 ± 0.10
Hypertension (%)	41.8	36.0
Type 2 diabetes (%)	7.9	7.8
Current smoker (%)	37.2	31.0
MONICA		
n	724	694
Age (years)	52.5 ± 13.9	51.6 ± 13.6
Obesity affection status (%)	9.9	16.1
BMI (kg/m^2)		
Total sample	27.0 ± 3.5	26.4 ± 4.7
Affected	34.6 ± 2.5 (50)	35.0 ± 2.7 (89)
Unaffected	24.9 ± 2.1 (457)	23.8 ± 2.5 (463)
Body weight (kg)	81.9 ± 11.5	68.8 ± 11.9
Waist-to-hip ratio	0.92 ± 0.06	0.80 ± 0.06
Hypertension (%)	41.9	33.4
Type 2 diabetes (%)	5.2	3.0
Current smoker (%)	30.2	23.6

Data are means ± SD, with (*n*) where appropriate, unless otherwise indicated. The obesity affection status is defined as a BMI >32 kg/m^2. Hypertension is defined as systolic blood pressure >140 mmHg, diastolic blood pressure >90 mmHg, or a history of hypertension. Because of nonindependent observations in TOPS families, characteristics are descriptive. Subjects classified as "unknown" are not shown.

Family data: genetic linkage and transmission disequilibrium of SNPs and haplotypes in the *GHSR* region. To test for linkage in families, we used the variance component methodology. Evidence for linkage with the quantitative phenotype BMI was detected for all five SNPs forming the high-LD block depicted in Fig. 2 ($P < 0.05$; data not shown). Thus, our SNP genotype data confirmed linkage to the previously shown QTL on chromosome 3q (4).

To test for transmission disequilibrium in families, we applied both the conventional TDT statistic (in the 148 trios with one randomly selected affected offspring) and the FBAT statistic (considering all family members) for each of the SNPs contributing to the haplotype. The TDT analysis revealed increased transmission for the minor alleles of the five SNPs to obese offspring (Table 2). A slightly stronger pattern of association of the single SNPs with the obesity-affection status was observed when the FBAT statistic was used ($P < 0.05$ for all five SNPs) (Table 2).

In addition, we observed transmission disequilibrium for the two most frequent haplotypes, one consisting of the five major alleles (haplotype 1) and the other consisting of the five minor alleles (haplotype 2) (Table 2). Corresponding to the "susceptible" haplotype, haplotype 2 had a greater number of transmissions to affected offspring ($P = 0.025$). In contrast, the transmission rate of haplotype 1 was significantly reduced in these offspring, suggesting that this haplotype is "nonsusceptible" or resistant to obesity ($P = 0.045$).

After reconstructing the individual haplotypes, we found suggestive evidence for linkage with the quantitative trait BMI ($P = 0.06$). Modeling linkage and association simultaneously resulted in no residual evidence for linkage at this haplotype marker ($P = 0.57$). This indicated that the evidence of linkage at this site was accounted for by association; that is, the haplotype marker contained the disease mutation itself or was in strong LD with it.

General population data: association of SNPs and haplotypes with obesity and BMI. Aiming to replicate the results obtained in the family cohort, we then performed an association analysis in an independent sample of the general population (MONICA Augsburg LVH substudy). Results of the association of the five SNPs are summarized in Table 4 for the entire study sample as well as for matched case and control subjects. Odds ratios were calculated for the comparison of allele frequencies and the "homozygous trait" and "allele positivity" comparisons. Overall, the five SNPs consistently showed nominally significant association with obesity in all three comparisons, in both the entire study sample and the matched case and control subjects (entire study sample, best $P = 0.0000$; matched sample, best $P = 0.0007$ for rs863441). When the result was corrected for multiple testing for SNPs in LD, most *P* values remained significant. In the entire study sample, the increased risk presented by the presence of the minor allele of these SNPs ranged between 41 ($P = 0.014$) and 56% ($P = 0.001$).

To further test for association, we next undertook full haplotype analysis. In agreement with our family data, haplotype 1 was more frequently present in nonobese individuals ("nonsusceptible," $P = 0.002$ after 50,000 permutations) and haplotype 2 was more frequently found in obese individuals ("susceptible," $P = 0.001$ after 50,000 permutations) (Table 3). Moreover, we observed a significant relation between the number of copies of haplotypes 1 and 2 with the qualitative trait obesity in matched case and control subjects and that in the entire study sample. Individuals homozygous for the "susceptible" haplotype 2 or lacking the "nonsusceptible" haplotype 1 presented more often with obesity than individuals with one or no copy of the respective haplotype (haplotype 2: $P = 0.003$ in the matched sample, $P = 0.006$ in the entire sample; haplotype 1: $P = 0.003$ in the matched sample, $P = 0.003$ in the entire sample after 50,000 permutations) (Fig. 3). When tested for association with BMI, individuals carrying two copies of the "susceptible" haplotype 2 or no copy of the "nonsusceptible" haplotype 1 analogously presented with higher BMI than individuals with one or no copy of the respective haplotype (haplotype 2: $P = 0.009$ in the matched sample, $P = 0.007$ in the entire sample; haplotype 1: $P = 0.005$ in the matched sample, $P = 0.006$ in the entire sample after 50,000 permutations) (Fig. 3). We also divided the MONICA LVH population into quartiles of BMI distribution and observed a significantly increasing frequency of the "susceptible" haplotype 2, from 23% in the lowest

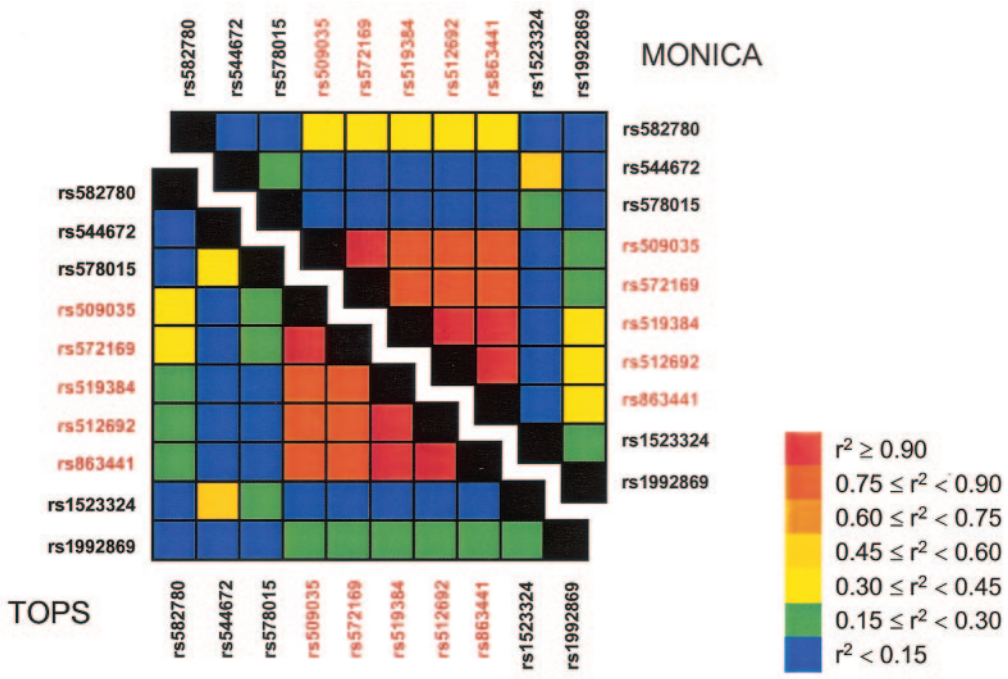


FIG. 2. Pattern of pairwise LD between SNPs in a 99.3-kb region encompassing the *GHSR* gene. The markers are plotted equidistantly. The measure of LD in the r^2 is shown in the lower triangle for the founders of the TOPS families and in the upper triangle for the MONICA Augsburg LVH substudy population. A scale for r^2 of the LD strength is provided on the right.

quartile to 30% in the highest quartile ($P < 0.004$ for trend; data not shown).

DISCUSSION

This study investigated the relation between common sequence variants and haplotypes covering the *GHSR* gene region with obesity phenotypes in families and in an independent sample of the general population. We were led to pursue the GHSR because *1*) the major physiological role of the GHSR appears to be in the central regulation of food intake and body weight; *2*) the *GHSR* gene is located within the QTL on chromosome 3q26-q29, which has been linked to six phenotypes of the metabolic syndrome, including BMI, in the same family cohort; and *3*) only one association study investigating the relation between *GHSR* gene variants and obesity has thus far been performed (42). Because of its location and biological function, the *GHSR* gene represents an excellent candidate gene that may contribute to the pathophysiology and risk of obesity.

The present study offers the first comprehensive analysis of LD, genetic variants, and haplotype structure across the entire *GHSR* gene region in two independent cohorts: families and the general population. The initial LD analysis in the 99.3-kb region revealed an LD block consisting of five SNPs in the *GHSR* gene region, which compared very well in both study cohorts. Subsequently, we focused on

TABLE 2
TDT of association with obesity-affection status in 148 trios selected from the TOPS families and family-based association test using FBAT analysis

SNP/haplotype	Frequency	Allele	Trio TDT (*n* probands with allele)			FBAT *P*
			Transmitted	Not transmitted	*P*	
rs509035	0.71	G	24	37	0.096 (0.137)	0.021 (0.030)
	0.29	A	37	24		
rs572169	0.71	G	23	38	0.055 (0.079)	0.032 (0.046)
	0.29	A	38	23		
rs519384	0.73	T	23	37	0.070 (0.100)	0.010 (0.014)
	0.27	A	37	23		
rs512692	0.74	A	21	38	0.027 (0.039)	0.033 (0.047)
	0.26	T	38	21		
rs863441	0.73	G	22	37	0.051 (0.073)	0.012 (0.017)
	0.27	C	37	22	—	—
Haplotype 1	0.69	G-G-T-A-G	—	—	—	0.025 (NA)
Haplotype 2	0.25	A-A-A-T-C	—	—	—	0.045 (NA)
Haplotype 3	0.03	A-A-T-A-G	—	—	—	0.905 (NA)

SNPs are shown as rs numbers from the dbSNP database. Haplotypes are derived from the five SNPs contributing to the high-LD block ($r^2 > 0.75$). *P* values were corrected for multiple testing (shown in parentheses). *P* values for trio TDT are based on the TDT statistic; *P* values for FBAT are based on the FBAT (−e option) statistic. NA, correction not applied.

TABLE 3
Haplotype structures of the LD block with their frequencies and association with obesity in the general population (MONICA Augsburg LVH substudy)

Haplotypes						Frequency			
ID	rs509035	rs572169	rs519384	rs512692	rs863441	Obese	Non-obese	Asymptotic *P*	Empirical *P*
1	G	G	T	A	G	0.61	0.69	0.002	0.002
2	A	A	A	T	C	0.33	0.25	0.002	0.001
3	A	A	T	A	G	0.04	0.04	0.474	0.452

Haplotypes are derived from the five SNPs contributing to the high-LD block ($r^2 > 0.75$). Asymptotic *P* values are based on haplotype trend regression analysis; empirical *P* values are based on 50,000 permutations. The permutation test was implemented in the haplotype trend regression analysis program.

these five SNPs and the five-SNP haplotypes. We report linkage between all five SNPs and BMI and, furthermore, provide weak yet suggestive evidence for transmission disequilibrium for the minor alleles of the SNPs as well as for the two most common five-SNP haplotypes with the obesity affection status. The replication of these findings in an independent sample of the general population further supports an association of *GHSR* gene variants with human obesity. Moreover, we report that our haplotypes or one in LD with them could account in part for the observed linkage signal. Thus, our data implicate common haplotypes in this gene region in the pathogenesis of human obesity.

We initially determined the extent of the high-LD region by covering the entire gene region, including the surrounding genomic regions close to the neighboring genes, with SNPs. The identified high-LD region encompasses part of the intron, exon 1, and the 5′ adjacent region extending 8.8 kb past the 5′ end of the gene, but not encompassing the flanking genes. Therefore, it is unlikely that the association between variants of the *GHSR* gene and obesity is seen because of LD with the proper causal mutation in one of the neighboring genes. In addition, we tested the SNPs that were not included in the high-LD block for association (data not shown) and found that none of the SNPs showed evidence for association with the obesity affection status or BMI. This observation supports the hypothesis that genetic variations within the LD block encompassing the *GHSR* gene, and not within neighboring genes, are related to our obesity phenotypes.

We focused only on common sequence variants, as it is more likely that these variants play a role in the general population. Furthermore, we included SNPs located in noncoding and intergenic regions rather than exclusively focusing on the coding region. This strategy was driven by the hypothesis that variations underlying complex diseases would not be limited to the structure of the encoded protein; that is, they could be due to variants leading to altered gene expression (43). Gene regulation is the result of the combinatorial action of multiple transcription factors binding at multiple sites in and near a gene and therefore can be affected by multiple SNPs. In fact, it has been recently demonstrated that gene regulatory elements reside in noncoding and intergenic regions (44,45). These enhancers are able to modulate gene expression over long distances, turning intergenic regions into reservoirs for sequence elements containing important functions (46). Little is known about the impact of sequence variations in these regions. In our study, SNPs located in the intergenic region past the 5′ end of the *GHSR* gene showed stronger association than the SNPs located in the coding or intronic region of the gene. These data suggest that the promoter, regulatory elements or transcriptional initiation could be involved. The exclusive screening of the coding region could be one possible explanation for the lack of association in the study by Wang et al. (42), who were not able to find a relation between two coding SNPs of the *GHSR* gene (one of them, rs 572169, was investigated in the present study) and obesity in children and adolescents. The extremely young age of Wang et al.'s population and the missing haplotype analysis offer further explanations for those researchers' negative results.

The fact that we showed a consistent association between the five SNPs and the two most frequent five-marker haplotypes with the obesity affection status and quantitative BMI in families, the entire MONICA LVH substudy sample, and matched case and control subjects argues against any reasonable likelihood that the findings are artifacts of population stratification or multiple testing. There were no differences with respect to the SNP or haplotype allele frequencies between the two study cohorts and almost all significant *P* values remained after correction for multiple testing using a method for SNPs in LD or running 50,000 permutations. Because our region showed high LD, we did not use the conventional Bonferroni correction, which assumes independence of markers and therefore would make the correction overly conservative (37).

Although the associated minor allele haplotype (haplotype 2) seems to confer susceptibility to obesity, the major allele haplotype (haplotype 1) acts in a reverse fashion by lowering the risk of obesity. Either effect is present independent of carrying one or two copies. However, the effect is strongest in those presenting with two copies of the respective haplotype and decreases with the number of copies.

It is interesting that the ghrelin receptor, encoded by the *GHSR* gene, along with its endogenous ligand ghrelin provides the only hormonal, appetite-stimulatory input that counterbalances a large number of inhibitory signals that are mediated by leptin, insulin, and $PYY_{3–36}$ (14,47,48). *GHSR* is expressed in neuropeptide Y− and agouti-related protein−containing neurons in the hypothalamus that respond to ghrelin by increasing their firing rate (16). Recently, it was shown that during fasting, *GHSR* expression is increased eightfold, which would be expected to result in an increase in receptor signaling and thereby an increase in appetite (49). Accordingly, genetic variations in the ghrelin receptor gene, and thereby altered expression of the receptor protein, should, in turn, result in altered

signaling and consequently altered regulation of appetite. Thus, increased ghrelin receptor expression should be expected to be associated with obesity. Furthermore, it was recently shown that the ghrelin receptor exhibits a high constitutive activity signal of ~50% efficacy between meals and thus provides a set point for food intake between meals (15). An increase in this constitutive activity based on genetic variation, such as the "susceptible" haplotype 2, could result in decreased sensitivity to the multiple inhibitory signals and consequently promote snack-eating behavior between meals. On the other hand, drugs blocking this constitutive activity of the ghrelin receptor might reduce the craving for desserts and intermeal snacks by increasing sensitivity to inhibitory signals (16).

Thus, it seems convincing that genetic variations in the ghrelin receptor gene may change either ghrelin receptor expression or receptor properties and thereby have an effect on appetite regulation by altered signaling, altered response to ghrelin, or an impaired capability to counterbalance inhibitory signals. A greater susceptibility to obesity could be the consequence. Additional functional studies are needed to clarify whether individuals carrying the "susceptible" haplotype present with altered receptor activity by, for example, measuring the inositol (1,4,5)-triphosphate turnover or determining the activation of cAMP responsive element–mediated gene transcription (15,50). Along a similar line, it should be investigated whether those individuals carrying the "nonsusceptible" haplotype exhibit more favorable receptor properties.

In summary, our work offers a comprehensive analysis of the LD structure, common genetic variants, and haplotypes within the *GHSR* gene region. To our knowledge, these data are the first to demonstrate linkage and association of genetic variants within the *GHSR* gene region and human obesity. The findings of linkage together with both transmission disequilibrium in families and the replication of this association in an independent population provide evidence that variants within the *GHSR* gene region might influence susceptibility to obesity and be involved in the pathogenesis of this complex disease. As we focus on common SNPs and haplotypes, the conclusions should be of great importance for a significant proportion of the population. Moreover, they provide background for the development of efficient antiobesity drugs, especially for individuals with the "susceptible" haplotype.

TABLE 4
Summary of the association of SNPs in the total MONICA Augsburg LVH substudy sample as well as in matched case and control subjects

SNP	Case subjects				Control subjects				Allele 2 vs. 1		Genotype 22 vs. 11		Genotypes 22 + 12 vs. 11	
	11	12	22	MAF	11	12	22	MAF	OR (95% CI)	P	OR (95% CI)	P	OR (95% CI)	P
rs509035														
Matched	47	59	17	0.38	173	166	21	0.29	1.50 (1.10–2.03)	0.009	2.98 (1.46–6.10)	0.002	1.50 (0.99–2.27)	0.058
Total study	49	65	17	0.38	424	375	69	0.30	1.45 (1.11–1.90)	0.007	2.13 (1.16–3.91)	0.013	1.60 (1.10–2.33)	0.015
rs572169														
Matched	50	62	17	0.37	182	157	25	0.28	1.49 (1.11–2.01)	0.009	2.48 (1.24–4.94)	0.009	1.58 (1.05–2.38)	0.028
Total study	52	67	17	0.37	451	371	80	0.29	1.42 (1.09–1.85)	0.010	1.84 (1.02–3.35)	0.042	1.62 (1.12–2.34)	0.011
rs519384														
Matched	56	60	15	0.34	207	155	16	0.25	1.59 (1.18–2.16)	0.003	3.47 (1.62–7.44)	0.0009	1.62 (1.09–2.42)	0.018
Total study	59	65	15	0.34	501	361	58	0.26	1.48 (1.13–1.94)	0.004	2.20 (1.17–4.12)	0.012	1.62 (1.13–2.33)	0.008
rs512692														
Matched	57	55	15	0.33	200	154	15	0.25	1.51 (1.11–2.06)	0.008	3.51 (1.62–7.61)	0.0009	1.45 (0.97–2.18)	0.070
Total study	60	60	15	0.33	482	356	56	0.26	1.41 (1.07–1.86)	0.014	2.15 (1.15–4.04)	0.015	1.46 (1.02–2.11)	0.040
rs863441														
Matched	56	56	18	0.35	204	149	13	0.24	1.74 (1.28–2.37)	0.0003	5.04 (2.33–10.9)	0.00001	1.66 (1.11–2.49)	0.013
Total study	59	61	18	0.35	493	355	55	0.26	1.56 (1.20–2.04)	0.001	2.74 (1.51–5.00)	0.0007	1.61 (1.12–2.31)	0.009

SNP	Allele 2 vs. 1		Genotype 22 vs. 11		Genotypes 22 + 12 vs. 11	
	$P_{(corr)}$		$P_{(corr)}$		$P_{(corr)}$	
rs509035						
Matched	0.013		0.003		0.085	
Total study	0.010		0.019		0.022	
rs572169						
Matched	0.013		0.013		0.041	
Total study	0.015		0.062		0.016	
rs519384						
Matched	0.004		0.001		0.026	
Total study	0.006		0.018		0.012	
rs512692						
Matched	0.012		0.001		0.103	
Total study	0.021		0.022		0.059	
rs863441						
Matched	0.0004		0.00002		0.019	
Total study	0.002		0.001		0.013	

SNPs shown as rs numbers from the dbSNP database. Odds ratios (ORs) with 95% CIs are based on the Armitage test for trend. MAF, minor allele frequency; $P_{(corr)}$, P value corrected for multiple testing.

ACKNOWLEDGMENTS

This study was supported by National Institutes of Health Grants R01-DK-54026 and M01-RR-00058 and grants from the Deutsche Forschungsgemeinschaft (BA2227/1-1) and the Deutsche Akademie fuer Naturforscher Leopoldina (BMBF-LPD9901/8-72, to M.F.).

We would like to thank all TOPS members and all volunteers of the MONICA Augsburg study for their participation. We also would like to acknowledge Jacqueline Marks, Roland James, and John Blimke for assistance in recruiting and phenotyping the patients, establishing the databases, and preparing the DNA samples. Moreover, we thank Jill Stein for helping to establish the invader assays as well as Gretta Borchardt, Regina Cole, Martina Köhler,

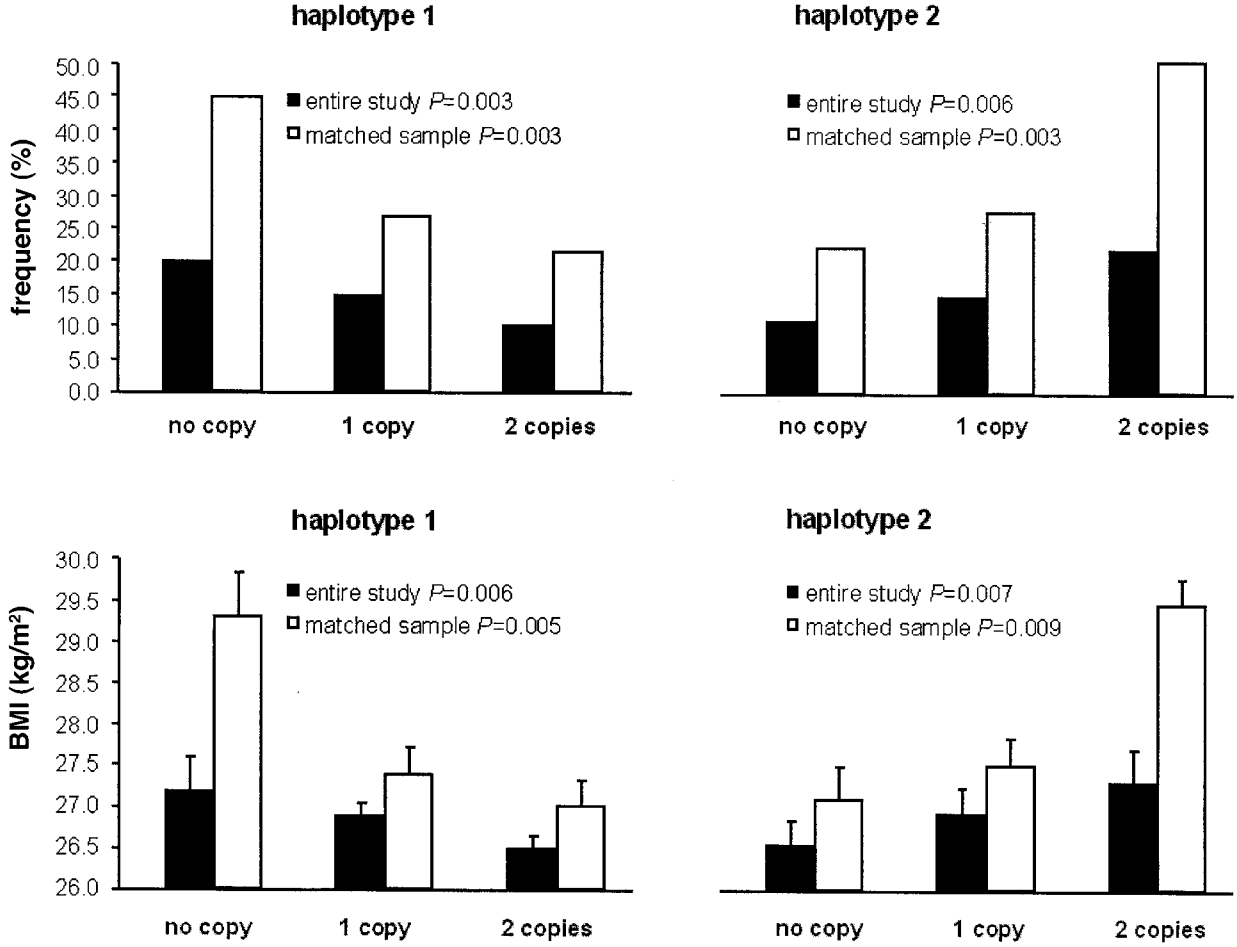


FIG. 3. Association of the number of copies of haplotypes 1 and 2 with obesity (*top*) and BMI (*bottom*) in matched case and control subjects and the entire study sample, respectively. *P* values were generated by trend regression and are based on 50,000 permutations.

and Josef Simon for excellent technical assistance and managing the lab supply.

REFERENCES

1. Frayn KN, Coppack SW: Insulin resistance, adipose tissue and coronary heart disease. *Clin Sci (Lond)* 82:1–8, 1992
2. Flegal KM, Carroll MD, Ogden CL, Johnson CL: Prevalence and trends in obesity among US adults, 1999–2000. *JAMA* 288:1723–1727, 2002
3. Allison DB, Kaprio J, Korkeila M, Koskenvuo M, Neale MC, Hayakawa K: The heritability of body mass index among an international sample of monozygotic twins reared apart. *Int J Obes Relat Metab Disord* 20:501–506, 1996
4. 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
5. Vionnet N, Hani El H, Dupont S, Gallina S, Francke S, Dotte S, De Matos F, Durand E, Lepretre 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 1q21–q24. *Am J Hum Genet* 67:1470–1480, 2000
6. Rice T, Chagnon YC, Perusse L, Borecki IB, Ukkola O, Rankinen T, Gagnon J, Leon AS, Skinner JS, Wilmore JH, Bouchard C, Rao DC: A genomewide linkage scan for abdominal subcutaneous and visceral fat in black and white families: the HERITAGE Family Study. *Diabetes* 51:848–855, 2002
7. Wu X, Cooper RS, Borecki I, Hanis C, Bray M, Lewis CE, Zhu X, Kan D, Luke A, Curb D: A combined analysis of genomewide linkage scans for body mass index from the National Heart, Lung, and Blood Institute Family Blood Pressure Program. *Am J Hum Genet* 70:1247–1256, 2002
8. Zhu X, Cooper RS, Luke A, Chen G, Wu X, Kan D, Chakravarti A, Weder A: A genome-wide scan for obesity in African-Americans. *Diabetes* 51:541–544, 2002
9. Luke A, Wu X, Zhu X, Kan D, Su Y, Cooper R: Linkage for BMI at 3q27 region confirmed in an African-American population. *Diabetes* 52:1284–1287, 2003
10. Kojima M, Hosoda H, Date Y, Nakazato M, Matsuo H, Kangawa K: Ghrelin is a growth-hormone-releasing acylated peptide from stomach. *Nature* 402:656–660, 1999
11. Petersenn S, Rasch AC, Penshorn M, Beil FU, Schulte HM: Genomic structure and transcriptional regulation of the human growth hormone secretagogue receptor. *Endocrinology* 142:2649–2659, 2001
12. Tschop M, Smiley DL, Heiman ML: Ghrelin induces adiposity in rodents. *Nature* 407:908–913, 2000
13. Nakazato M, Murakami N, Date Y, Kojima M, Matsuo H, Kangawa K, Matsukura S: A role for ghrelin in the central regulation of feeding. *Nature* 409:194–198, 2001
14. Cummings DE, Schwartz MW: Genetics and pathophysiology of human obesity. *Annu Rev Med* 54:453–471, 2003
15. Holst B, Cygankiewicz A, Jensen TH, Ankersen M, Schwartz TW: High constitutive signaling of the ghrelin receptor–identification of a potent inverse agonist. *Mol Endocrinol* 17:2201–2210, 2003
16. Holst B, Schwartz TW: Constitutive ghrelin receptor activity as a signalling set-point in appetite regulation. *Trends Pharmacol Sci* 25:113–117, 2004
17. Wren AM, Seal LJ, Cohen MA, Brynes AE, Frost GS, Murphy KG, Dhillo WS, Ghatei MA, Bloom SR: Ghrelin enhances appetite and increases food intake in humans. *J Clin Endocrinol Metab* 86: 5992, 2001
18. Wren AM, Small CJ, Abbott CR, Dhillo WS, Seal LJ, Cohen MA, Batterham RL, Taheri S, Stanley SA, Ghatei MA, Bloom SR: Ghrelin causes hyperphagia and obesity in rats. *Diabetes* 50:2540–2547, 2001
19. Broglio F, Arvat E, Benso A, Gottero C, Muccioli G, Papotti M, van der Lely AJ, Deghenghi R, Ghigo E: Ghrelin, a natural GH secretagogue produced by

the stomach, induces hyperglycemia and reduces insulin secretion in humans. *J Clin Endocrinol Metab* 86:5083–5086, 2001
20. Tschop M, Weyer C, Tataranni PA, Devanarayan V, Ravussin E, Heiman ML: Circulating ghrelin levels are decreased in human obesity. *Diabetes* 50:707–709, 2001
21. Cummings DE, Weigle DS, Frayo RS, Breen PA, Ma MK, Dellinger EP, Purnell JQ: Plasma ghrelin levels after diet-induced weight loss or gastric bypass surgery. *N Engl J Med* 346:1623–1630, 2002
22. Marzullo P, Verti B, Savia G, Walker GE, Guzzaloni G, Tagliaferri M, Di Blasio A, Liuzzi A: The relationship between active ghrelin levels and human obesity involves alterations in resting energy expenditure. *J Clin Endocrinol Metab* 89:936–939, 2004
23. WHO MONICA Project Principal Investigators: The World Health Organization MONICA Project (Monitoring Trends and Determinants in Cardiovascular Disease): a major international collaboration. *J Clin Epidemiol* 41:105–114, 1988
24. Keil U, Stieber J, Doring A, Chambless L, Hartel U, Filipiak B, Hense HW, Tietze M, Gostomzyk JG: The cardiovascular risk factor profile in the study area Augsburg: results from the first MONICA survey 1984/85. *Acta Med Scand Suppl* 728:119–128, 1988
25. Reich DE, Gabriel SB, Altshuler D: Quality and completeness of SNP databases. *Nat Genet* 33:457–458, 2003
26. Carlson CS, Eberle MA, Rieder MJ, Smith JD, Kruglyak L, Nickerson DA: Additional SNPs and linkage-disequilibrium analyses are necessary for whole-genome association studies in humans. *Nat Genet* 33:518–521, 2003
27. Olivier M, Chuang LM, Chang MS, Chen YT, Pei D, Ranade K, de Witte A, Allen J, Tran N, Curb D, Pratt R, Neefs H, de Arruda Indig M, Law S, Neri B, Wang L, Cox DR: High-throughput genotyping of single nucleotide polymorphisms using new biplex invader technology. *Nucleic Acids Res* 30: e53, 2002
28. Devlin B, Risch N: A comparison of linkage disequilibrium measures for fine-scale mapping. *Genomics* 29:311–322, 1995
29. Ardlie KG, Kruglyak L, Seielstad M: Patterns of linkage disequilibrium in the human genome. *Nat Rev Genet* 3:299–309, 2002
30. O'Connell JR, Weeks DE: PedCheck: a program for identification of genotype incompatibilities in linkage analysis. *Am J Hum Genet* 63:259–266, 1998
31. 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
32. 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
33. Horvath S, Wei E, Xu X, Palmer LJ, Baur M: Family-based association test method: age of onset traits and covariates. *Genet Epidemiol* 21 (Suppl 1):S403–S408, 2001
34. Lake SL, Blacker D, Laird NM: Family-based tests of association in the presence of linkage. *Am J Hum Genet* 67:1515–1525, 2000
35. Slatkin M, Excoffier L: Testing for linkage disequilibrium in genotypic data using the expectation-maximization algorithm. *Heredity* 76:377–383, 1996
36. Horvath S, Xu X, Lake SL, Silverman EK, Weiss ST, Laird NM: Family-based tests for associating haplotypes with general phenotype data: application to asthma genetics. *Genet Epidemiol* 26:61–69, 2004
37. Nyholt DR: A simple correction for multiple testing for single-nucleotide polymorphisms in linkage disequilibrium with each other. *Am J Hum Genet* 74:765–769, 2004
38. Fulker DW, Cherny SS, Sham PC, Hewitt JK: Combined linkage and association sib-pair analysis for quantitative traits. *Am J Hum Genet* 64:259–267, 1999
39. Zhang K, Sun F, Zhao H: HAPLORE: a program for haplotype reconstruction in general pedigrees without recombination [article online], 2004. Available from http://bioinformatics.med.yale.edu. Accessed 1 August 2004
40. Sasieni PD: From genotypes to genes: doubling the sample size. *Biometrics* 53:1253–1261, 1997
41. Zaykin DV, Westfall PH, Young SS, Karnoub MA, Wagner MJ, Ehm MG: Testing association of statistically inferred haplotypes with discrete and continuous traits in samples of unrelated individuals. *Hum Hered* 53:79–91, 2002
42. Wang HJ, Geller F, Dempfle A, Schauble N, Friedel S, Lichtner P, Fontenla-Horro F, Wudy S, Hagemann S, Gortner L, Huse K, Remschmidt H, Bettecken T, Meitinger T, Schafer H, Hebebrand J, Hinney A: Ghrelin receptor gene: identification of several sequence variants in extremely obese children and adolescents, healthy normal-weight and underweight students, and children with short normal stature. *J Clin Endocrinol Metab* 89:157–162, 2004
43. Knight JC: Functional implications of genetic variation in non-coding DNA for disease susceptibility and gene regulation. *Clin Sci (Lond)* 104:493–501, 2003
44. Pennacchio LA, Rubin EM: Genomic strategies to identify mammalian regulatory sequences. *Nat Rev Genet* 2:100–109, 2001
45. Mattick JS: RNA regulation: a new genetics? *Nat Rev Genet* 5:316–323, 2004
46. Nobrega MA, Ovcharenko I, Afzal V, Rubin EM: Scanning human gene deserts for long-range enhancers. *Science* 302:413, 2003
47. Cummings DE, Foster KE: Ghrelin-leptin tango in body-weight regulation. *Gastroenterology* 124:1532–1535, 2003
48. Barazzoni R, Zanetti M, Stebel M, Biolo G, Cattin L, Guarnieri G: Hyperleptinemia prevents increased plasma ghrelin concentration during short-term moderate caloric restriction in rats. *Gastroenterology* 124:1188–1192, 2003
49. Kim MS, Yoon CY, Park KH, Shin CS, Park KS, Kim SY, Cho BY, Lee HK: Changes in ghrelin and ghrelin receptor expression according to feeding status. *Neuroreport* 14:1317–1320, 2003
50. Matthews RP, Guthrie CR, Wailes LM, Zhao X, Means AR, McKnight GS: Calcium/calmodulin-dependent protein kinase types II and IV differentially regulate CREB-dependent gene expression. *Mol Cell Biol* 14:6107–6116, 1994