

explain our data. First, overall similar expression of the *RET* genes could be achieved through assemblage of analogously acting, although not orthologous, enhancers. A second, more parsimonious, explanation is that orthologous enhancer elements control expression of both *RET* genes, but have evolved beyond recognition through small changes in TFBSs, rearrangement of sites within enhancers, or multiple coevolved changes. Examination of enhancer evolution in *Drosophila* species reveals examples of these types of sequence changes, confounding traditional sequence alignment approaches while preserving enhancer function across species (29–31). Comparison of human and mouse enhancer sequences suggests that similar widespread turnover of TFBSs is observed in vertebrate evolution (28), although there is no corresponding functional data to confirm that such changes occur while preserving the function of the enhancers. Our data cannot distinguish between these two models; however, it must be the case that largely the same set of transcription factors regulate expression of either gene, and the binding of these is conserved from mammalian to teleost enhancer elements, which allows the HCSs to function in zebrafish. These data may now significantly alter the manner in

which the biological relevance of vertebrate noncoding sequences is evaluated.

References and Notes

1. D. Boffelli, M. A. Nobrega, E. M. Rubin, *Nat. Rev. Genet.* **5**, 456 (2004).
2. A. Woolfe *et al.*, *PLoS Biol.* **3**, e7 (2005).
3. M. A. Nobrega, I. Ovcharenko, V. Afzal, E. M. Rubin, *Science* **302**, 413 (2003).
4. S. Bagheri-Fam, C. Ferraz, J. Demaille, G. Scherer, D. Pfeifer, *Genomics* **78**, 73 (2001).
5. N. Baroukh *et al.*, *Mamm. Genome* **16**, 91 (2005).
6. F. Poulin *et al.*, *Genomics* **85**, 774 (2005).
7. E. de la Calle-Mustienes *et al.*, *Genome Res.* **15**, 1061 (2005).
8. A. Sandelin *et al.*, *BMC Genomics* **5**, 99 (2004).
9. G. Bejerano *et al.*, *Science* **304**, 1321 (2004).
10. V. Veeramachaneni, W. Makalowski, *Nucleic Acids Res.* **33**, D442 (2005).
11. A. S. McCallion, A. Chakravarti, in *Inborn Errors of Development*, C. Epstein, R. Erikson, A. Wynshaw-Boris, Eds. (Oxford Univ. Press, Oxford, 2004).
12. M. Hahn, J. Bishop, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 1053 (2001).
13. C. V. Marcos-Gutierrez, S. W. Wilson, N. Holder, V. Pachnis, *Oncogene* **14**, 879 (1997).
14. B. W. Bigrove, D. W. Raible, V. Walter, J. S. Eisen, D. J. Grunwald, *J. Neurobiol.* **33**, 749 (1997).
15. V. Pachnis, B. Mankoo, F. Costantini, *Development* **119**, 1005 (1993).
16. E. S. Emison *et al.*, *Nature* **434**, 857 (2005).
17. A. S. McCallion *et al.*, *Cold Spring Harb. Symp. Quant. Biol.* **68**, 373 (2003).
18. C. S. Kashuk *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 8949 (2005).
19. K. A. Frazer, L. Pachter, A. Poliakov, E. M. Rubin, I. Dubchak, *Nucleic Acids Res.* **32**, W273 (2004).
20. Materials and methods are available as supporting material on *Science* Online.
21. E. A. Grice, E. S. Rochelle, E. D. Green, A. Chakravarti, A. S. McCallion, *Hum. Mol. Genet.* **14**, 3837 (2005).
22. S. Ellingsen *et al.*, *Development* **132**, 3799 (2005).
23. S. Fisher, E. A. Grice, R. M. Vinton, S. L. Besling, A. S. McCallion, unpublished observations.
24. See http://lagan.stanford.edu/lagan_web/index.shtml.
25. See <http://genome.ucsc.edu/cgi-bin/hgBlat>.
26. DanRer2 (www.sanger.ac.uk/Projects/D_eriol/), accessed June 2004.
27. B. Gottgens *et al.*, *Nat. Biotechnol.* **18**, 181 (2000).
28. L. A. Pennacchio *et al.*, *Science* **294**, 169 (2001).
29. B. P. Berman *et al.*, *Genome Biol.* **5**, R61 (2004).
30. M. Z. Ludwig, C. Bergman, N. H. Patel, M. Kreitman, *Nature* **403**, 564 (2000).
31. M. Z. Ludwig *et al.*, *PLoS Biol.* **3**, e93 (2005).
32. E. T. Dermitzakis, A. G. Clark, *Mol. Biol. Evol.* **19**, 1114 (2002).
33. The authors thank K. Kawakami for his generous gift of the pTZIG vector. This study was supported by a Basil O'Connor award from the March of Dimes (A.S.M.) and grant AR048101 from the NIH (S.F.). Zebrafish sequence data (DanRer2) were produced by the *Danio rerio* Sequencing Group at the Sanger Institute and can be obtained by ftp transfer (<ftp://ftp.ensembl.org/pub/assembly/zebrafish/Zv4release>).

Supporting Online Material

www.sciencemag.org/cgi/content/full/1124070/DC1

Materials and Methods

Tables S1 to S3

References

19 December 2005; accepted 8 March 2006

Published online 23 March 2006;

10.1126/science.1124070

Include this information when citing this paper.

A Common Genetic Variant Is Associated with Adult and Childhood Obesity

Alan Herbert,^{1*} Norman P. Gerry,¹ Matthew B. McQueen,² Iris M. Heid,^{3,4} Arne Pfeufer,^{5,6} Thomas Illig,^{3,4} H.-Erich Wichmann,^{3,4,7} Thomas Meitinger,^{5,6} David Hunter,^{2,8,9} Frank B. Hu,^{2,8,9} Graham Colditz,^{8,9} Anke Hinney,¹⁰ Johannes Hebebrand,¹⁰ Kerstin Koberwitz,^{6,10} Xiaofeng Zhu,¹¹ Richard Cooper,¹¹ Kristin Ardlie,¹² Helen Lyon,^{13,14,15} Joel N. Hirschhorn,^{13,14,15} Nan M. Laird,¹⁶ Marc E. Lenburg,¹ Christoph Lange,^{9,13} Michael F. Christman^{1*}

Obesity is a heritable trait and a risk factor for many common diseases such as type 2 diabetes, heart disease, and hypertension. We used a dense whole-genome scan of DNA samples from the Framingham Heart Study participants to identify a common genetic variant near the *INSIG2* gene associated with obesity. We have replicated the finding in four separate samples composed of individuals of Western European ancestry, African Americans, and children. The obesity-predisposing genotype is present in 10% of individuals. Our study suggests that common genetic polymorphisms are important determinants of obesity.

Obesity is associated with an increased risk of type 2 diabetes mellitus, heart disease, metabolic syndrome, hypertension, stroke, and some forms of cancer (1). It is commonly assessed by calculating an individual's body mass index (BMI) [weight/(height)² in kg/m²] as a surrogate measurement. Individuals with a BMI \geq 25 kg/m² are classified as overweight, and those with a BMI \geq 30 kg/m² are considered obese. Having a BMI

over 25 kg/m² increases the risk of death (2). Presently, 65% of Americans are overweight and 30% are obese (3). Genetic factors contribute significantly to the etiology of obesity (4, 5), with estimates of the heritability of BMI ranging from 30 to 70% (6–9).

To identify common genetic variants associated with elevated BMI, we have studied individuals from the National Heart, Lung, and Blood Institute (NHLBI)–Framingham Heart

Study (FHS) (10). The participants were enrolled from the community without being selected for a particular trait or disease and were followed over 24 years (table S1). In this population, heritability estimates for BMI range between 37 and 54% (11, 12).

Using families from this sample, we performed a genome-wide association analysis, using a testing strategy for quantitative traits in

¹Department of Genetics and Genomics, Boston University Medical School, E613, 715 Albany Street, Boston, MA 02118, USA. ²Department of Epidemiology, Harvard School of Public Health, 677 Huntington Avenue, Boston, MA 02115, USA. ³Institute of Epidemiology, ⁴KORA Group, ⁵Institute of Human Genetics, Forschungszentrum für Umwelt und Gesundheit, D-85764 Neuherberg, Germany. ⁶Institute of Human Genetics, Technical University Munich, D-81671 Munich, Germany. ⁷Institute for Medical Informatics, Biometry and Epidemiology, Ludwig Maximilians University, Munich, Germany. ⁸Nurses Health Study, ⁹Channing Laboratory, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, 181 Longwood Avenue, Boston, MA 02115, USA. ¹⁰Department of Child and Adolescent Psychiatry of the University of Duisburg-Essen, D-45147 Essen, Germany. ¹¹Department of Preventive Medicine and Epidemiology, Loyola University Medical Center, Maywood, IL 60153, USA. ¹²Genomics Collaborative, SeraCare Life Sciences, Cambridge, MA, USA. ¹³Program in Genomics and Divisions of Genetics and Endocrinology, Children's Hospital, Boston, MA 02115, USA. ¹⁴Department of Genetics, Harvard Medical School, Boston, MA 02115, USA. ¹⁵Broad Institute of Harvard and MIT, Cambridge, MA 02139, USA. ¹⁶Department of Biostatistics, Harvard School of Public Health, 655 Huntington Avenue, Boston, MA 02115, USA.

*To whom correspondence should be addressed. E-mail: aherbert@bu.edu (A.H.); mfc@bu.edu (M.F.C.)

family-based designs (13, 14). To bypass the multiple comparisons, we used a two-stage testing strategy implemented in the software package PBAT (15, 16). In the screening step, parental genotypes are used to select single-nucleotide polymorphisms (SNPs) and genetic models that best predict offspring phenotypes. Formally, the screening step estimates the power to detect an association in the second step (Fig. 1 and fig. S1). The second step uses a family-based association test (FBAT) (17) to test the selected SNPs for association with BMI levels using measured offspring genotypes. The FBAT is a generalization of the transmission disequilibrium test (18) and assesses whether the over- or undertransmission of an allele is correlated with offspring phenotypes. Because the allele that is transmitted from the parent to the offspring is selected stochastically, the test step is statistically independent of the screening step, which conditions only on parental genotypes (19, 20). Thus, the power estimates from the screening

step do not bias the significance level of any subsequently computed FBAT statistic in the test step (13, 14, 19, 20). Hence, the FBAT results need to be adjusted only for the number of comparisons performed during the test step. A SNP that reaches significance after adjustment for the number of tests performed in the test step is considered significant at a genome-wide level (13). Following the guidelines of Van Steen *et al.* (13), we tested the 10 SNPs with the highest power estimates in the screening step for association, using the FBAT (21, 22).

We genotyped 116,204 SNPs in 694 participants from the FHS offspring cohort (20). After exclusions, 86,604 SNPs were tested for association with BMI. We used FBAT-PC, which incorporates BMI data across multiple exams and increases our power to detect a genetic effect (14). Of the top 10 SNPs tested under a recessive model [which we determined during the screening procedure had greatest power (Table 1)], only SNP rs7566605 reached overall significance (unadjusted FBAT-PC *P* value, 0.0026). Other genetic models that did not capture the underlying biology of the SNP were less robust. In such cases, rs7566605 was not ranked in the top 10 for power. The frequency of the rs7566605 C allele is 0.37, and

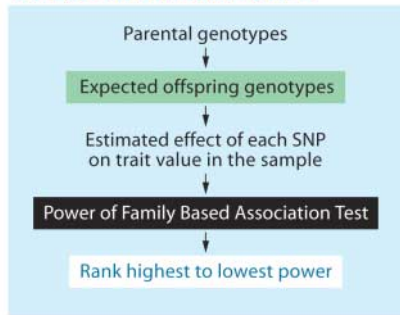
the SNP is in Hardy-Weinberg equilibrium (Table 2). We found no association with BMI for Affymetrix SNPs close to rs7566605 (Fig. 2A and Table 2). This result was expected because none of these SNPs are good proxies for rs7566605 at a r^2 threshold of 0.8 (Fig. 2B) (23).

Another analysis based on a larger sample of 923 FHS individuals, including those already tested, indicated that the SNP is a significant predictor of BMI. For all exams, rs7566605 CC homozygotes are about 1 BMI unit heavier than individuals with GC or GG genotypes ($P < 0.0001$), regardless of sex or age (Fig. 3). CC homozygotes were also more likely to be obese than non-obese [for example, at exam 5, odds ratio (OR) = 1.33, 95% confidence interval (CI) (1.20 to 1.48)].

Many reported associations with common gene polymorphisms have not been replicated, presumably because of factors such as population stratification, inadequate statistical power, and genotyping errors (24). Therefore, we sought to confirm the association between BMI and rs7566605 in five additional unrelated samples of varying ethnicity and age.

We examined the association of rs7566605 with obesity in 3996 participants of the KORA S4 cohort, a population of Western European

A Screen all available SNPs



B Test top 10 SNPs from screen

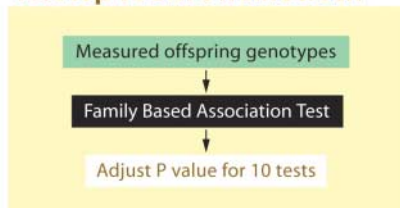


Fig. 1. Family-based design for the whole-genome SNP scan. The same data set was used to both screen and test SNP-phenotype associations (13). **(A)** In the screening step, the difference in offspring trait values due to a particular SNP allele is estimated based on the expected offspring genotypes, computed assuming Mendelian transmission of the parental genotypes. This information, along with the minor allele frequency, is used to compute the power of the FBAT to detect an association. **(B)** In the test step, the top 10–ranked SNPs are evaluated and the FBAT statistic is adjusted for 10 tests. SNPs that are significant after adjustment are also significant genome-wide. The screening and test steps use statistically independent components of the data; hence, the *P* values computed at the test step are not biased by the screening step (20).

Table 1. Screening and testing of SNPs for association with BMI. Genome-wide SNPs (86,604) were screened using parental genotypes to find those likely to affect offspring BMI. The top 10 SNPs from the screening step (ranked by power from most likely to least likely) are shown. These SNPs were tested using offspring genotypes for association with BMI using the FBAT. The rs7566605 SNP is highlighted in bold.

Ranking from screen	SNP	Chromosome	Frequency	Informative families	<i>P</i> value FBAT
1	rs3897510	20p12.3	0.36	30	0.2934
2	rs722385	2q32.1	0.16	15	0.1520
3	rs3852352	8p12	0.33	34	0.7970
4	rs7566605	2q14.1	0.37	39	0.0026
5	rs4141822	13q33.3	0.29	27	0.0526
6	rs7149994	14q21.1	0.35	31	0.0695
7	rs1909459	14q21.1	0.39	38	0.2231
8	rs10520154	15q15.1	0.36	38	0.9256
9	rs440383	15q15.1	0.36	38	0.8860
10	rs9296117	6p24.1	0.40	44	0.3652

Table 2. SNPs genotyped in the region of *INSIG2* along with their minor allele frequency (MAF). Results for the FBAT are given for a recessive model without covariate adjustment. SNPs were also tested for Hardy-Weinberg equilibrium (HWE). The coordinates for *INSIG2a* are 118,562,280 to 118,583,824 and for *INSIG2b* are 118,570,224 to 118,582,624. The rs7566605 SNP is highlighted in bold.

	FBAT (<i>P</i> value)	Physical position	Distance from rs7566605	HWE (<i>P</i> value)	MAF
rs1385923	0.33562	118550089	2166	0.107	0.057
rs10490626	0.07777	118552071	184	0.118	0.087
rs7566605	0.00258	118552255	0	0.197	0.373
rs10490625	0.34606	118575786	23531	0.251	0.063
rs10490624	0.20954	118578722	26467	0.02	0.082

ancestry in a town near Munich, Germany (table S2) (25). Genetic determinants of BMI have been studied in this cohort using candidate gene-based approaches (26). The mean BMI of rs7566605 CC homozygotes under a recessive model was 0.60 kg/m² higher than for GC and GG genotypes combined ($P = 0.008$). The CC genotype was also significantly more frequent in obese individuals [BMI ≥ 30 kg/m², OR = 1.32, 95% CI (1.05 to 1.66), $P = 0.0167$] than in non-obese individuals, again with a recessive

model. These results confirmed the association that we found in the original genome scan of the FHS sample.

We also tested for an association in a case-control study consisting of self-described white or Caucasian subjects from Poland and the United States (1775 cases and 926 controls, table S3). Cases were drawn from the 90th to 97th percentile of the BMI distribution and controls were drawn from the 5th to 12th percentile, with distributions determined for each

combination of gender, country of origin, and age by decade. The results for rs7566605 showed no evidence of heterogeneity between the Polish and U.S. subsamples, so we performed a combined analysis using a Mantel-Haenszel procedure, as described (27). The CC genotype showed a significant association of the SNP with obesity under a recessive model [OR = 1.37, 95% CI (1.05 to 1.78), $P = 0.02$], with an odds ratio similar to that estimated from the KORA data.

The SNP was also genotyped in the Nurses Health Study cohort, using the 2726 control DNAs from two nested case-control studies examining diabetes (28) and breast cancer cases (29) (table S4). The participants are, by self-report, >95% Caucasian. No significant association or trend was observed between rs7566605 and BMI levels. There are fewer individuals with a high BMI in this sample as compared to the KORA and FHS samples (20). The lack of replication thus may reflect a different BMI distribution in the Nurses Health Study or differences in environment and lifestyle.

The rs7566605 SNP was also typed in 368 Western European parent-child trios in which either a child or adolescent offspring was obese (mean BMI percentile 98.4 ± 1.93) (30) (table S5). Analysis using the transmission disequilibrium test (18) revealed an overtransmission of the C allele to the obese offspring ($P = 0.0017$), indicating that rs7566605 is associated with BMI from an early age.

The finding was also replicated in two samples from a self-reported African American population from Maywood, Illinois (table S6). The samples consisted of 866 individuals from a family-based sample consisting of nuclear families and sibships and 402 unrelated individuals (186 chosen from the top and bottom quartiles of the BMI distribution, plus the parents from the family-based sample, whose

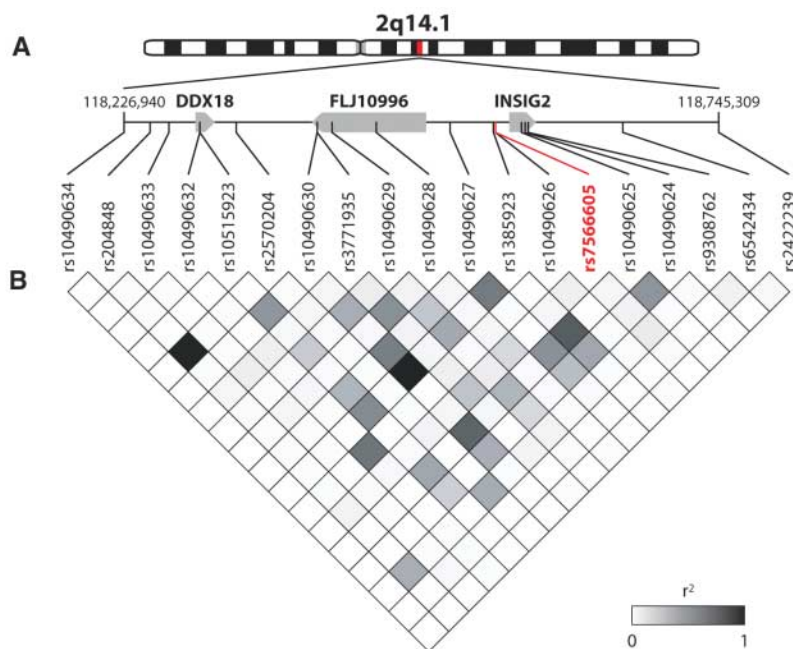


Fig. 2. SNPs present on the Affymetrix 100K chip near 2q14.1 Mapping of 100K SNPs in the chromosomal region from 2q14.1 to 2q14.2 is shown. (A) Known genes in the region of association. (B) A plot showing the pairwise r^2 values indicating the correlation between the SNP genotypes in our sample over this region prepared in Haploview (23). The rs7566605 SNP is shown in red. Black squares, $r^2 = 1$; white squares, $r^2 = 0$; squares in shades of gray, $0 < r^2 < 1$ (the intensity of the gray is proportional to r^2) (for example, r^2 between rs7566605 and rs3771935 equals 0.286).

Table 3. Summary of study results and meta-analysis. All values given are for a recessive model. NHS, Nurses Health Study; TDT, transmission disequilibrium test; PBAT, tools for FBATs.

Study	Design	Total genotyped	Obese	Non-obese	Test	Total number of families	P value
FHS	Family	694			PBAT	288	0.0030
Maywood	Family	866	361	505	PBAT dichotomous	342	0.0090
Maywood	Family	866			PBAT quantitative	342	0.0700
Essen children/adolescents	Trios	1104	368	–	TDT	368	0.0020
<i>Adjustments</i>							
KORA	Cohort	3996			Linear regression	Sex, age	0.0080
NHS	Cohort	2726			Linear regression	Age	–
KORA	Cohort	3996	935	3061	Logistic regression	OR	0.0167
NHS	Cohort	2726	503	2223	Chi-squared test	95% CI	–
American/Polish Caucasian	Case-control	2761	1835	926	Chi-squared test	1.40	0.0200
Maywood	Case-control	398	216	182	Fischer's exact test	2.36	0.0400
Pooled OR (all)		9881	3445	6426	2-tailed Mantel-Haenszel	1.22	0.0080

phenotypic information is not used in the family-based analysis). The samples were dichotomized into obese (BMI ≥ 30 kg/m²) and lean (BMI < 30 kg/m²) individuals. The CC genotype was associated with obesity under recessive models in both the family-based ($P = 0.009$) and unrelated samples (OR = 2.36, $P = 0.04$ by Fisher's exact test), adding further confirmation to our finding.

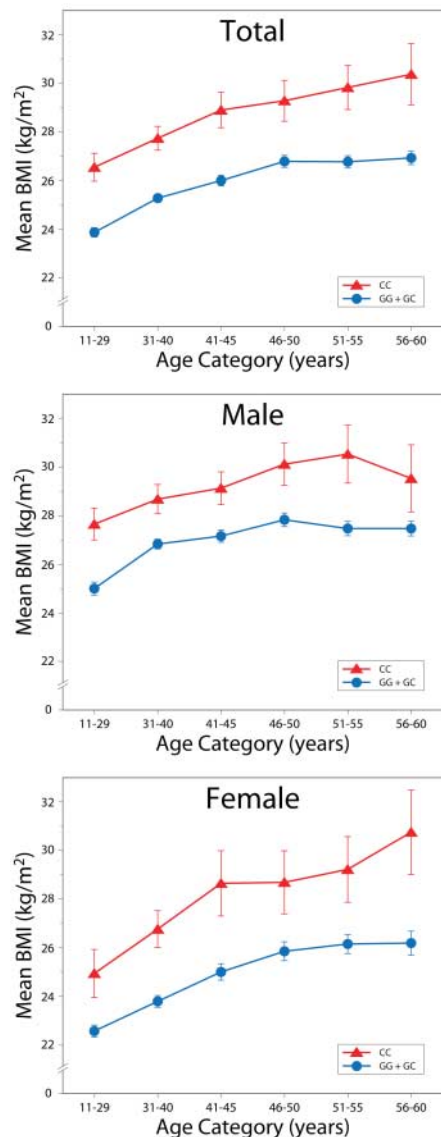


Fig. 3. BMI as a function of age, sex, and rs7566605 genotype. Unadjusted mean values for BMI comparing individuals homozygous for the minor allele (CC) with those that are either heterozygous or homozygous for the major allele (CG or GG) are shown. Data are for 923 related offspring and combine all the measurements from the first five offspring exams. Panels are for the pooled sample and for sex-specific analysis. Data are shown as mean \pm SEM for each age category. Data from individuals over age 60 are omitted because few of those individuals have the CC genotype.

The rs7566605 CC genotype was associated with obesity in three different family-based samples and three studies of unrelated individuals. A meta-analysis of all the case-control samples showed that the CC genotype was significantly associated with obesity under a recessive model, with an OR of 1.22 [95% CI (1.05 to 1.42), $P = 0.008$; Table 3].

The result from both Western Europeans and African Americans suggests that the risk allele predates human migration out of Africa. By comparing chromosomes of individuals with different ancestry from the HapMap project (31) in a 750-kilobase (kb) region near *INSIG2* (*insulin-induced gene 2*), it is apparent that rs7566605 is present on a single haplotype in the European samples (haplotype 2) and on two haplotypes in the Yoruba samples (haplotypes 2 and 8) (Fig. 4; the rs7566605 C allele is shown in red). All of these haplotypes extend to include the intergenic region near *INSIG2* and the nearby *FLJ10996* gene. We also typed the rs11684454 SNP on European haplotype 2, located within an intron of *FLJ10996*. It too was associated with BMI in KORA samples ($P = 0.0093$) and in Western European parent-child trio samples ($P = 0.0015$), confirming that the associated region in Europeans extends at least 70 kb. Nonsynonymous SNPs (rs2229616 and rs17512204) in *FLJ10996* were not associated with BMI in Western European parent-child trio samples. We are not aware of any nonsynonymous SNPs in *INSIG2*. The region harboring the causal variant (Fig. 4) may be narrowed by detailed analysis of African haplotypes 2 and 8, because they are only similar close to rs7566605.

rs7566605 is 10 kb upstream of the transcription start site of *INSIG2*. *INSIG2* is an attractive candidate gene for the quantitative trait locus affecting BMI because its protein product inhibits the synthesis of fatty acid and cholesterol (32). For example, in the ZDF (*fa/fa*) rat model, overexpression of *insig2* in the liver reduces plasma triglyceride levels (33). A model in which altered *insig2* activity leads to obesity by elevating plasma triglyceride levels with subsequent storage in adipose tissue is certainly plausible. Indeed, the *INSIG2* region has been implicated as a factor in obesity by linkage studies in mice (34) and humans (35).

By studying a population not selected for a particular phenotype (the FHS Cohort), we hoped to identify common genetic variants affecting BMI levels. The high frequency of the rs7566605 C allele suggests that it is ancient and is consistent with the hypothesis that such alleles have become deleterious only in modern times (36). Although these variants are likely to carry low relative risk, their impact on health is substantial because of their prevalence in the population.

We expect that other common variants affecting BMI were not detected in our screen because of either lack of power in the screening

step or insufficient coverage of the genome at the density provided by the Affymetrix 100K SNP chip. Taken together, the findings in four of the five samples from different populations using two different study designs confirm a consistent association between the rs7566605 polymorphism and obesity.

References and Notes

1. NIH Publication No. 98-4083 (National Institutes of Health, Bethesda, MD, 1998).
2. J. E. Manson et al., *N. Engl. J. Med.* **333**, 677 (1995).
3. A. A. Hedley et al., *JAMA* **291**, 2847 (2004).
4. J. M. Friedman, *Nat. Med.* **10**, 563 (2004).
5. H. N. Lyon, J. N. Hirschhorn, *Am. J. Clin. Nutr.* **82**, 2155 (2005).
6. C. G. Bell, A. J. Walley, P. Froguel, *Nat. Rev. Genet.* **6**, 221 (2005).
7. I. S. Farooqi, S. O'Rahilly, *Int. J. Obes.* **29**, 1149 (2005).
8. J. Hebebrand, S. Friedel, N. Schauble, F. Geller, A. Hinney, *Obes. Rev.* **4**, 139 (2003).
9. K. Schousboe et al., *Twin Res.* **6**, 409 (2003).
10. W. B. Kannel, *J. Atheroscler. Thromb.* **6**, 60 (2000).
11. The lower estimate is that for narrow-sense heritability that estimates the proportion of phenotypic variance due to additive genetic factors, whereas the upper estimate is for broad-sense heritability, where all genetic factors, such as recessive and dominant alleles, are considered.
12. L. D. Atwood et al., *Am. J. Hum. Genet.* **71**, 1044 (2002).
13. K. Van Steen et al., *Nat. Genet.* **37**, 683 (2005).
14. C. Lange, D. DeMeo, E. K. Silverman, S. T. Weiss, N. M. Laird, *Am. J. Hum. Genet.* **74**, 367 (2004).
15. C. Lange, D. DeMeo, E. K. Silverman, S. T. Weiss, N. M. Laird, *Am. J. Hum. Genet.* **73**, 801 (2003).
16. C. Lange et al., *Hum. Hered.* **56**, 10 (2003).
17. N. M. Laird, S. Horvath, X. Xu, *Genet. Epidemiol.* **19** (suppl. 1), S36 (2000).
18. R. S. Spielman, R. E. McGinnis, W. J. Ewens, *Am. J. Hum. Genet.* **52**, 506 (1993).
19. C. Lange, E. K. Silverman, X. Xu, S. T. Weiss, N. M. Laird, *Biostatistics* **4**, 195 (2003).
20. Supporting material is available on Science Online.
21. S. Horvath, X. Xu, N. M. Laird, *Eur. J. Hum. Genet.* **9**, 301 (2001).
22. C. Lange, D. L. DeMeo, N. M. Laird, *Am. J. Hum. Genet.* **71**, 1330 (2002).
23. J. C. Barrett, B. Fry, J. Maller, M. J. Daly, *Bioinformatics* **21**, 263 (2005).
24. J. N. Hirschhorn, K. Lohmueller, E. Byrne, K. Hirschhorn, *Genet. Med.* **4**, 45 (2002).
25. H. E. Wichmann, C. Gieger, T. Illig, *Das Gesundheitswesen Special Issue* **1**, S26 (2005).
26. I. M. Heid et al., *J. Med. Genet.* **42**, e21 (2005).
27. K. E. Lohmueller, C. L. Pearce, M. Pike, E. S. Lander, J. N. Hirschhorn, *Nat. Genet.* **33**, 177 (2003).
28. F. B. Hu et al., *Diabetes* **53**, 209 (2004).
29. V. W. Setiawan, S. E. Hankinson, G. A. Colditz, D. J. Hunter, I. De Vivo, *Cancer Epidemiol. Biomarkers Prev.* **13**, 213 (2004).
30. A. Hinney et al., *J. Clin. Endocrinol. Metab.* **88**, 4258 (2003).
31. The International HapMap Consortium, *Nature* **426**, 789 (2003).
32. D. Yabe, M. S. Brown, J. L. Goldstein, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 12753 (2002).
33. K. Takaishi, L. Duplomb, M. Y. Wang, J. Li, R. H. Unger, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 7106 (2004).
34. J. M. Cheverud et al., *Diabetes* **53**, 3328 (2004).
35. H. W. Deng et al., *Am. J. Hum. Genet.* **70**, 1138 (2002).
36. D. E. Reich, E. S. Lander, *Trends Genet.* **17**, 502 (2001).
37. The authors thank the participants of the FHS for their contribution and the NHLBI-FHS investigators for providing DNA samples and phenotypic data for our

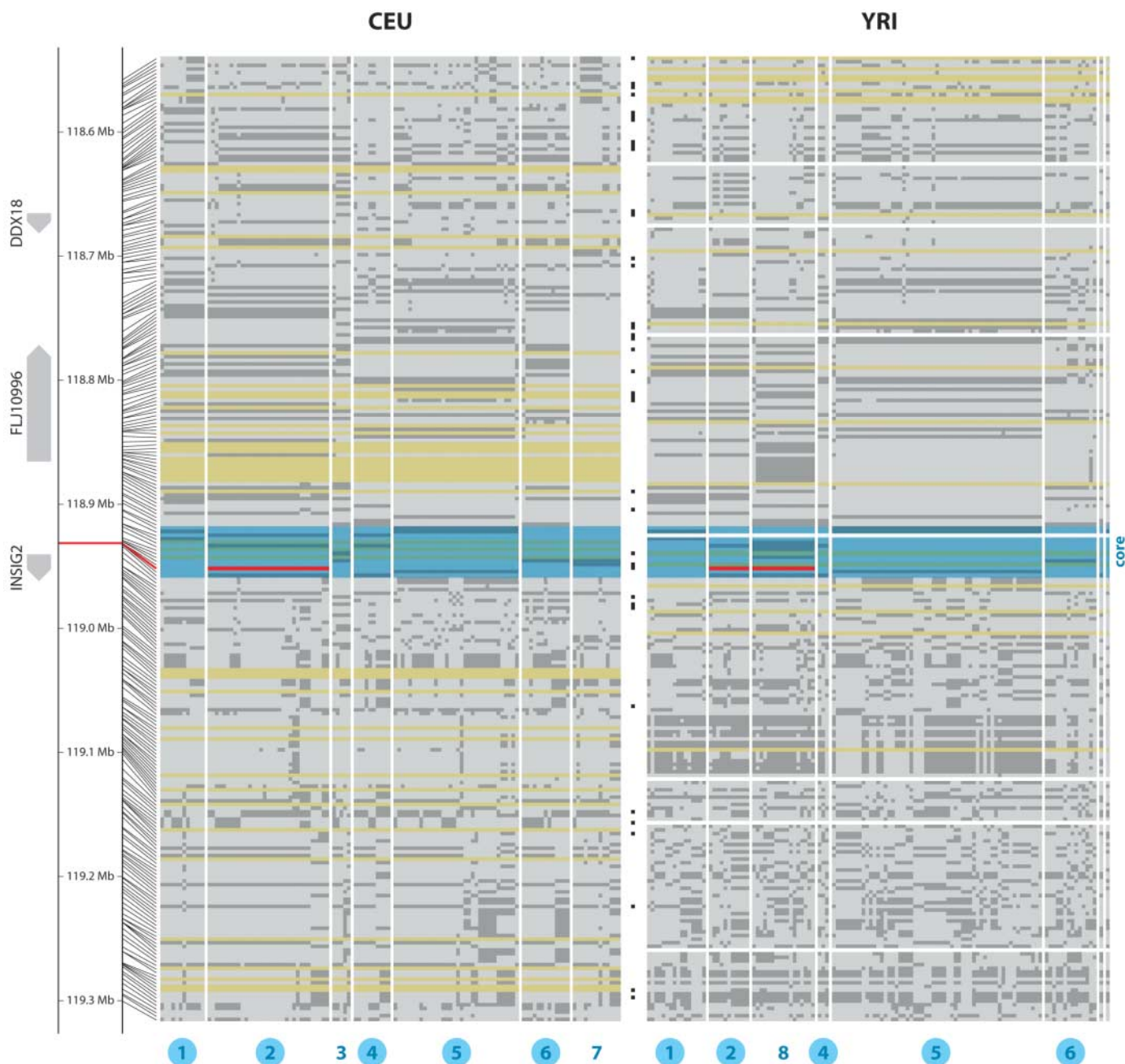


Fig. 4. Linkage disequilibrium around rs7566605 Alignment of the phased HapMap genotypes (release 16c.1) for 120 chromosomes from Utah residents of Northern and Western European ancestry (CEU) and 120 chromosomes from Yoruba in Ibadan, Nigeria (YRI) over a 750-kb region surrounding rs7566605. A core haplotype block containing rs7566605 was identified using Haploview (23) and was used to organize the chromosomes from left to right. The region containing this core haplotype block is highlighted in blue, and each haplotype with a frequency greater than 1% is designated with a

number at the bottom. Haplotypes shared between the CEU and YRI samples are circled. The SNP rs7566605 minor allele is present on one major haplotype in Utah residents (haplotype 2) and on two in Yoruba samples (haplotypes 2 and 8). Minor alleles are represented with dark grey boxes and major alleles with light gray boxes, with the exception of the rs7566605 minor allele (shown in red). SNPs not found to be polymorphic in one of the samples are shaded yellow. SNPs present on the Affymetrix Mapping 100K set are indicated in black between the CEU and YRI panels.

analysis. This work was supported by a Leadership Award from the Whitaker Foundation (A.H. and M.F.C.); grants from the American Diabetes Association to A.H. and an ADA Smith Family Pinnacle Program Project to J.N.H.; and partially by the following NIH grants: R01GM046877 (M.F.C.), K23DK067288 (H.L.), R01HL54485 and R01HL074166 (R.C. and X.Z.), P30DK46200 (principal investigator B. Corkey), and R01MH59532, U01HL65899, and R01HL66289 (N.M.L. and C. L.). A.P., I.H., T.I., T.M., H.-E.W., A. Hi., J. He., and K.K. were

supported by the German Ministry of Education and Research (BMBF, grant 031U212C) through the National Genome Research Network (NGFN, grants 01GR0411, N2NV-S31T10, PGE-S04T13, PGE-S15T04, NHK-S19T19, and NHK-S38T01) and Bioinformatics for the Analysis of Mammalian Genomes (BFAM) and by DFG WI621/12-1. The Nurses Health Study is funded by National Cancer Institute grant CA87969 and thanks C. Guo for data analysis. We thank D. Pellman, N. Levin, and J. Kirkland for comments on the manuscript.

Supporting Online Material

www.sciencemag.org/cgi/content/full/312/5771/279/DC1
Materials and Methods
Fig. S1
Tables S1 to S6
References

10 January 2006; accepted 14 March 2006
10.1126/science.1124779