Genetic Association Analysis of 30 Genes Related to Obesity in a European American Population

Peng Li, M.S.[1], **Hemant K. Tiwari, Ph.D.**[1], **Wan-Yu Lin, Ph.D.**[2], **David B. Allison, Ph.D.**[3], **Wendy K. Chung, M.D.**[4,5], **Rudolph L. Leibel, M.D.**[4,5], **Nengjun Yi, PhD**[1,*], and **Nianjun Liu, PhD**[1,*]

[1]Department of Biostatistics, School of Public Health, University of Alabama at Birmingham, Birmingham, AL

[2]Institute of Epidemiology and Preventive Medicine, College of Public Health, National Taiwan University, Taipei, Taiwan

[3]Nutrition Obesity Research Center, University of Alabama at Birmingham, Birmingham, AL

[4]Columbia University Medical Center, New York, NY

[5]Naomi Berrie Diabetes Center, New York, NY

Abstract

Objective—Obesity, which is frequently associated with diabetes, hypertension, and cardiovascular diseases, is primarily the result of a net excess of caloric intake over energy expenditure. Human obesity is highly heritable, but the specific genes mediating susceptibility in non-syndromic obesity remain unclear. We tested candidate genes in pathways related to food intake and energy expenditure for association with body mass index (BMI).

Methods—We re-analyzed 355 common genetic variants of 30 candidate genes in 7 molecular pathways related to obesity in 1,982 unrelated European Americans from the New York Health Project. Data were analyzed by using a Bayesian hierarchical generalized linear model. The BMIs were log-transformed and then adjusted for covariates including age, age^2, gender, and diabetes status. The single nucleotide polymorphisms (SNPs) were modeled as additive effects.

Results—With the stipulated adjustments, nine SNPs in eight genes were significantly associated with BMI: *GHRL* (rs35683), *AGRP* (rs5030980), *CPE* (rs1946816 and rs4481204), *GLP1R*

---

*Correspondence to: Nianjun Liu. Ph.D. Department of Biostatistics, Ryals Public Health Bldg, 420A, University of Alabama at Birmingham, Birmingham, AL 35294. Phone: (205) 975-9190; Fax: (205) 975-2540; nliu@uab.edu, Nengjun Yi, Ph.D. Department of Biostatistics, Ryals Public Health Bldg, 317F, University of Alabama at Birmingham, Birmingham, AL 35294. Phone: (205) 934-4924; Fax: (205) 975-2540; nyi@uab.edu.

(rs2268641), *HTR2A* (rs912127), *NPY5R* (Y5R1c52), *SOCS3* (rs4969170), and *STAT3* (rs4796793). We also found a gender-by-SNP interaction (rs1745837 in *HTR2A*), which indicated that variants in the gene *HTR2A* had a stronger association with BMI in males. In addition, *NPY1R* was detected as having a significant gene effect even though none of the SNPs in this gene was significant.

Conclusion—Variations in genes *AGRP*, *CPE*, *GHRL*, *GLP1R*, *HTR2A*, *NPY1R*, *NPY5R*, *SOCS3*, and *STAT3* showed modest associations with BMI in European Americans. The pathways in which these genes participate regulate energy intake and thus these associations are mechanistically plausible in this context.

Keywords

Obesity; genetic association; single nucleotide polymorphism (SNP); Bayesian hierarchical generalized linear model (BhGLM)

Introduction

Obesity is primarily a result of a greater intake of calories than the energy expended. Usually defined in adults as a body mass index (BMI) greater than 30 kg/m^2, obesity has become a leading public health concern for both genders, all ages, and all ethnic groups.[1] With current trends, by 2030, some researchers project that 86.3% of American adults will be overweight (25 < BMI 30) or obese (BMI > 30) and that overall 51.1% will be obese.[2] Furthermore, it is predicted that there will be 1.12 billion obese individuals globally.[3] Obesity is highly associated with several major co-morbidities, including diabetes, hypertension, cardiovascular diseases, and kidney diseases.[4] The total healthcare costs attributable to overweight and obesity are large and are projected to grow.[2]

In addition to important environmental factors, evidence from animal models to human studies suggests that genetic makeup also plays an important role in the pathogenesis of obesity.[5, 6] Leptin, a 16-kilodalton, adipocyte-derived hormone, is an important factor in regulating energy intake and energy expenditure through signaling to the hypothalamic pathways that play critical roles in obesity.[7] Advances in genotyping technology have made entire genome screening possible. With genome-wide association studies (GWAS), many associations between obesity and genetic variants have been identified.[8-11]

GWAS are powerful in identifying the association between diseases (or traits) and genetic variants, such as single nucleotide polymorphisms (SNPs). Because the number of SNPs is usually much larger than the number of subjects, traditional multiple regression including all measured SNPs as predictors cannot be used. In published GWAS, simple single-marker analysis of one SNP at a time is commonly used for detecting associations. Yet, the flaws of this strategy are manifold. Because most complex traits are associated with multiple genetic variants, and each variant contributes only a very small portion of risk, single-marker analysis may not have enough power to detect the weak association of each individual variant.[12] Furthermore, human genome is well known for "genetic redundancy" - an abundance of genes having similar roles. Disruption of a single gene may often be selectively neutral.[13] Consequently it may take the malfunction of several genes to produce

a particular disorder.[14] In addition, multiple-testing penalties become a critical issue when the number of SNPs is large. The many tests, of which most correspond to unassociated SNPs, will reduce the required per-test significance level to a very small value, and therefore the power to detect truly associated SNPs is greatly compromised. To address these issues, instead of using the single-marker analysis, a promising approach would be to simultaneously model multiple SNPs in a candidate gene or a pathway.[15] Various statistical methods have been developed to analyze multiple SNPs simultaneously,[16] in which penalized[17, 18] and Bayesian[19, 20] estimation procedures have been proposed and used for variable selection or shrinkage in *large p small n* regressions.

Chung et al.[21] previously investigated 355 common genetic variants in 30 candidate genes in 7 molecular pathways related to obesity in 1,982 unrelated European Americans from the New York Health Project (NYHP) and found that two SNPs in the ghrelin gene (*GHRL*) (rs35682 and rs35683) were associated with BMI in this population. In that study, the 355 SNPs were tested individually by using ordinary least squares regression combined with a false discovery rate (FDR) control for multiple-testing correction.[22] To test the associations of these 355 SNPs on BMI simultaneously, we re-analyze these data here by using a Bayesian hierarchical generalized linear model (BhGLM).[19]

Methods

The New York BMI Data

European American participants were recruited from the NYHP,[21] a prospective cohort study undertaken under the auspices of the Academic Medicine Development Corporation (AMDeC) of New York to analyze general health and cancer-related epidemiology. Approximately 18,000 subjects were enrolled between January 2000 and December 2002 at 14 sites across the five boroughs of New York City. The institutional review boards at each of the 14 sites and at Columbia University approved the protocol. Enrollees had to be 30 years of age or older; reside in New York, New Jersey, or Connecticut; and have a literacy level sufficient to complete a consent and simple follow-up questionnaire. A total of 17,709 subjects with valid data were recruited. Of these, 10,260 (57.9%) were European Americans (non-Hispanic white). Individuals at the extremes of the BMI distribution are potentially the most informative in this context and were therefore selected for initial analysis. Cancer history was by self-report and was confirmed in state cancer registries when possible. All subjects with a history of cancer were eliminated from the study groups to eliminate the possible effect of cancer and its treatment on body weight. European American subjects were sorted by ascending BMI. Subjects below the 10th percentile for BMI were not included to avoid the confounding effects of medical conditions that themselves might cause weight loss. In total, 1,003 gender- and age-matched pairs of European Americans with BMI between the 10th and 30th percentiles or above the 80th percentile were selected for genotyping.

The BMI data set used in this analysis included the descriptive and genetic information of 1,982 unrelated European Americans with valid genotype data from the NYHP.[21] The phenotypic characteristics of the 1,982 individuals are summarized in Table 1. A total of 355 SNPs in 30 candidate genes selected from hypothalamic-arcuate pathways were analyzed.[21]

Briefly, the candidate genes were drawn from seven functional groups related to energy homeostasis: (i) the leptin pathway, which includes the hormone leptin (*LEP*), its receptor (*LEPR*), and downstream signaling molecules (*JAK2*, *SOCS3*, and *STAT3*); (ii) the melanocortin pathway, which includes the agouti-related peptide (*AGRP*), melanocortin 4 receptor (*MC4R*), proopiomelanocortin (*POMC*), and carboxypeptidase E (*CPE*) that processes preprohormones; (iii) the ghrelin pathway, which includes ghrelin (*GHRL*) and its receptor *GHSR*; (iv) the glucagon-like peptide 1 pathway, which includes *GCG* and its receptor *GLP1R*; (v) the neuropeptide Y pathway, which includes the receptors *NPY1R* and *NPY5R*; (vi) the serotonin pathway, which includes the serotonin receptors *HTR2A* and *HTR2C*; (vii) the Bardet-Biedl group, which includes the genes identified for *BBS1, BBS2, ARL6, BBS4, BBS5, MKKS, BBS7*, and *TTC8*. The list of the SNPs and the selection criteria can be found in Chung et al.'s article and supplementary materials.[21]

Statistical Analysis

We jointly estimated the effects of all the SNPs and some covariates, including age, age^2, gender, and diabetes status, and gender-SNP interactions. Because the model included a large number of effects, a classic generalized linear model approach would have unstable estimation. Therefore, we used a Bayesian hierarchical modeling approach, in which the effects are themselves modeled. The hierarchical level which models the effects themselves is referred to as prior distributions.[19] Because the priors constrain effects to lie in a reasonable range and provide additional information to estimate them, the hierarchical approach can yield stable and more precise estimates.[23, 24] The Bayesian hierarchical models described in Yi et al.[19] can simultaneously estimate the genetic effects of individual variants and the overall effects of genetic variants within a gene (referred to as gene effects). With the estimated coefficients and their standard deviations, the p-values can be calculated by using Wald statistics as in the classical framework to test whether an estimated coefficient is significantly different from zero. These p-values have similar interpretation to the classical p-values; however, they are calculated based on both data and priors. The prior distributions and the computational algorithm are fully described in Yi et al. [19] and were implemented in the R package BhGLM (http://www.ssg.uab.edu/bhglm). The package BhGLM provides several prior distributions for setting up the hierarchical models and a computationally efficient algorithm for fitting the models. Because the analyses with different priors yielded similar results (data not shown), we provided the results from the analysis using the commonly used double-exponential prior distributions. For genetic association analyses, it is reasonable to assume that most of the genetic markers have no or very weak effects on the phenotype, whereas only some have noticeable effects. The double-exponential prior with one scale parameter has higher mass around zero (leading to stronger shrinkage towards zero for the effect estimates of markers with no or very weak effects) and thicker tails (leading to less shrinkage for estimates of markers with sizable effects), and the larger scale parameter forces more coefficients to be near zero. In the BhGLM package, the scale parameter of double-exponential prior can be considered as a random variable with a gamma distribution and estimated from the data with an EM algorithm. The estimation of the scale parameter greatly decreases the arbitrariness of the shrinkage than using fixed values for the prior.

We transformed all the input variables to a roughly common scale. The BMIs were log-transformed to normalize the residual distribution and then standardized by centering and dividing by two standard deviations. The covariates included age, gender, and diabetes status, the same as in Chung et al.'s article [21]. Age was rescaled to have a mean of 0 and a standard deviation of 0.5. There were no missing data for the covariates.

Among the 355 SNPs, 17 were not polymorphic in the sense that all study subjects were homozygous for the major allele. These SNPs were excluded from further analyses. Among the remaining 338 SNPs, 154 SNPs contained missing values and the missing rate for each SNP ranged from 0.05% to 0.75%. The SNPs were modeled with additive effects, i.e., the common homozygote, the heterozygote, and the rare homozygote were code as 0, 1, and 2, respectively. The missing genotypes were replaced with an observed mean of the predictor in the whole sample.

To assess the overall effect of multiple genetic variants in a particular gene on the phenotype, the 338 SNPs were grouped into 30 genes on the basis of their chromosomal positions (The gene names, chromosomal locations, coordinates according to NCBI Human Genome Build 35, http://www.ncbi.nlm.nih.gov/mapview/stats/BuildStats.cgi?taxid=9606&build=;35&ver=;1). Gender-SNP interactions were also investigated.

Results

The phenotypic characteristics of the 1,982 European Americans are summarized in Table 1. Among the 1,982 individuals, 1,124 were females and 858 were males. There was no significant gender difference in diabetes prevalence or BMI. The average age of the females was 3 years older than that of the males ($p < 0.0001$).

Gender, age, age^2, and diabetes status were all significant predictors (at a significance level of 0.05) after control for the genetic predictors (Table 2). The positive estimate of the age effect ($p = 0.0156$) and the negative estimate of the age^2 effect ($p = 0.0266$) suggested a positive effect of age on BMI untill a turning point which is about 70 years old. The diabetic individuals had larger BMI values than did the nondiabetic individuals ($p < 0.0001$). After control for other predictors, the males had slightly but significantly larger BMI values than did the females ($p = 0.0028$).

Among the 338 SNPs, rs5030980 in gene *AGRP*, rs1946816 and rs4481204 in *CPE*, rs35683 in *GHRL*, rs912127 in *HTR2A*, rs2268641 in *GLP1R*, Y5R1c52 in *NPY5R*, rs4969170 in *SOCS3*, and rs4796793 in *STAT3* were significantly associated with BMI, either positively or negatively (Table 3).

The associations with individual SNPs could be weak and therefore were difficult to detect when the SNPs were tested separately. The BhGLM method allowed us to investigate the association of a gene by testing the overall association of all SNPs in that gene with the phenotype and at the same time pinpointing the significant SNPs. In our analysis, four genes were significantly associated with BMI: *AGRP*, *CPE*, *GHRL*, and *NPY1R* (Table 4).

A significant interaction between SNP rs1745837 in *HTR2A* and gender was also detected (Table 5). This finding suggested that variants in *HTR2A* could have stronger effects on BMI in males than in females.

Discussion

Obesity, an important risk factor for diabetes, hypertension, cardiovascular diseases, and kidney diseases, has become a global public health concern. Genetic association studies may provide insight into the pathogenesis and prevention of obesity. Chung et al.[21] performed a genetic study on BMI in European Americans and tested the effects of 355 selected SNPs individually, with target genes in target pathways. They found that two SNPs in *GHRL* (rs35682 and rs35683) were associated with BMI in the sample of New York European Americans. Using the same data, we tested the effects of the SNPs simultaneously with a new Bayesian hierarchical model, and found more statistically significant genetic associations. In addition to the SNP in *GHRL* (rs35683), we found that SNPs in *AGRP* (rs5030980), *CPE* (rs1946816 and rs4481204), *GLP1R* (rs2268641), *HTR2A* (rs912127), *NPY5R* (Y5R1c52), *SOCS3* (rs4969170), and *STAT3* (rs4796793) were significantly associated with BMI in the same sample of New York European Americans. Gene effects on BMI were also detected in *AGRP*, *CPE*, *GHRL*, and *NPY1R*. Our results further suggested a gender-by-SNP interaction for rs1745837 in *HTR2A*, which implied that variants in *HTR2A* might have stronger effects on BMI in men than in women. However, only 3% of the variance in BMI could be explained by the variants of these genes on the basis of our models (each SNP contributed only $0.2 \sim 0.3\%$ of the variance on average), which suggests that more genes, their interactions, and their interactions with enviromental factors could be involved in the pathogenesis of obesity.

Ghrelin (*GHRL*) is an endogenous peptide that stimulates growth hormone secretion and enhances appetite. Polymorphisms in *GHRL* have previously been reported to be associated with obesity in different populations.[21, 25-34] In Chung et al.,[21] both rs35682 and rs35683 were found to be significantly associated with BMI on the basis of the individual SNP tests. Because these two SNPs are closely located (within 600 base pairs) in intron 2 and intron 3, they are highly correlated ($r^2 = 0.96$, P < 0.0001, a high level of linkage disequilibrium). Either one can be considered to be the surrogate variable of the other in a multiple regression. Therefore, when the effects of rs35682 and rs35683 were simultaneously modeled with the BhGLM method, only one (rs35683) stood out.

Agouti-related peptide (*AGRP*) is an endogenous antagonist of melanocortin signaling and therefore stimulates appetite. The SNP rs5030980, causing the Thr67Ala mutant, was shown to have a significant association with BMI in our study. A causal effect is plausible, because this is a nonsynonymous mutant. However, the effects of the Thr67Ala mutation are not consistent in the literature.[35-38] The Thr67Ala *AGRP* polymorphism was reported to be associated with lower body weight in Quebec families[37] and in a West African population,[35] whereas no association was found between rs5030980 and BMI in Dutch adults[38] or in a Latvian population.[36] These inconsistent conclusions could be due to the different statistical models used for the analyses, different sample sizes, or different effects in different populations.

The glucagon-like peptide-1 receptor (*GLP1R*) is a key physiological regulator of insulin secretion and is a major therapeutic target for treatment of type II diabetes. Variation in *GLP1R* has been reported to alter insulin secretion in response to exogenous GLP-1 in humans[39] and to influence food intake in mice.[40] Treatment with GLP1R agonists also leads to weight loss in overweight or obese patients with or without type 2 diabetes mellitus.[41] Our results further suggest that genetic variation in *GLP1R* is associated with obesity in European Americans.

Leptin and its downstream signaling molecules *JAK2*, *SOCS3*, and *STAT3* play an important role in regulating energy intake and energy expenditure.[7, 42] Polymorphisms near *SOCS3* were reported to be associated with obesity and glucose homeostasis traits in Hispanic Americans,[43] and *STAT3* polymorphisms were reported to be associated with abdominal obesity in French adults.[44] Our findings further suggest the association of polymorphisms in *SOCS3* and *STAT3* with BMI in European Americans living in the New York area.

Serotonin (5-hydroxytryptamine, or 5-HT) is a neurotransmitter regulating many physiologic processes, such as sleep, appetite, and hormone secretion. Serotonin receptor 2A (*HTR2A*) and 5-HT are involved in regulating cortisol secretion,[45] which may play a pathogenetic role in abdominal obesity.[46] Polymorphisms in *HTR2A* have been found to be associated with abdominal obesity and BMI in Swedish adults[47, 48] and with increased energy and fat intake in French children.[49] We found a novel interaction of an SNP in this gene with gender, which indicated that variants in this gene might have different effects on BMI in men and women.

Neuropeptide Y and neuropeptide Y receptor (*NPY1R* and *NPY5R*) play a key role in the physiological control of energy homeostasis.[50] It has been reported that polymorphisms in *NPY5R* are associated with obesity in Pima Indians.[51] Our models established an association between BMI and polymorphisms in *NPY5R* in a European American population. The effects of individual variants in *NPY1R* might be weak; however, a significant gene effect was detected.

Carboxypeptidase E (*CPE*) is an enzyme that functions in the production and activation of many neuropeptides and peptide hormones.[52] Two SNPs (rs1946816 and rs4481204) were found to be associated with BMI in the European American population by the BhGLM method. Interestingly and consistently, these two SNPs were located in a haplotype block that was associated with BMI by use of a haplotype similarity-based multimarker association test.[53] The BhGLM method also identified a signficant gene effect for *CPE*. In fact, some previous studies have shown that a single point mutation in the *CPE* gene is sufficient to cause multiple disorders in animals, including obesity.[54, 55] In addition, other studies have reported an association of the *CPE* gene with human obesity and diabetes.[56, 57]

Because no genetic associations with obesity for *GLP1R* and *NPY1R* have been reported in humans so far, confirmation in independent study populations is needed, although our results are biologically plausible. Note that most of the SNPs we found to be significantly associated with BMI are located in introns or nontranscribed regions (Table 3); these SNPs may cause changes in splicing or expression. The association of these SNPs may also derive

from the effects of causal variants that are in linkage disequilibrium with them. Because our cross-sectional data came from a prospective cohort, when new data become avalable, a further prospective analysis will be conducted to reinforce the strength of our findings.

The most commonly used approach in genetic association studies is to consider one or a few genetic variants at a time. Despite the ease of implementation, this approach can produce misleading results. In the present study, we used the BhGLM method to simultaneously test the associations of multiple variants. This method put prior information on the parameters to control the compexity of the model and handle high-dimensional and correlated variables. The method has been implemented in the freely available R package BhGLM.[19] This method can be powerful for detecting weak associations of individual SNPs with joint consideration of multiple SNPs. BhGLM provides flexibility by including most versions of the penalized regression procedures with different prior specification or further modeling on the hyperparameters.[19, 20, 58-60] Another advantage of BhGLM is that the parameters and the hyperparameters can be estimated from the data by use of an EM algorithm rather than by giving fixed values, which greatly eliminates the arbitrariness of the shrinkage. Some genetic effects may be too weak to be detected individually and therefore the overall (gene or group) effects will be preferred. This is testified too in our study by the detection of a gene effect of *NPY1R* but no individual SNP effect.

In summary, we used a BhGLM method to test the genetic association of BMI in European Americans living in New York and found that genetic variants in or close to the genes *AGRP*, *CPE*, *GHRL*, *GLP1R*, *HTR2A*, *NPY1R*, *NPY5R*, *SOCS3*, and *STAT3* were significantly associated with BMI. The underlying mechanism of this association, however, requires further investigation.

Acknowledgments

References

1. Flegal KM, Carroll MD, Ogden CL, Curtin LR. Prevalence and Trends in Obesity Among US Adults, 1999-2008. Jama-J Am Med Assoc. 2010; 303(3):235–241.

2. Wang YF, Beydoun MA, Liang L, Caballero B, Kumanyika SK. Will all Americans become overweight or obese? Estimating the progression and cost of the US obesity epidemic. Obesity. 2008; 16(10):2323–2330. [PubMed: 18719634]

3. Kelly T, Yang W, Chen CS, Reynolds K, He J. Global burden of obesity in 2005 and projections to 2030. Int J Obesity. 2008; 32(9):1431–1437.

4. Mokdad AH, Ford ES, Bowman BA, Dietz WH, Vinicor F, Bales VS, et al. Prevalence of obesity, diabetes, and obesity-related health risk factors, 2001. Jama-J Am Med Assoc. 2003; 289(1):76–79.

5. Bell CG, Walley AJ, Froguel P. The genetics of human obesity. Nat Rev Genet. 2005; 6(3):221–234. [PubMed: 15703762]

6. Chung WK, Leibel RL. Considerations Regarding the Genetics of Obesity. Obesity. 2008; 16:S33–S39. [PubMed: 19037210]

7. Mantzoros CS. The role of leptin in human obesity and disease: A review of current evidence. Ann Intern Med. 1999; 130(8):671–680. [PubMed: 10215564]

8. Hofker M, Wijmenga C. A supersized list of obesity genes. Nat Genet. 2009; 41(2):139–140. [PubMed: 19174833]

9. Meyre D, Delplanque J, Chevre JC, Lecoeur C, Lobbens S, Gallina S, et al. Genome-wide association study for early-onset and morbid adult obesity identifies three new risk loci in European populations. Nat Genet. 2009; 41(2):157–159. [PubMed: 19151714]

10. Thorleifsson G, Walters GB, Gudbjartsson DF, Steinthorsdottir V, Sulem P, Helgadottir A, et al. Genome-wide association yields new sequence variants at seven loci that associate with measures of obesity. Nat Genet. 2009; 41(1):18–24. [PubMed: 19079260]

11. Willer CJ, Speliotes EK, Loos RJF, Li SX, Lindgren CM, Heid IM, et al. Six new loci associated with body mass index highlight a neuronal influence on body weight regulation. Nat Genet. 2009; 41(1):25–34. [PubMed: 19079261]

12. Hoggart CJ, Whittaker JC, De Iorio M, Balding DJ. Simultaneous Analysis of All SNPs in Genome-Wide and Re-Sequencing Association Studies. Plos Genet. 2008; 4(7)

13. Kimura, M. The Neutral Theory of Molecular Evolution. Cambridge University Press; Cambridge: p. 1985

14. Shi M, Weinberg CR. How much are we missing in SNP-by-SNP analyses of genome-wide association studies? Epidemiology. 2011; 22(6):845–7. [PubMed: 21878815]

15. Yang X, Deignan JL, Qi H, Zhu J, Qian S, Zhong J, et al. Validation of candidate causal genes for obesity that affect shared metabolic pathways and networks. Nat Genet. 2009; 41(4):415–23. [PubMed: 19270708]

16. de Los Campos G, Hickey JM, Pong-Wong R, Daetwyler HD, Calus MP. Whole-genome regression and prediction methods applied to plant and animal breeding. Genetics. 2013; 193(2): 327–45. [PubMed: 22745228]

17. Tibshirani R. Regression shrinkage and selection via the Lasso. J Roy Stat Soc B Met. 1996; 58(1): 267–288.

18. Wu TT, Chen YF, Hastie T, Sobel E, Lange K. Genome-wide association analysis by lasso penalized logistic regression. Bioinformatics. 2009; 25(6):714–21. [PubMed: 19176549]

19. Yi N, Liu N, Zhi D, Li J. Hierarchical generalized linear models for multiple groups of rare and common variants: jointly estimating group and individual-variant effects. Plos Genet. 2011; 7(12):e1002382. [PubMed: 22144906]

20. Yi N, Xu S. Bayesian LASSO for quantitative trait loci mapping. Genetics. 2008; 179(2):1045–55. [PubMed: 18505874]

21. Chung WK, Patki A, Matsuoka N, Boyer BB, Liu N, Musani SK, et al. Analysis of 30 genes (355 SNPS) related to energy homeostasis for association with adiposity in European-American and Yup'ik Eskimo populations. Human heredity. 2009; 67(3):193–205. [PubMed: 19077438]

22. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate - a Practical and Powerful Approach to Multiple Testing. J Roy Stat Soc B Met. 1995; 57(1):289–300.

23. Gelman, A.; Carlin, J.; Stern, H.; Rubin, D. Bayesian data analysis. Chapman and Hall; London: p. 2003

24. Gelman, AaH. J Data Analysis Using Regression and Multilevel/Hierarchical Models. Cambridge University Press; New York: p. 2007

25. Ando T, Ichimaru Y, Konjiki F, Shoji M, Komaki G. Variations in the preproghrelin gene correlate with higher body mass index, fat mass, and body dissatisfaction in young Japanese women. Am J Clin Nutr. 2007; 86(1):25–32. [PubMed: 17616759]

26. del Giudice EM, Santoro N, Cirillo G, Raimondo P, Grandone A, D'Aniello A, et al. Molecular screening of the ghrelin gene in Italian obese children: the Leu72Met variant is associated with an earlier onset of obesity. Int J Obesity. 2004; 28(3):447–450.

27. Hinney A, Hoch A, Geller F, Schafer H, Siegfried W, Goldschmidt H, et al. Ghrelin gene: Identification of missense variants and a frameshift mutation in extremely obese children and adolescents and healthy normal weight students. J Clin Endocr Metab. 2002; 87(6):2716–2719. [PubMed: 12050239]

28. Jo DS, Kim SL, Kim SY, Hwang PH, Lee KH, Lee DY. Preproghrelin Leu72Met polymorphism in obese Korean children. J Pediatr Endocr Met. 2005; 18(11):1083–1086.

29. Kuzuya M, Ando F, Iguchi A, Shimokata H. Preproghrelin Leu72Met variant contributes to overweight in middle-aged men of a Japanese large cohort. Int J Obesity. 2006; 30(11):1609–1614.

30. Mager U, Lindi V, Lindstrom J, Eriksson JG, Valle TT, Hamalainen H, et al. Association of the Leu72Met polymorphism of the ghrelin gene with the risk of Type 2 diabetes in subjects with impaired glucose tolerance in the Finnish Diabetes Prevention Study. Diabetic Med. 2006; 23(6):685–689. [PubMed: 16759313]

31. Steinle NI, Pollin TI, O'Connell JR, Mitchell BD, Shuldiner AR. Variants in the ghrelin gene are associated with metabolic syndrome in the old order amish. J Clin Endocr Metab. 2005; 90(12):6672–6677. [PubMed: 16204371]

32. Ukkola O, Ravussin E, Jacobson P, Perusse L, Rankinen T, Tschop M, et al. Role of ghrelin polymorphisms in obesity based on three different studies. Obes Res. 2002; 10(8):782–791. [PubMed: 12181387]

33. Vartiainen J, Kesaniemi YA, Ukkola O. Sequencing analysis of ghrelin gene 5 ' flanking region: relations between the sequence variants, fasting plasma total ghrelin concentrations, and body mass index. Metabolism. 2006; 55(10):1420–1425. [PubMed: 16979415]

34. Vivenza D, Rapa A, Castellino N, Bellone S, Petri A, Vacca G, et al. Ghrelin gene polymorphisms and ghrelin, insulin, IGF-I, leptin and anthropometric data in children and adolescents. Eur J Endocrinol. 2004; 151(1):127–133. [PubMed: 15248833]

35. Bonilla C, Panguluri RK, Taliaferro-Smith L, Argyropoulos G, Chen G, Adeyemo AA, et al. Agouti-related protein promoter variant associated with leanness and decreased risk for diabetes in West Africans. Int J Obesity. 2006; 30(4):715–721.

36. Kalnina I, Kapa I, Pirags V, Ignatovica V, Schioth HB, Klovins J. Association between a rare SNP in the second intron of human Agouti related protein gene and increased BMI. Bmc Med Genet. 2009; 10

37. Marks DL, Boucher N, Lanouette CM, Perusse L, Brookhart G, Comuzzie AG, et al. Ala67Thr polymorphism in the agouti-related peptide gene is associated with inherited leanness in humans. Am J Med Genet A. 2004; 126A(3):267–271. [PubMed: 15054840]

38. van Rossum CTM, Pijl H, Adan RAH, Hoebee B, Seidell JC. Polymorphisms in the NPY and AGRP genes and body fatness in Dutch adults. Int J Obesity. 2006; 30(10):1522–1528.

39. Sathananthan A, Man CD, Micheletto F, Zinsmeister AR, Camilleri M, Giesler PD, et al. Common Genetic Variation in GLP1R and Insulin Secretion in Response to Exogenous GLP-1 in Nondiabetic Subjects A pilot study. Diabetes care. 2010; 33(9):2074–2076. [PubMed: 20805279]

40. Kumar KG, Poole AC, York B, Volaufova J, Zuberi A, Richards BKS. Quantitative trait loci for carbohydrate and total energy intake on mouse chromosome 17: congenic strain confirmation and candidate gene analyses (Glo1, Glp1r). Am J Physiol-Reg I. 2007; 292(1):R207–R216.

41. Vilsboll T, Christensen M, Junker AE, Knop FK, Gluud LL. Effects of glucagon-like peptide-1 receptor agonists on weight loss: systematic review and meta-analyses of randomised controlled trials. Bmj. 2012; 344:d7771. [PubMed: 22236411]

42. Morris DL, Rui LY. Recent advances in understanding leptin signaling and leptin resistance. Am J Physiol-Endoc M. 2009; 297(6):E1247–E1259.

43. Talbert ME, Langefeld CD, Ziegler J, Mychaleckyj JC, Haffner SM, Norris JM, et al. Polymorphisms near SOCS3 are associated with obesity and glucose homeostasis traits in Hispanic Americans from the Insulin Resistance Atherosclerosis Family Study. Human genetics. 2009; 125(2):153–62. [PubMed: 19083014]

44. Phillips CM, Goumidi L, Bertrais S, Field MR, Peloso GM, Shen J, et al. Dietary Saturated Fat Modulates the Association between STAT3 Polymorphisms and Abdominal Obesity in Adults. Journal of Nutrition. 2009; 139(11):2011–2017. [PubMed: 19776189]

45. Rittenhouse PA, Bakkum EA, Levy AD, Li Q, Carnes M, Vandekar LD. Evidence That Acth-Secretion Is Regulated by Serotonin(2a/2c) (5-Ht2a/2c) Receptors. J Pharmacol Exp Ther. 1994; 271(3):1647–1655. [PubMed: 7996480]

46. Bjorntorp P, Rosmond P. Obesity and cortisol. Nutrition. 2000; 16(10):924–936. [PubMed: 11054598]

47. Rosmond R, Bouchard C, Bjorntorp P. 5-HT2A receptor gene promoter polymorphism in relation to abdominal obesity and cortisol. Obes Res. 2002; 10(7):585–589. [PubMed: 12105278]

48. Rosmond R, Bouchard C, Bjorntorp P. Increased abdominal obesity in subjects with a mutation in the 5-HT2A receptor gene promoter. Lipids and Insulin Resistance: The Role of Fatty Acid Metabolism and Fuel Partitioning. 2002; 967:571–575.

49. Herbeth B, Aubry E, Fumeron F, Aubert R, Cailotto F, Siest G, et al. Polymorphism of the 5-HT2A receptor gene and food intakes in children and adolescents: the Stanislas Family Study. Am J Clin Nutr. 2005; 82(2):467–470. [PubMed: 16087994]

50. Mashiko S, Moriya R, Ishihara A, Gomori A, Matsushita H, Egashira S, et al. Synergistic interaction between neuropeptide Y-1 and Y-5 receptor pathways in regulation of energy homeostasis. Eur J Pharmacol. 2009; 615(1-3):113–117. [PubMed: 19482021]

51. Jenkinson CP, Cray K, Walder K, Herzog H, Hanson R, Ravussin E. Novel polymorphisms in the neuropeptide-Y Y5 receptor associated with obesity in Pima Indians. Int J Obesity. 2000; 24(5):580–584.

52. Cool DR, Normant E, Shen FS, Chen HC, Pannell L, Zhang Y, et al. Carboxypeptidase E is a regulated secretory pathway sorting receptor: Genetic obliteration leads to endocrine disorders in Cpe(fat) mice. Cell. 1997; 88(1):73–83. [PubMed: 9019408]

53. Lin WY, Tiwari HK, Gao G, Zhang K, Arcaroli JJ, Abraham E, et al. Similarity-based multimarker association tests for continuous traits. Annals of human genetics. 2012; 76(3):246–60. [PubMed: 22497480]

54. Fricker LD, Berman YL, Leiter EH, Devi LA. Carboxypeptidase E activity is deficient in mice with the fat mutation - Effect on peptide processing. Journal of Biological Chemistry. 1996; 271(48):30619–30624. [PubMed: 8940036]

55. Naggert JK, Fricker LD, Varlamov O, Nishina PM, Rouille Y, Steiner DF, et al. Hyperproinsulinaemia in obese fat/fat mice associated with a carboxypeptidase E mutation which reduces enzyme activity. Nat Genet. 1995; 10(2):135–42. [PubMed: 7663508]

56. Jackson RS, Creemers JW, Ohagi S, Raffin-Sanson ML, Sanders L, Montague CT, et al. Obesity and impaired prohormone processing associated with mutations in the human prohormone convertase 1 gene. Nat Genet. 1997; 16(3):303–6. [PubMed: 9207799]

57. Jeffrey KD, Alejandro EU, Luciani DS, Kalynyak TB, Hu XK, Li H, et al. Carboxypeptidase E mediates palmitate-induced beta-cell ER stress and apoptosis. Proceedings of the National Academy of Sciences of the United States of America. 2008; 105(24):8452–8457. [PubMed: 18550819]

58. Friedman J, Hastie T, Tibshirani R. Regularization Paths for Generalized Linear Models via Coordinate Descent. J Stat Softw. 2010; 33(1):1–22. [PubMed: 20808728]

59. Kyung M, Gill J, Ghosh M, Casella G. Penalized Regression, Standard Errors, and Bayesian Lassos. Bayesian Anal. 2010; 5(2):369–411.

60. Sun W, Ibrahim JG, Zou F. Genomewide Multiple-Loci Mapping in Experimental Crosses by Iterative Adaptive Penalized Regression. Genetics. 2010; 185(1):349–U529. [PubMed: 20157003]

List of abbreviations

BMI	Body mass index
NYHP	New York Health Project
GLM	generalized linear model
GWAS	genome-wide association studies
SNPs	single nucleotide polymorphisms
BhGLM	Bayesian hierarchical generalized linear model
GHRL	ghrelin

AGRP	agouti-related peptide
GLP1R	glucagon-like peptide-1 receptor
CPE	carboxypeptidase E
NPY1R **and** *NPY5R*	neuropeptide Y receptor
HTR2A	serotonin receptor 2A
SOCS3	suppressor of cytokine signaling 3
STAT3	signal transducer and activator of transcription 3

Table 1

Descriptive Statistics for the 1982 Unrelated European Americans.

	Male	Female	Total
Number	858	1124	1982
Subjects with diabetes, n (%)	30 (3.50%)	40 (3.56%)	70 (3.53%)
Age (years)	45.55 ± 10.13	$48.23 \pm 10.59^*$	47.07 ± 10.48
BMI (kg/m^2)	29.77 ± 6.38	29.24 ± 8.44	29.47 ± 7.62

*
P < 0.0001 (t-test)

Table 2

The Adjusted Association of BMI with Gender, Age, and Diabetes Status in the 1982 European Americans, by BhGLM.

Parameter (standardized)	Estimate	Standard deviation	P value
Gender	0.0622	0.0208	0.0028
Age	0.3130	0.1295	0.0156
Age^2	-0.2868	0.1294	0.0266
Diabetes	0.4457	0.0590	<0.0001

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 3

SNPs Significantly Associated With BMI in the 1982 European Americans, by a BhGLM.

Gene name	SNP name	Variation	Location	Parameter estimate	P value
AGRP	rs5030980	A>G	Thr67Ala	0.0696	0.0233
CPE	rs1946816	A>C	intron	-0.0700	0.0046
CPE	rs4481204	A>G	intron	0.0479	0.0393
GHRL	rs35683	A>C	intron	0.0358	0.0107
GLP1R	rs2268641	A>G	intron	0.0339	0.0363
HTR2A	rs912127	A>G	intron	-0.0425	0.0135
NPY5R	Y5R1c52	A>G	UTR	-0.0812	0.0056
SOCS3	rs4969170	A>G	UTR	0.0308	0.0257
STAT3	rs4796793	C>G	locus	-0.0676	0.0053

Table 4

Genes Significantly Associated With BMI in the 1982 European Americans, by a BhGLM.

Gene name	Numbers of SNPs in the gene	Parameter estimate	P Value
AGRP	2	0.0595	0.0357
CPE	22	0.0229	0.0004
GHRL	7	0.0130	0.0144
NPY1R	6	0.0133	0.0313

Table 5

Gender-by-SNP Interaction Associated with BMI in the 1982 European Americans, by a GhGLM.

Gene name	SNP name	Interaction	Parameter estimate	P value
HTR2A	rs1745837	X Gender	0.0571	0.0072