BMC Genetics

BioMed Central

Proceedings

Open Access

Structural equation model-based genome scan for the metabolic syndrome

Catherine M Stein*[1], Yeunjoo Song[1], Robert C Elston[1], Gyungah Jun[1], Hemant K Tiwari[2] and Sudha K Iyengar[1]

Address: [1]Department of Epidemiology and Biostatistics, Case Western Reserve University 2500 MetroHealth Dr., R258 Rammelkamp Building, Cleveland, Ohio USA and [2]Department of Biostatistics, 327C Ryals Public Health Building, 1665 University Boulevard, University of Alabama at Birmingham, Birmingham, Alabama USA

Email: Catherine M Stein* - kasia@darwin.cwru.edu; Yeunjoo Song - song@darwin.cwru.edu; Robert C Elston - rce@darwin.cwru.edu; Gyungah Jun - gyungah@darwin.cwru.edu; Hemant K Tiwari - HTiwari@ms.soph.uab.edu; Sudha K Iyengar - ski@po.cwru.edu

* Corresponding author

from Genetic Analysis Workshop 13: Analysis of Longitudinal Family Data for Complex Diseases and Related Risk Factors
New Orleans Marriott Hotel, New Orleans, LA, USA, November 11–14, 2002

Published: 31 December 2003

BMC Genetics 2003, **4**(Suppl 1):S99

This article is available from: http://www.biomedcentral.com/1471-2156/4/s1/S99

Abstract

Background: The metabolic syndrome is characterized by the clustering of several traits, including obesity, hypertension, decreased levels of HDL cholesterol, and increased levels of glucose and triglycerides. Because these traits cluster, there are likely common genetic factors involved.

Results: We used a multivariate structural equation model (SEM) approach to scan the genome for loci involved in the metabolic syndrome. We found moderate evidence for linkage on chromosomes 2, 3, 11, 13, and 15, and these loci appear to have different relative effects on the component traits of the metabolic syndrome.

Conclusion: Our results suggest that the metabolic syndrome components, diabetes, obesity, and hypertension, are under the pleiotropic control of several loci.

Background

Metabolic syndrome (MSX) is characterized by the aggregation of several risk factors, including obesity, impaired glucose tolerance, elevated triglyceride levels and blood pressure, and low HDL cholesterol [1]. These factors promote the development of insulin resistance, diabetes, renal insufficiency, and cardiovascular disease. The prevalence of MSX in the United States has been estimated to be 18–24% [2,3]. Studies of the components of MSX demonstrate a substantial contribution of both genetic and environmental factors to disease risk [4].

Methods that account for shared environmental influence on the components of MSX are necessary if the genetic variance is not to be overestimated. Structural equation models (SEM) comprise a valuable method for partitioning the variance into its genetic and shared environmental components. A previous SEM analysis has suggested that the components of MSX are pleiotropically influenced by common genetic factors [5]. Thus, a multivariate analysis of the traits involved may increase power to detect quantitative trait loci (QTL) related to MSX. Although data reductive techniques such as principal component analysis have been performed (e.g. [6,7]), few comprehensive multivariate analyses have been done. In this paper, we analyzed the data from the Framingham offspring study using SEM, which enabled us to model traits related to

MSX simultaneously with genetic data and perform a genome scan.

Methods

Data

We analyzed data from the offspring cohort of the Framingham Heart Study, provided for Problem 1 of Genetic Analysis Workshop 13. The analysis was performed on all individuals with complete genotypic and phenotypic data for the fifth time point; this sample consisted of 1097 individuals from 381 pedigrees for a total of 1220 sib-pairs. The proportion of males and females was roughly equal, and the mean age was 51.33 (± 10.01) years. We modeled MSX using the measurements taken at the fifth time point for systolic blood pressure (SBP), fasting plasma glucose (GLUC), triglycerides (TG), HDL cholesterol (HDL), and body mass index (BMI), which was calculated as weight(kg)/(height(m))^2. Data from the final time point were used because we hypothesized that the study subjects would best demonstrate any progression to MSX by this time. Both GLUC and TG were log-transformed to reduce skewness. Because the distribution of GLUC was leptokurtotic after log transformation, we further transformed this variable using a generalized modulus power transformation [8]. The shape and scale parameters were estimated using ASSOC [9]. These five phenotypes were covariate-adjusted for age, sex, number of cigarettes smoked per day, number of drinks of alcohol per day, and two-way interaction terms significant at the $\alpha = 0.10$ level using a stepwise regression model; there were no significant three-way interactions. Standardized residuals were obtained from the regression model and used in the subsequent analyses. Both individual and sibling correlations were estimated using FCOR [9].

Structural equation model (SEM)

Five adjusted phenotypes were used in multivariate linkage analyses that simultaneously incorporate the phenotypic and genetic marker information into a single SEM, as proposed by Eaves and colleagues [10] and implemented in the Mx package [11]. This approach utilizes the full cross-trait covariance structure between siblings to better separate genetic from within-family environmental effects, which offered several advantages. First, the incorporation of multiple traits in a linkage analysis offers a considerable increase in power, particularly when there is shared environmental variation, even for traits with low heritability [12]. It has been shown that continuous traits have considerably more power for linkage analysis than do binary traits [13]. Finally, this approach is particularly applicable for the measurement of putative "endophenotypes" and is capable of detecting multiple QTL with pleiotropic effects [10].

The methodology is fully described in Eaves et al. [10]. Briefly, using their notation, the sibpair covariance matrix is expected to be

$$\begin{bmatrix} AA' + EE' & 0.5AA' \\ 0.5AA' & AA' + EE' \end{bmatrix},$$

where AA' and EE' are the covariance matrices due to additive genetic and within-family environmental effects, respectively. The information contained about the QTL contained in cross-trait covariances is used to estimate the effects of a putative QTL. The resulting sib-pair covariance matrix is

$$k_{(s)} \begin{bmatrix} AA' + EE' + QQ' & 0.5AA' + f_k QQ' \\ 0.5AA' + f_k QQ' & AA' + EE' + QQ' \end{bmatrix},$$

where k is the number of alleles shared identical by descent (IBD) at the location s, f_k is the expected proportion of pairs in a random sample sharing k alleles IBD (1/4, 1/2, 1/4), and AA' and EE' are residual additive genetic and environmental effects, after accounting for QQ', the contribution of the QTL effects at the genetic location in question. The SEM including QTL effects is illustrated in Figure 1. The proportion of alleles shared IBD at a genetic locus determines the correlation between the siblings' QTL effects. Thus, two models are evaluated: one with the SEM alone, and another weighted by the IBD probabilities for each sibpair for each genetic locus. Twice the difference between the log likelihood values for these two models gives a likelihood ratio statistic. The contributions of the putative genetic loci to each phenotype are represented as orthogonal factors. If these values (provided in Table 2) are squared, they represent the heritability of the phenotype due to that QTL. We estimated multi-point IBD probabilities using GENIBD [9] at 2cM intervals along the genome, using the parental genotypes from the original Framingham cohort. The linkage analysis was performed using Mx (script available on the Mx home page [14]).

Results

The individual- and sibling-specific correlations of the adjusted variables are provided in Table 1. This correlation matrix demonstrates that MSX phenotypes cluster within individuals, and that the five phenotypes are correlated between siblings. For each linked region from the SEM-based genome scan, the most significant *p*-values and associated QTL contributions for each phenotype are provided (Table 2). Though no regions attained statistical significance by conventional criteria [15], loci on chromosomes 2, 3, 11, 13, and 15 demonstrated tentative evidence for linkage at markers D2S1353, AFM306yg5, D11S2008, D13S793, D15S165, and D15S642, respectively. Weaker but potential evidence for linkage was seen on chromosomes 4 and 17 at markers D4S1647 and AFM290vc9, respectively. The QTL contributions to each

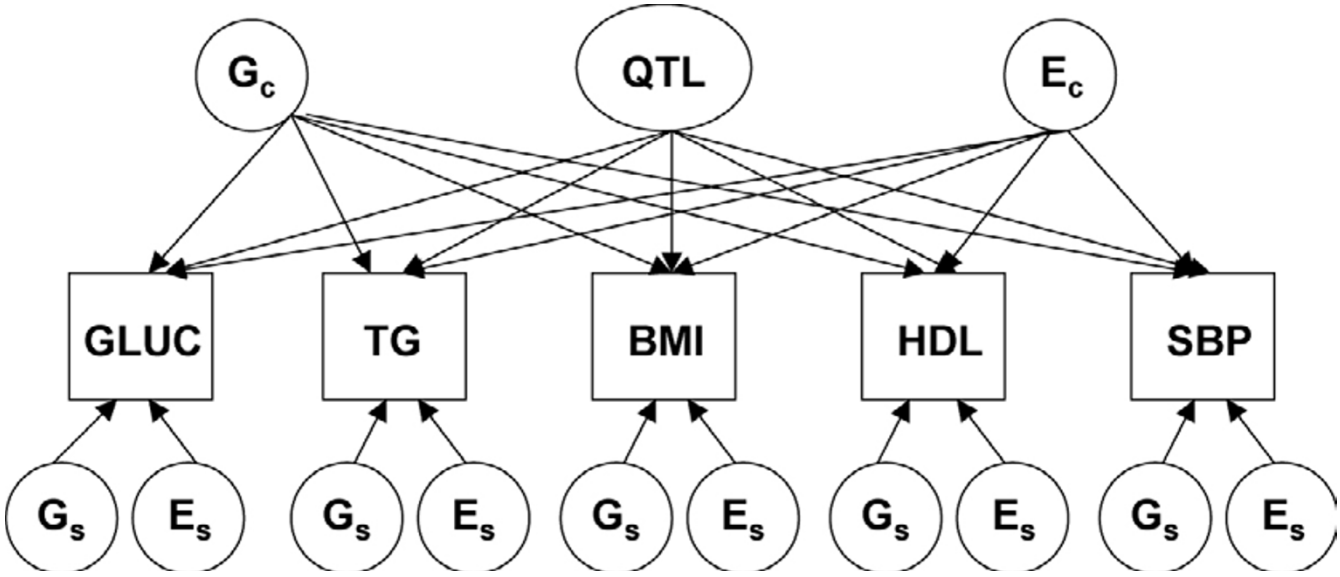

Figure 1
Path diagram of structural equation model. G_c and E_c are the genetic and environmental influences common to all the components of MSX, respectively, while G_s and E_s are the genetic and environmental influences specific to each component.

Table 1: Individual- and sibling-specific correlations.

	GLUC	TG	BMI	HDL	SBP
GLUC	**0.2306**	0.2118	0.3187	-0.1784	0.2491
TG	**0.0670**	**0.2247**	0.2917	-0.5326	0.3316
BMI	**0.1130**	**0.0973**	**0.2921**	-0.2883	0.2206
HDL	**-0.0276**	**-0.1911**	**-0.0624**	**0.2622**	-0.1024
SBP	**0.1143**	**0.0786**	**0.1027**	**-0.0620**	**0.1825**

Unbolded values are individual-specific correlations; bolded values are sibling-specific correlations.

phenotype were estimated. Negative QTL contributions imply that the putative QTL acts to lower the trait value (for example, low HDL increases risk for MSX). For most linked regions, GLUC had the largest contribution, though BMI and SBP were substantially influenced by the QTL as well. Interestingly, TG occasionally had negative QTL contributions and HDL had positive contributions for some loci. Though counterintuitive at first glance, this is most likely an artifact of imprecise estimation. In general, TG and HDL did not yield as strong evidence for linkage as did GLUC, BMI, and SBP.

Discussion

Previous multivariate analyses of MSX phenotypes have found that insulin and adiposity are highly correlated, and suggested that common genetic factors influence them [5,6]. Utilizing genome-scan data from the Framingham offspring data in conjunction with SEM, we have identified possible loci on several chromosomes that have pleiotropic effects on the component traits of MSX. SEM indicated that the greatest impact of the QTLs was on GLUC, although BMI and SBP were also influenced by these loci, supporting the hypothesis that obesity, glucose intolerance, and hypertension are key factors in MSX [5-7]. Previous factor analyses for MSX are inconsistent with these observations, as glucose and obesity variables frequently loaded together, but there was seldom overlap with blood pressure factors [16].

We observed moderate linkage on chromosomes 2, 3, 4, 11, 13, and 15, although the relative effect of these putative QTL on each MSX phenotype appears to differ.

Table 2: Linked regions for MS identified using SEM.[A]

			QTL contribution				
Region	Location (cM)	*p*-Value	GLUC	TG	BMI	HDL	SBP
2q24.2	165	0.003	0.4633	0.1344	0.3305	-0.1086	0.3044
3p21.31-3p13	71–112	0.003	0.4537	0.2567	0.4101	-0.0987	0.2894
4q24.2	88	0.024	0.2258	-0.0360	0.3348	-0.0823	0.4879
11p15.5-11q23.3	9–113	0.003	0.3352	0.2877	0.3639	-0.0588	0.3417
12q14.3-12q15	78	0.029	0.2242	0.0290	0.3839	0.0366	0.4476
13q32	76	0.009	0.3393	0.1286	0.2361	0.0127	0.3627
15q11.2-15q31.1	5–20	0.005	0.3221	-0.0547	0.0433	-0.2067	0.2927
15q26.3	122	0.011	0.4494	0.1068	0.2088	-0.0454	0.3349
17p12	31–32	0.030	0.2102	0.1363	0.2268	-0.0275	0.4786

[A]Marker names are provided within the text.

Because the literature related to linkage signals for MSX components is vast, we have summarized reports within a 30cM window of our linkage signals [17]. Though we conducted a truly multivariate analysis of MSX, our results can be considered confirmations of these previous reports. Body size [7], diabetes mellitus [18], and insulin resistance [19] have been linked to chromosome 11; this region also contains several candidate genes, including insulin (OMIM 17673), *SHIP2* (OMIM 600829), and the uncoupling protein 2 (OMIM 601693) [20]. An adiposity-insulin factor attained suggestive linkage to chromosome 2 [6], and a separate study found linkage to BMI 20 cM away [21]. Insulin resistance [19], BMI [19], and high blood pressure [22] have been linked to chromosome 3. Leptin levels have been linked to the 20-cM location on chromosome 15 [21]. Finally, *IRS-2*, a diabetes candidate gene, has been mapped to chromosome 13 (102 cM) [23].

Given the multifactorial nature of MSX, the method of linkage analysis most suitable is one that incorporates all available trait information. Here we use SEM with multivariate data to increase power, account for shared sibling environment within a sib-pair linkage analysis, and depict the combined effects of the components of MSX simultaneously. This approach is uniquely capable of estimating the contribution of each phenotype to the QTL, which can illustrate facets of MSX biology. Another advantage of our approach is the direct use of quantitative traits. Currently, there are two different clinical definitions of MSX [2], which require the presence of at least two (or three) of five characteristics. By analyzing the full quantitative scale for each MSX component, not only do we gain power statistically, but we also avoid classifying individuals together who have moderate to extreme trait values. As observed in Table 1, these traits are not highly correlated with each other, but our genome scan results suggest that they may be influenced by common genetic factors.

There are, however, limitations to these analyses. Although there were data available on blood pressure medication, there were no data on diabetes medication. Since both were not available, we did not adjust our quantitative variables for these covariates. Because of missing genotype and/or phenotype data, approximately 700 sibpairs were excluded from this analysis. Though our sample size was still substantial, it is unknown how these sibpairs might have increased power. There is an upward bias in the estimation of the QTL contributions. The sum of the squared QTL contributions to GLUC for all distinct linked regions is greater than 1, which is not possible if these values are interpreted to be QTL-specific heritabilities. Also, when partitioning genetic and shared-environmental components of variance, gene×environment interaction was unaccounted for, and allowing for such interaction would only exacerbate the discrepancy. Finally, in the SEM, all sibpairs were treated as though they were independent. It is unclear how to handle larger sibships within the SEM; however, the average sibship size is less than three, so this should not have a great impact on the analysis.

Another point worth revisiting is our approach to nonnormal data. The SEM assumes multivariate normality and, with increasing sample size, departure from this assumption will affect the likelihood ratio test adversely. In our analysis, the GLUC variable was transformed using a generalized modulus power transformation [8] to remove skewness and kurtosis. Though the departure of this variable from normality may be regarded as biologically based [24], a leptokurtotic distribution leads to an inflated type I error rate [24], and the scale of measurement may drastically affect results of linkage analysis [25]. In fact, when this analysis was performed without transformation, most of the linkage results were more significant, and the QTL contributions were incorrectly inflated

(data not shown). After transformation, the relative effects on the other MSX traits increased, indicating that the variance was partitioned more evenly when univariate normality of the five traits was attained.

Conclusions

A number of loci appear to be linked to MSX, most notably regions on chromosomes 2, 3, 11, 13, and 15. Diabetes, obesity, and hypertension were most influenced by these loci. These regions appear to influence the components of MSX in different ways, and they warrant further analysis, both in reference to the metabolic syndrome itself and to other disorders associated with insulin-resistance.

Acknowledgments

This study was supported in part by U.S. Public Health Service Grants HL07567 from the National Heart, Lung, and Blood Institute, HG01577 from the National Human Genome Research Institute, RR03655 from the National Center for Research Resources, GM28356 from the National Institute of General Medicine Sciences, DK54644 and DK57292 from the National Institute of Diabetes and Digestive and Kidney Diseases, and NIH grants RR08123 and MH01458. Some of the results of this paper were obtained by using the program package S.A.G.E., which is supported by a U.S. Public Health Resource Grant (1 P41 RR 03655). CMS sincerely thanks Dr. Michael Neale for all his help using the Mx software and Courtney Gray-McGuire for helpful discussion.

References

1. Expert Panel on the Detection Evaluation and Treatment of High Blood Cholesterol in Adults: **Executive Summary of the Third Report of the National Cholesterol Education Treatment Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III).** *JAMA* 2001, **285:**2486-2497.
2. Alexander CM, Landsman PB, Teutsch SM, Haffner SM: **Prevalence of metabolic syndrome and coronary heart disease using NCEP and WHO criteria [abstract].** *Diabetes* 2002, **51(suppl 2):**A218.
3. Ford ES, Giles WH, Dietz WH: **Prevalence of the metabolic syndrome among US adults: findings from the third national health and nutrition examination survey.** *JAMA* 2002, **287:**356-359.
4. Groop L: **Genetics of the metabolic syndrome.** *Br J Nutr* 2000, **83(suppl):**S39-S48.
5. Hong Y, Pedersen NL, Brismar K, de Faire U: **Genetic and environmental architecture of the features of the insulin-resistance syndrome.** *Am J Hum Genet* 1997, **60:**143-152.
6. Arya R, Blangero J, Williams K, Almasy L, Dyer TD, Leach RJ, O'Connell P, Stern MP, Duggirala R: **Factors of insulin-resistance syndrome-related phenotypes are linked to genetic locations on chromosomes 6 and 7 in nondiabetic Mexican-Americans.** *Diabetes* 2002, **51:**841-847.
7. Imperatore G, Knowler WC, Kobes S, Bennett PH, Hanson RL: **Genome-wide linkage analysis of factors characterizing the metabolic syndrome among Pima Indians [abstract].** *Diabetes* 1999, **48(suppl):**A182.
8. George VT, Elston RC: **Generalized modulus power transformations.** *Commun Statist Theory Meth* 1988, **17:**2933-2952.
9. SA.G.E., version 4.2.: **Statistical Analysis for Genetic Epidemiology.** *A computer package available from Statistical Solutions, Cork, Ireland* 2002.
10. Eaves LJ, Neale MC, Maes H: **Multivariate multipoint linkage analysis of quantitative trait loci.** *Behav Genet* 1996, **26:**519-525.
11. Neale MC: **Mx: Statistical Modeling, 5th edition.** *Medical College of Virginia, Richmond, VA* 1999.
12. Schmitz S, Cherny SS, Fulker DW: **Increase in power through multivariate analyses.** *Behav Genet* 1998, **28:**357-363.
13. Duggirala R, Williams JT, Williams-Blangero S, Blangero J: **A variance components approach to dichotomous trait linkage analysis using a threshold model.** *Genet Epidemiol* 1997, **14:**987-992.
14. **Mx software homepage** [http://views.vcu.edu/mx/]
15. Lander E, Kruglyak L: **Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results.** *Nat Genet* 1995, **11:**241-247.
16. Meigs JB: **Invited commentary: Insulin resistance syndrome? Syndrome X? Multiple metabolic syndrome? A syndrome at all? Factor analysis reveals patterns in the fabric of correlated metabolic risk factors.** *Am J Epidemiol* 2000, **152:**908-911.
17. Cordell HJ: **Sample size requirements to control for stochastic variation in magnitude and location of allele-sharing linkage statistics in affected sibling pairs.** *Ann Hum Genet* 2001, **65:**491-502.
18. Ghosh S, Watanabe RM, Valle TT, Hauser ER, Magnuson VL, Langefeld CD, Ally DS, Mohlke KL, Silander K, Kohtamaki K, Chines P, Balow J Jr, Birznieks G, Chang J, Eldridge W, Erdos MR, Karanjawala ZE, Knapp JI, Kudelko K, Martin C, Morales-Mena A, Musick A, Musick T, Pfahl C, Porter R, Rayman JB: **The Finland-United States investigation of non-insulin-dependent diabetes mellitus genetics (FUSION) study. I. An autosomal genome scan for genes that predispose to type 2 diabetes.** *Am J Hum Genet* 2000, **67:**1174-1185.
19. Watanabe RM, Ghosh S, Langefeld CD, Valle TT, Hauser ER, Magnuson VL, Mohlke KL, Silander K, Ally DS, Chines P, Blaschak-Harvan J, Douglas JA, Duren WL, Epstein MP, Fingerlin TE, Kaleta HS, Lange EM, Li C, McEachin RC, Stringham HM, Trager E, White PP, Balow J Jr, Birznieks G, Chang J, Eldridge W: **The Finland-United States investigation of non-insulin-dependent diabetes mellitus genetics (FUSION) study. II. An autosomal genome scan for diabetes-related quantitative trait loci.** *Am J Hum Genet* 2000, **67:**1186-1200.
20. **Online Mendelian Inheritance in Man** [http://www3.ncbi.nlm.nih.gov/omim/]
21. Cheng LS-C, Davis RC, Raffel LJ, Xiang AH, Wang N, Quiñones M, Wen P-Z, Toscano E, Diaz J, Pressman S, Henderson PC, Azen SP, Hsueh WA, Buchanan TA, Rotter JI: **Coincident linkage of fasting plasma insulin and blood pressure to chromosome 7q in hypertensive Hispanic families.** *Circulation* 2001, **104:**1255-60.
22. Francke S, Manraj M, Lacquemant C, Leoceur C, Leprêtre F, Passa P, Hebe A, Corset L, Yan SLK, Lahmidi S, Jankee S, Gunness TK, Ramjuttun US, Balgobin V, Dina C, Froguel P: **A genome-wide scan for coronary heart disease suggests in Indo-Mauritians a susceptibility locus on chromosome 16p13 and replicates linkage with the metabolic syndrome on 3q27.** *Hum Mol Genet* 2001, **10:**2751-65.
23. Kalidas K, Wasson J, Glaser B, Meyer JM, Duprat LJ, White MF, Permutt MA: **Mapping of the human insulin receptor substrate-2 gene, identification of a linked polymorphic marker and linkage analysis in families with Type II diabetes: no evidence for a major susceptibility role.** *Diabetologia* 1998, **41:**1389-91.
24. Allison DB, Neale MC, Zannolli R, Schork NJ, Amos CI, Blangero J: **Testing the robustness of the likelihood-ratio test in a variance-component quantitative-trait loci-mapping procedure.** *Am J Hum Genet* 1999, **65:**531-544.
25. Wilson AF, Elston RC, Tran LD, Siervogel RM: **Use of the robust sib-pair method to screen for single-locus, multiple-locus, and pleiotropic effects: application to traits related to hypertension.** *Am J Hum Genet* 1991, **48:**862-872.