Original Article
EPIDEMIOLOGY/GENETICS

Obesity

Modeling Metabolic Syndrome Through Structural Equations of Metabolic Traits, Comorbid Diseases, and GWAS Variants

Rebekah Karns[1], Paul Succop[1], Ge Zhang[2], Guangyun Sun[1], Subba R. Indugula[1], Dubravka Havas-Augustin[3], Natalija Novokmet[3], Zijad Durakovic[3], Sanja Music Milanovic[4], Sasa Missoni[3], Silvije Vuletic[4], Ranajit Chakraborty[5], Pavao Rudan[3] and Ranjan Deka[1]

Objective: To provide a quantitative map of relationships between metabolic traits, genome-wide association studies (GWAS) variants, metabolic syndrome (MetS), and metabolic diseases through factor analysis and structural equation modeling (SEM).

Design and Methods: Cross-sectional data were collected on 1,300 individuals from an eastern Adriatic Croatian island, including 14 anthropometric and biochemical traits, and diagnoses of type 2 diabetes, coronary heart disease, gout, kidney disease, and stroke. MetS was defined based on Adult Treatment Panel III criteria. Forty widely replicated GWAS variants were genotyped. Correlated quantitative traits were reduced through factor analysis; relationships between factors, genetic variants, MetS, and metabolic diseases were determined through SEM.

Results: MetS was associated with obesity ($P < 0.0001$), dyslipidemia ($P < 0.0001$), glycated hemoglobin (HbA1c; $P = 0.0013$), hypertension ($P < 0.0001$), and hyperuricemia ($P < 0.0001$). Of metabolic diseases, MetS was associated with gout ($P = 0.024$), coronary heart disease was associated with HbA1c ($P < 0.0001$), and type 2 diabetes was associated with HbA1c ($P < 0.0001$) and obesity ($P = 0.008$). Eleven GWAS variants predicted metabolic variables, MetS, and metabolic diseases. Notably, rs7100623 in HHEX/IDE was associated with HbA1c ($\beta = 0.03$; $P < 0.0001$) and type 2 diabetes ($\beta = 0.326$; $P = 0.0002$), underscoring substantial impact on glucose control.

Conclusions: Although MetS was associated with obesity, dyslipidemia, glucose control, hypertension, and hyperuricemia, limited ability of MetS to indicate metabolic disease risk is suggested.

Obesity (2013) **00**, 0000-0000. doi:10.1002/oby.20445

Introduction

Over one-third of adult Americans meet the diagnostic criteria of metabolic syndrome [MetS; (1)], a clustering of metabolic disturbances that increase risk for type 2 diabetes, cardiovascular diseases, and even cancer (2-8). The worldwide prevalence of MetS is growing, most notably in developing nations where it follows obesity trends, sedentary lifestyles, and excessive caloric intake (9-16). Although formal definitions of MetS from the World Health Organization (17), National Cholesterol Education Program Third Adult Treatment Panel III [NCEP ATPIII (18)], American Heart Association/National Heart, Lung and Blood Institute (19), and International Diabetes Federation (20) differ somewhat, the most relevant elements of MetS diagnosis are well established. Irrespective of diagnostic cut points, all definitions include four key traits: obesity, glucose intolerance, dyslipidemia, and hypertension. Previous multivariate analyses of MetS have provided models of their associations and confirmed the involvement of these traits (21-25), which may be measured by numerous highly correlated measures (e.g., obesity by waist–hip ratio, waist circumference, hip circumference, and body mass index). As evidenced by the number of discordant MetS definitions and recent publications questioning the utility of the diagnosis *per se* (26-28), the development, progression, and subsequent disease risks of MetS remain in question. Investigation of a

[1] Division of Epidemiology and Biostatistics, Department of Environmental Health, University of Cincinnati, Cincinnati, Ohio, USA. Correspondence: Ranjan Deka (ranjan.deka@uc.edu) [2] Human Genetics Division, Cincinnati Children's Hospital, Cincinnati, Ohio, USA [3] Institute for Anthropological Research, Zagreb, Croatia [4] Andrija Stampar School of Public Health, University of Zagreb, School of Medicine, Zagreb, Croatia [5] Center for Computational Genomics, University of North Texas Health Science Center, Forth Worth, Texas, USA

Funding agencies: The study was supported by grants from the National Institutes of Health, USA (R01 DK069845, R01 DK069845-4S, and P30 ES006096) and from the Croatian Ministry of Science, Education and Sports (196-1962766-2751, 196-1962766-2747, and 196-0342282-0291). R.K. was supported by a training grant fellowship from the National Institutes of Environmental Health Sciences, USA (T32 ES010957).

Disclosure: The authors declared no conflict of interest.

Received: 24 February 2012 **Accepted:** 19 February 2013 **Published online 20 March 2013. doi:10.1002/oby.20445**

comprehensive set of metabolic traits may improve our understanding of the MetS network and its long-term implications.

In addition to lifestyle and environmental influences, MetS has a strong genetic component with heritability estimated between 6 and 50% (8,29). Although genome-wide association studies (GWAS) have identified numerous single-nucleotide polymorphisms (SNPs) associated with individual metabolic traits and diseases, the impact of these variants on the MetS network and its downstream metabolic diseases is not well investigated. Furthermore, traditional single-trait genetic association testing may be inadequate to determine a variant's impact in a complex disease system. We aimed to improve our understanding of the MetS network through multivariate analyses of an expanded set of metabolic traits and associated GWAS-derived genetic variants involved in the development of MetS and metabolic diseases in an isolated population with homogenous genetic background and environmental exposures, and a high prevalence of obesity, hypertension, and MetS (15,16,30,31). This objective was accomplished through two steps: 1) determining latent (unmeasured) factors underlying correlated anthropometric and biochemical traits using exploratory and confirmatory factor analyses (EFA and CFA, respectively) and 2) describing the impact of manifest (measured) traits, latent factors, and GWAS-derived variants on MetS and metabolic disease risk through structural equation modeling (SEM). This study provides a quantitative map of the impact and relationships between an expansive set of metabolic parameters, GWAS variants, MetS, and downstream metabolic diseases.

Methods and Procedures

Population and phenotypic measurements

Our study population derives from a genetic epidemiology study in an isolated island population in the eastern Adriatic coast of Croatia, with obesity, hypertension, and MetS rates similar to US populations (15,16,30,31). Croatian islanders are predominantly of Slavic descent, having last emigrated to the Adriatic islands between the 15th and 18th centuries (32). Since that time, the island populations have remained relatively isolated because of geographic confinement, resulting in a population with a homogenous genetic background and similar dietary and environmental exposures.

This study was conducted in the middle Dalmatian island of Hvar, the largest among the Adriatic islands, with a population of 11,103 individuals according to the 2001 census. Anthropometric measurements, blood pressure, and blood samples were obtained from 1,300 individuals aged 18-80 years, with no consideration for disease status or medication use. The study sample was obtained through nonprobability sampling methods; village leaders advertised the epidemiological study and volunteers meeting age requirements were accepted. Height (Ht), weight (Wt), waist circumference (WC), and hip circumference (HC) were measured, from which BMI (Wt in kg)/(Ht in m^2) and waist–hip ratio (WC/HC) were computed. The second and third of three blood pressure measurements (using mercury sphygmomanometer at a resting position) were used to calculate mean systolic and diastolic blood pressures (SBP and DBP, respectively). Blood samples were drawn following a 12-h fast; serum was separated and kept frozen until shipped to the clinical chemistry laboratory in Zagreb, Croatia. Standard biochemical tests were performed to measure fasting plasma glucose (FPG; enzymatic hexokinase assay CHOD-PAP), glycated hemoglobin (HbA1c; immunoinhibition assay), high-density lipoprotein (HDL; homogeneous enzyme inhibition assay), low-density lipoprotein (LDL; FRIEDWALD calculation), total cholesterol (TC; photometric color test CHOD-PAP), triglycerides (TG; photometric color test GPO-PAP), and uric acid (UA; enzymatic color test) levels. Through a self-administered questionnaire, diagnoses of stroke, type 2 diabetes, gout, coronary heart disease, and kidney disease were obtained, and were confirmed through assessment of medications prescribed, chart review, and clinical diagnostic data. Study protocols were approved by the Institutional Review Board of the University of Cincinnati and the ethics committee of the Institute for Anthropological Research, Zagreb. Informed written consent was obtained from the participants.

Data cleaning and preparation

The distribution of 14 biochemical and anthropometric traits was assessed through visual examinations of histograms, Q–Q plots, and through formal Kolmogorov–Smirnov tests using SAS v9.2. Conformity with normal distribution was achieved through log-transformation (HDL and TG) and the Box-Cox method (FPG and HbA1c, with transformation power parameters of -2.5 and -3, respectively; Ref. 33). UA was recoded as a five-level categorical variable to resolve overdispersion. Based on the sample distribution, UA quintile cut-off points were the following: 1) $\leq$240 μmol/l; 2) 241-280 μmol/l; 3) 281-315 μmol/l; 4) 316-363 μmol/l; and 5) $\geq$364 μmol/l. Missing values (0.7% of the dataset, missing completely at random; Ref. 34) were imputed using Gibbs sampling methods and Bayesian linear regression in the R package MICE (v2.10.1). All variables were adjusted for age prior to the application of MetS diagnostic criteria. MetS diagnosis was based on the criteria of NCEP ATPIII (18), which defines MetS as the co-occurrence of three or more of the following five risk factors: 1) WC $>$ 102 cm in men and $>$88 cm in women; 2) TG $\geq$ 1.69 mmol/l; 3) HDL $<$ 1.03 mmol/l in men and $<$1.29 mmol/l in women; 4) blood pressure $\geq$130/85 mm Hg; and 5) FPG $\geq$6.1 mmol/l. Although the diagnostic criteria of NCEP ATPIII are somewhat less inclusive than other definitions, they classify a population with more severe MetS, thus aiming at capturing only individuals with high risk for metabolic disease rather than those who are relatively healthy or at low risk.

Genome-wide SNP genotyping was performed using the Affymetrix Human SNP Array 5.0 at the University of Cincinnati's Center for Genome Information Core Genotyping Laboratory according to the manufacturer's protocol. Among chips that passed the DM QC call rate (>0.86), genotype calls were determined using the CRLMM algorithm. A cleaned data set of 344,512 SNPs was obtained following quality control filtering (MAF $>$ 0.02, HWE $P >$ 0.0001, call rate $>$ 95%). Genotype imputation was performed on the cleaned data set using MACH and the reference haplotype from Phase II CEU HapMap, yielding 2.2 million SNPs. The NCBI GWAS catalog (35) was queried for SNPs with widely replicated associations reported with the 14 anthropometric and biochemical measures included in our study; highly significant and replicated SNPs were extracted from our GWAS data set. Genotype and allele frequencies and linkage disequilibrium patterns were determined through Haploview. To confirm association in our study population, we performed a replication study of 40 GWAS-derived SNPs (35) with age- and gender-adjusted phenotypic traits based on an additive model (1 df test). SNPs with nominal significant association ($P \leq 0.05$) with any phenotypic trait were included as exogenous predictors in a final

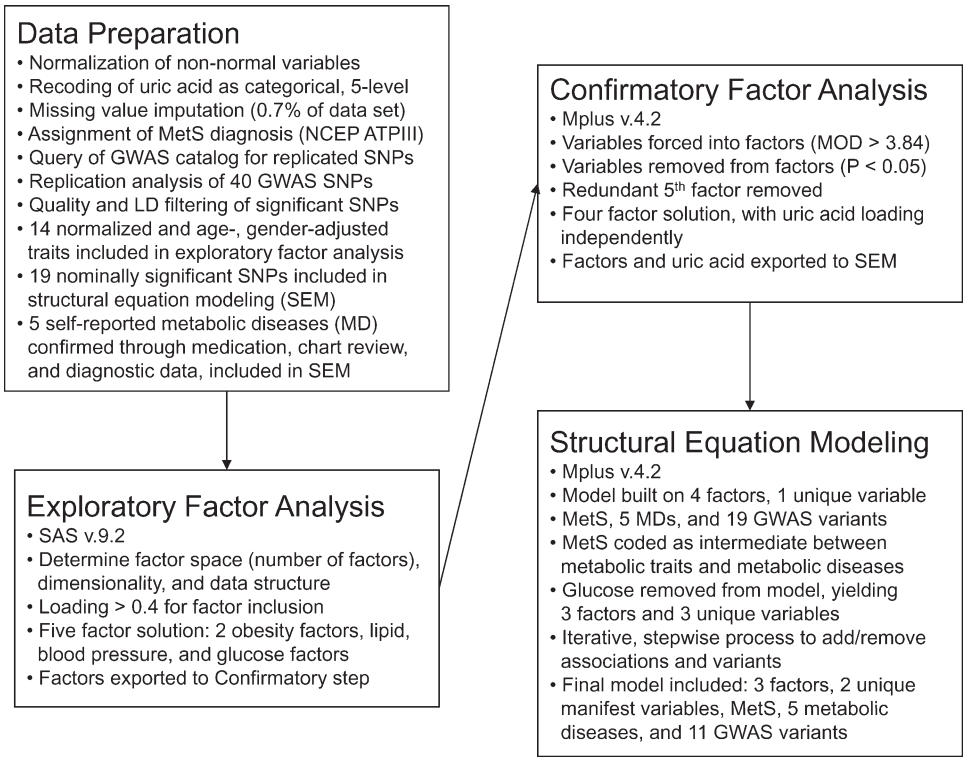

FIGURE 1 Flow chart of analysis process.

SEM. In regions of moderate to high linkage disequilibrium (LD; $r > 0.6$), only the SNP with the most significant association signals within the LD block was included in SEM.

Statistical methods

A diagram of the analytic process is presented in Figure 1. As the data structure was unknown, we performed both EFA and CFA to determine the factor structure that best fit the data. In step 1, EFA was performed in SAS v.9.2 to provide dimensionality and structure of latent factors underlying redundant and collinear manifest variables. Creation of latent factors allows consideration and inclusion of all possible information provided by the correlated variables while eliminating collinearity. With an oblique factor rotation to allow factor correlation (which also allows factor loadings >1.0), a minimum loading of 0.4 was required for inclusion in a factor. CFA based on EFA latent factors was performed in Mplus v.4.2. In a stepwise and iterative process, manifest variables were forced into latent factors according to model modification indices (MODs) $\geq$ 3.84 and removed if loading was insignificant ($P > 0.05$). To ensure similar factor loadings in males and females, EFA and CFA were initially performed separately for each gender; they were combined because of nearly identical factors and factor loadings. EFA and CFA model estimates were based on the maximum likelihood procedure and model fit was based on a χ^2 statistic, the Tucker and Lewis's reliability coefficient (TLI), and root-mean-square error of approximation (RMSEA), each of which provide information on specific aspects of model fit. However, we relied on TLI and RMSEA rather than χ^2 because factor creation may not explain all covariation when strong associations exist between factors. The CFA is unable to explain all variability and is simply a starting point for SEM.

To provide a multivariate regression model of relationships between GWAS variants, unique manifest variables (continuous, measured variables that do not load on any factor), latent factors (continuous), MetS (binary), and metabolic diseases (binary), a first-order SEM was built in Mplus v4.2 based on the factor structures obtained from EFA and CFA. We have not used an explanatory/fixed model (where variables, loadings, and estimates are fixed, based on previous models) because our study population is isolated and may exhibit phenotypic and genotypic characteristics distinct from previous populations included in SEM of MetS (21-25). SNPs that showed nominal significance in the replication association analysis with top signals in regions of moderate to high LD were included as predictors (assuming additive effects) of unique manifest variables, latent factors, MetS, and metabolic diseases. As MetS results from the dysregulation of multiple metabolic traits and the diagnosis is hypothesized to predict metabolic disease risk, we considered MetS as the potential link between individual metabolic traits and metabolic diseases. Therefore, MetS was fitted as the intermediate between unique manifest variables, latent factors, and metabolic diseases. SEM building proceeded in a multistep, multimodel, iterative process in which associations were added and removed in backwards elimination according to MOD and the model was tested until no further improvement could be found. Correlation of residual variances between variables was specified during model building, because

TABLE 1 Gender-specific descriptive statistics of anthropometric and biochemical measures, and prevalence (%) of metabolic diseases

	Female ($n = 737$)		Male ($n = 563$)		
Trait	**Mean**	**SD**	**Mean**	**SD**	***P*-value**
Age (years)	54.33	9.87	56.25	10.77	<0.0001
Ht (cm)	164.3	6.01	177.9	6.84	<0.0001
Wt (kg)	72.61	12.53	87.57	12.86	<0.0001
BMI (kg/m^2)	26.94	4.29	27.63	3.38	<0.0001
WC (cm)	90.69	11.31	99.29	9.15	<0.0001
HC (cm)	104.64	9.64	103.1	7.32	0.0015
WHR	0.86	0.07	0.96	0.06	<0.0001
FPG (mmol/l)	5.82	1.25	6.11	1.31	<0.0001
HbA1c (%)	5.78	0.82	5.71	0.83	N.S.
DBP (mmHg)	80.62	9.17	83.53	8.79	<0.0001
SBP (mmHg)	130.9	18.98	134.44	17.93	0.0006
HDL (mmol/l)	1.53	0.41	1.29	0.32	<0.0001
LDL (mmol/l)	3.8	0.99	3.57	0.99	<0.0001
TC (mmol/l)	5.95	1.14	5.66	1.15	<0.0001
TG (mmol/l)	1.38	0.78	1.7	1.33	<0.0001
UA (μmol/l)	266.83	72.87	359.1	78.16	<0.0001
Stroke (%)	2.51	0.18	4.75	0.89	<0.0001
Coronary heart disease (%)	9.26	1.06	8.96	1.2	N.S.
Gout (%)	2.25	0.54	8.61	1.18	<0.0001
Type 2 diabetes (%)	15.06	1.31	19.72	1.67	<0.0001
Kidney disease (%)	2.38	0.56	3.51	0.77	<0.0001
MetS (%)	32.01	1.72	35.15	2.01	N.S.

P-value indicates significant male–female differences.

TABLE 2 Factor loading in exploratory and confirmatory factor analysis

Exploratory factor analysis		Confirmatory factor analysis	
	Loading		**Loading**
Factor 1		*Factor 1*	
Wt	0.89	Wt	0.95
WC	0.68	WC	0.88
HC	0.94	HC	0.87
BMI	1.02	BMI	0.91
Factor 2		WHR	0.39
LDL	1.06	*Factor 2*	
TC	0.83	LDL	0.31
Factor 3		TC	0.41
DBP	0.75	TG	0.86
SBP	0.96	HDL	−0.51
Factor 4		*Factor 3*	
FPG	0.97	DBP	0.95
HbA1c	0.99	SBP	0.75
Factor 5		*Factor 4*	
WHR	1.0	FPG	1.93
WC	0.48	HbA1c	0.51

the traits included in the model are highly correlated even if not loaded in the same factors. SEM allows simultaneous inclusion of continuous and categorical outcome variables; the model therefore used multiple regression for statements with continuous outcomes, and logistic regression for statements with MetS and metabolic diseases as outcomes. As χ^2 statistics may be overly conservative in models with relatively large sample sizes and categorical outcomes, goodness of fit was based on comparative fit index (CFI), TLI, and RMSEA. CLI and TFI values of $\geq$0.95 indicate good fit and RMSEA values of $\leq$0.06 indicate close fit (36). Figure 1 provides a flow chart of steps taken in the statistical analysis.

Results

Table 1 provides descriptive statistics, including gender-specific mean values and standard deviations for all phenotypic traits, significance of male–female mean differences, and prevalence of MetS and five metabolic diseases. Of traits involved in NCEP ATPIII MetS diagnosis, mean values of FPG in males (6.11 mmol/l) were in excess of the hyperglycemia cut-off point, blood pressures bordered on hypertensive in both males and females (134.44/83.53 and 130.90/80.62 mmHg SBP/DBP, respectively), and WC was above the obese cut-off point in females (90.69 cm) and was borderline in males (99.29 cm). Males had significantly higher mean values for all anthropometric and biochemical measures except HbA1c and HDL.

EFA and CFA factor loadings are presented in Table 2. Five factors were extracted from EFA, which included 1) obesity factor underlying BMI, WC, Wt, and HC; 2) lipid factor underlying LDL and TG; 3) blood pressure factor manifested by SBP and DBP; 4) glucose factor underlying FPG and HbA1c; and 5) a second obesity factor manifested by WHR and WC, with UA, TC, and HDL uniquely loaded on independent factors. Through the confirmatory steps of the analysis, WHR was forced into the first obesity factor and HDL and TC were forced into the lipid factor. Therefore, four interpretable latent factors were identified from step 1 of the analysis (EFA and CFA), which included 13 of the original 14 measured variables, as uric acid loaded independently as a unique variable. The four factors were as follows: 1) obesity factor underlying BMI, WC, WHR, Wt, and HC; 2) lipid factor underlying LDL, TG, TC, and HDL; 3) blood pressure factor underlying SBP and DBP; and 4) glucose factor underlying FPG and HbA1c. Within the lipid factor, HDL was positively loaded, whereas TG, TC, and LDL loaded negatively, resulting in an inverse relationship with MetS. As expected, the final CFA fit was poor according to the χ^2 (99.05; $P < 0.0001$); however, the TLI and RMSEA (1.00 and 0.03, respectively) indicated factor adequacy and reliability.

During step 2 (SEM fitting of latent and unique manifest variables, GWAS variants, MetS, and metabolic diseases), the glucose factor was dismantled because neither HbA1c nor FPG loaded significantly. Following inclusion of HbA1c and FPG as unique manifest

TABLE 3 Summary of studied GWAS SNPs extracted from the GWAS catalog

NCBI SNP ID	Gene	Chromosome	Position	Phenotypes in GWAS catalog	Minor allele	Major allele	MAF	CEU MAF
rs646776[1]	CELSR2	1	109818530	LDL, TC	C	T	0.225	0.254
rs599839[1]	PSRC1	1	109822166	LDL	G	A	0.225	0.279
rs6754295	APOB	2	21206183	TG, HDL	G	T	0.226	0.212
rs693	APOB	2	21232195	LDL, TG, TC	A	G	0.451	0.491
rs515135[1]	APOB	2	21286057	LDL	T	C	0.276	0.225
rs780094[1]	GCKR	2	27741237	LDL, TG, TC	T	C	0.491	0.392
rs419076	MDS	3	169100886	SBP	T	C	0.498	0.418
rs874432[1]	SLC2A9	4	9920606	UA	A	T	0.314	0.224
rs16998073	FGF5	4	81184341	DBP	T	A	0.233	0.192
rs3846662[1]	HMGCR	5	74651084	LDL, TC	G	A	0.411	0.425
rs328	LPL	8	19819724	TG, HDL	G	C	0.113	0.125
rs10096633[1]	LPL	8	19830921	TG, HDL	T	C	0.155	0.142
rs2083637	LPL	8	19865175	HDL	G	A	0.238	0.274
rs11014166[1]	CACN2B	10	18708798	DBP, SBP	T	A	0.457	0.367
rs7100623[1]	HHEX/IDE	10	94327156	HbA1c	T	C	0.221	0.195
rs1004467	CYP17A1	10	104594507	SBP	G	A	0.156	0.084
rs11191548	NT5C2	10	104846178	SBP	C	T	0.144	0.059
rs12243326[1]	TCF7L2	10	114788815	HbA1c	C	T	0.282	0.243
rs4938303	RPL15P15	11	116584987	TG	C	T	0.329	0.217
rs12272004[1]	RPL15P15	11	116603724	LDL, TG, TC	A	C	0.088	0.036
rs3184504	SH2B3	12	111884608	DBP	C	T	0.475	0.445
rs4775041	LIPC	15	58674695	HDL	C	G	0.288	0.280
rs1532085	LIPC	15	58683366	HDL, TC	A	G	0.342	0.367
rs1800588[1]	LIPC	15	58723675	HDL	T	C	0.243	0.258
rs1378942[1]	CSK	15	75077367	DBP, SBP	C	A	0.399	0.330
rs6495122	CSK	15	75125645	DBP	C	A	0.454	0.351
rs1421085	FTO	16	53800954	Obesity	C	T	0.482	0.460
rs1121980[1]	FTO	16	53809247	Obesity	A	G	0.484	0.482
rs17817449	FTO	16	53813367	Obesity	G	T	0.451	0.460
rs8050136	FTO	16	53816275	Obesity	A	C	0.451	0.460
rs3751812	FTO	16	53818460	Obesity	T	G	0.451	0.460
rs9939609	FTO	16	53820527	Obesity	A	T	0.451	0.460
rs7190492[1]	FTO	16	53828752	Obesity	A	G	0.308	0.351
rs173539	CETP	16	56988044	HDL	T	C	0.341	0.308
rs1800775	CETP	16	56995236	TG, HDL	C	A	0.447	0.487
rs1532624[1]	CETP	16	57005479	TC, HDL	A	C	0.458	0.460
rs6511720[1]	LDLR	19	11202306	LDL, TC	T	G	0.085	0.106
rs2228671[1]	LDLR	19	11210912	TC	T	C	0.076	0.106
rs2075650[1]	TOMM40	19	45395619	TC	G	A	0.075	0.159
rs4420638	APOE	19	45422946	LDL, TG	G	A	0.042	0.183

[1]SNPs included in the initial SEM, i.e., those that passed quality filters, exhibited nominal significance, and were top SNPs in an LD block.

variables, FPG was excluded from the model because of insignificant associations with all factors, MetS and metabolic diseases; its removal significantly improved model fit. The SEM was built using the combined dataset of males and females while including sex as a covariate in all regression statements. We extracted 40 highly replicated SNPs uncovered through GWAS of obesity, glucose, blood pressure, and lipid phenotypes (Table 3). Of these 40, 19 exhibited nominal significance with metabolic traits through a replication association study in our population (association *P*-values are provided in Table 4) and were included in the initial SEM.

In the final SEM (Table 5 and Figure 2), the obesity factor was significantly associated with all metabolic traits except HbA1c, including the lipid factor ($\beta = -0.35$, $P < 0.0001$), the blood pressure

TABLE 4 Replication P-values of studied GWAS variants

CHR	SNP	Weight	BMI	WC	HC	WHR	SBP	DBP	UA	TC	TG	HDL	LDL	HbA1c
1	rs646776	**0.041**	**0.043**	0.119	0.061	0.760	0.095	0.138	0.615	**0.041**	0.642	0.155	**0.007**	0.884
1	rs599839	**0.044**	**0.049**	0.123	0.068	0.734	0.094	0.135	0.598	**0.033**	0.666	0.155	**0.006**	0.888
2	rs6754295	0.631	0.646	0.593	0.522	0.943	0.534	0.937	0.956	0.499	0.076	0.264	0.819	0.388
2	rs693	0.278	0.681	0.344	0.396	0.633	0.257	0.594	0.102	0.763	0.339	0.208	0.742	0.370
2	rs515135	0.094	0.396	0.391	0.062	0.316	0.435	0.087	0.067	0.057	0.763	0.519	**0.004**	0.735
2	rs780094	0.120	0.438	0.212	0.284	0.616	0.995	0.918	0.936	0.133	**0.009**	0.056	0.313	0.364
3	rs419076	0.839	0.365	0.831	0.936	0.768	0.343	0.595	0.324	0.559	0.599	0.468	0.219	0.280
4	rs874432	0.483	0.122	0.497	0.509	0.783	0.258	0.163	**1E−12**	0.635	0.730	0.893	0.368	0.504
4	rs16998073	0.528	0.474	0.167	0.655	0.145	0.735	0.418	0.396	0.074	0.142	0.496	0.082	0.809
5	rs3846662	0.854	0.086	0.156	0.702	**0.011**	0.328	0.661	0.555	0.102	0.487	0.101	0.070	0.125
8	rs328	0.841	0.167	0.319	0.487	0.409	0.402	0.658	0.748	0.755	0.465	0.053	0.964	0.459
8	rs10096633	0.541	0.245	0.448	0.610	0.391	0.580	0.749	0.726	0.519	0.705	**0.014**	0.776	0.681
8	rs2083637	0.429	0.665	0.970	0.461	0.311	0.788	0.807	0.644	0.591	0.618	0.315	0.367	0.642
10	rs11014166	**0.048**	0.405	0.553	0.135	0.345	0.608	0.752	0.090	**0.001**	**0.004**	0.941	**0.005**	0.302
10	rs7100623	0.554	0.391	0.831	0.415	0.395	0.826	0.518	0.312	0.060	**0.017**	0.230	**0.018**	**0.003**
10	rs1004467	0.351	0.937	0.226	0.414	0.362	0.914	0.582	0.115	0.937	0.867	0.400	0.804	0.165
10	rs11191548	0.411	0.709	0.159	0.449	0.181	0.350	0.610	0.282	0.652	0.993	0.353	0.930	0.215
10	rs12243326	0.231	0.167	0.372	0.393	0.681	0.549	0.857	0.897	0.729	0.474	0.500	0.613	**3.5E−10**
11	rs4938303	0.965	0.634	0.806	0.821	0.525	0.302	0.497	0.513	0.119	0.103	0.158	0.051	0.497
11	rs12272004	0.244	0.682	0.464	0.145	0.585	**0.037**	**0.006**	**0.035**	0.693	**0.0003**	**0.044**	0.107	0.561
12	rs3184504	0.516	0.897	0.812	0.484	0.559	0.050	0.093	0.739	0.944	0.376	0.471	0.772	0.545
15	rs4775041	0.280	0.335	0.335	0.186	0.986	0.718	0.158	0.691	0.598	0.402	0.643	0.572	0.866
15	rs1532085	0.534	0.642	0.719	0.434	0.657	0.873	0.273	0.299	0.837	0.688	0.919	0.535	0.395
15	rs1800588	0.564	0.949	0.469	0.612	0.623	**0.030**	**0.033**	0.445	0.729	**0.047**	**0.000**	0.088	0.569
15	rs1378942	0.561	0.886	0.525	0.688	0.543	**0.018**	0.221	0.148	0.051	0.978	0.291	0.110	0.721
15	rs6495122	0.954	0.518	0.984	0.814	0.671	0.340	0.478	0.298	**0.034**	0.259	0.363	0.166	0.321
16	rs1421085	**0.017**	**0.010**	**0.005**	**0.019**	0.126	0.400	0.393	0.450	0.323	0.225	0.063	0.170	0.667
16	rs1121980	**0.018**	**0.010**	**0.005**	**0.018**	0.146	0.415	0.396	0.467	0.372	0.232	0.088	0.189	0.676
16	rs17817449	**0.022**	**0.020**	**0.010**	**0.014**	0.316	0.490	0.506	0.603	0.372	0.192	0.060	0.180	0.923
16	rs8050136	**0.022**	**0.020**	**0.010**	**0.014**	0.316	0.490	0.506	0.603	0.372	0.192	0.060	0.180	0.923
16	rs3751812	**0.022**	**0.020**	**0.010**	**0.014**	0.316	0.490	0.506	0.603	0.372	0.192	0.060	0.180	0.923
16	rs9939609	**0.022**	**0.020**	**0.010**	**0.014**	0.316	0.490	0.506	0.603	0.372	0.192	0.060	0.180	0.923
16	rs7190492	**0.009**	**0.013**	**0.004**	**0.010**	0.103	**0.039**	0.222	0.388	0.058	0.245	0.209	0.146	0.719
16	rs173539	0.486	0.277	0.775	0.542	0.699	0.473	0.656	0.805	0.465	0.706	**0.001**	0.795	0.717
16	rs1800775	0.265	0.512	0.436	0.185	0.688	0.453	0.515	0.505	0.159	0.410	**0.0003**	0.687	0.656
16	rs1532624	0.795	0.816	0.963	0.964	0.977	0.565	0.677	0.482	0.236	0.446	**0.0003**	0.993	0.478
19	rs6511720	0.166	0.208	0.350	**0.028**	0.158	0.603	0.783	0.504	0.679	0.283	0.897	0.595	0.076
19	rs2228671	0.117	0.136	0.109	**0.040**	0.825	0.217	0.352	0.916	0.562	0.832	0.929	0.367	0.217
19	rs2075650	0.150	0.250	0.242	0.275	0.691	0.692	0.589	0.611	0.309	0.683	0.854	0.706	**0.024**
19	rs4420638	0.195	0.124	0.527	0.451	0.941	0.907	0.639	0.682	0.560	0.199	0.522	0.160	0.082

SNPs with nominal significance ($P < 0.05$) are given in bold font.

factor ($\beta = 0.27$, $P < 0.0001$), UA (per-level OR = 1.18, $P < 0.0001$), MetS (OR = 1.68, $P < 0.0001$), and type 2 diabetes (OR = 1.13, $P = 0.008$). The factor with the largest effect (based on the regression coefficient) on MetS, however, was the lipid factor (OR = 0.54, $P < 0.0001$), which was also associated with UA (per-level OR = 0.80, $P < 0.0001$) and HbA1c ($\beta = -0.02$, $P < 0.0001$). The blood pressure factor and UA were also strongly associated with MetS (OR = 1.33, $P < 0.0001$ and $\beta = 1.25$, $P < 0.0001$, respectively); furthermore, UA was also associated with gout (OR = 1.405, $P < 0.0001$). HbA1c was highly associated with type 2 diabetes (OR = 1.03, $P < 0.0001$), coronary heart disease (OR = 1.01, $P < 0.0001$), and somewhat less significantly associated with MetS (OR = 1.01, $P = 0.0013$).

In relation to metabolic diseases, MetS was significantly associated only with gout (OR = 1.16, $P = 0.0245$). Among the metabolic diseases, coronary heart disease was associated with risk for kidney disease (OR = 2.43, $P < 0.0001$) and type 2 diabetes was associated

TABLE 5 Parameter estimates of structural equations model

Outcome	Predictor	Estimate	S.E.	T-value	P-value
Obesity factor[1]	Weight	3.48	0.08	45.33	<0.0001
	Waist circumference	10.07	0.22	46.64	<0.0001
	Hip circumference	6.98	0.18	37.97	<0.0001
	BMI	12.16	0.25	48.79	<0.0001
	Waist–hip ratio	0.05	0.00	39.88	<0.0001
Lipid factor[1]	LDL	−0.17	0.03	−5.46	<0.0001
	Cholesterol	−0.28	0.05	−6.09	<0.0001
	Triglycerides	−0.46	0.05	−8.52	<0.0001
	HDL	0.16	0.02	7.37	<0.0001
BP factor[1]	SBP	16.852	0.31	54.419	<0.0001
	DBP	8.407	0.169	49.665	<0.0001
Obesity factor	rs10096633 (LPL)	0.04	0.01	3.14	0.0022
	rs12243326 (TCF7L2)	−0.07	0.03	−2.43	0.0169
	rs1121980 (FTO)	0.10	0.04	2.32	0.0224
Lipid factor	Obesity factor	−0.35	0.05	−7.05	<0.0001
	rs1532624 (CETP)	0.14	0.05	2.83	0.0055
	rs515135 (APOB)	0.10	0.05	2.13	0.035
	rs1227204 (RPL15P15)	−0.29	0.08	−3.84	0.0002
	rs780094 (GCKR)	−0.12	0.05	−2.73	0.0074
BP factor	Obesity factor	0.27	0.02	12.15	<0.0001
HbA1c	Lipid factor	−0.02	0.00	−4.22	<0.0001
	rs12243326 (TCF7L2)	0.02	0.01	3.03	0.003
	rs7100623 (HHEX/IDE)	0.03	0.00	11.40	<0.0001
	rs2075650 (TOMM40)	−0.01	0.00	−2.55	0.012
	rs1800588 (LIPC)	0.01	0.00	4.46	<0.0001
Uric acid	SEX	1.35	0.08	17.77	<0.0001
	Obesity factor	0.17	0.03	5.81	<0.0001
	Lipid factor	−0.22	0.04	−5.93	<0.0001
	rs874432 (SLC2A9)	−0.37	0.06	−6.20	<0.0001
MetS	Obesity factor	0.52	0.05	9.56	<0.0001
	HbA1c	0.01	0.00	3.29	0.0013
	Uric acid	0.23	0.05	4.95	<0.0001
	Lipid factor	−0.62	0.11	−5.74	<0.0001
	BP factor	0.29	0.05	6.52	<0.0001
Type 2 diabetes	Obesity factor	0.12	0.05	2.68	0.0084
	HbA1c	0.03	0.00	24.26	<0.0001
	rs7100623 (HHEX/IDE)	0.33	0.08	3.87	0.0002
Gout	MetS	0.15	0.06	2.28	0.0245
	Uric acid	0.34	0.06	5.69	<0.0001
Heart disease	HbA1c	0.01	0.00	4.16	<0.0001
Stroke	Type 2 diabetes	0.47	0.11	4.42	<0.0001
Kidney disease	Heart disease	0.89	0.22	4.03	0.0001

[1]The predictor estimates are factor loading estimates for SEM factors.

with stroke (OR = 1.599, $P < 0.0001$). Gender was significant only in the association involving UA as the outcome.

Of the 19 SNPs with nominal replication associations, 11 were significantly associated with latent factors, unique manifest variables, and type 2 diabetes (Table 5); the eight GWAS SNPs that were not significantly associated with variables in the model were removed through an iterative, stepwise backwards elimination process. The majority of variants were associated only with latent factors or unique manifest variables underlying their original GWAS traits. Two SNPs, however, revealed associations beyond their original traits: 1) HbA1c-associated SNP rs7100623 in HHEX/IDE with type 2 diabetes (OR = 1.39, $P = 0.0002$) and 2) HbA1c-associated SNP rs12143326 in TCF7L2 with the obesity factor ($\beta = -0.066$, $P =$

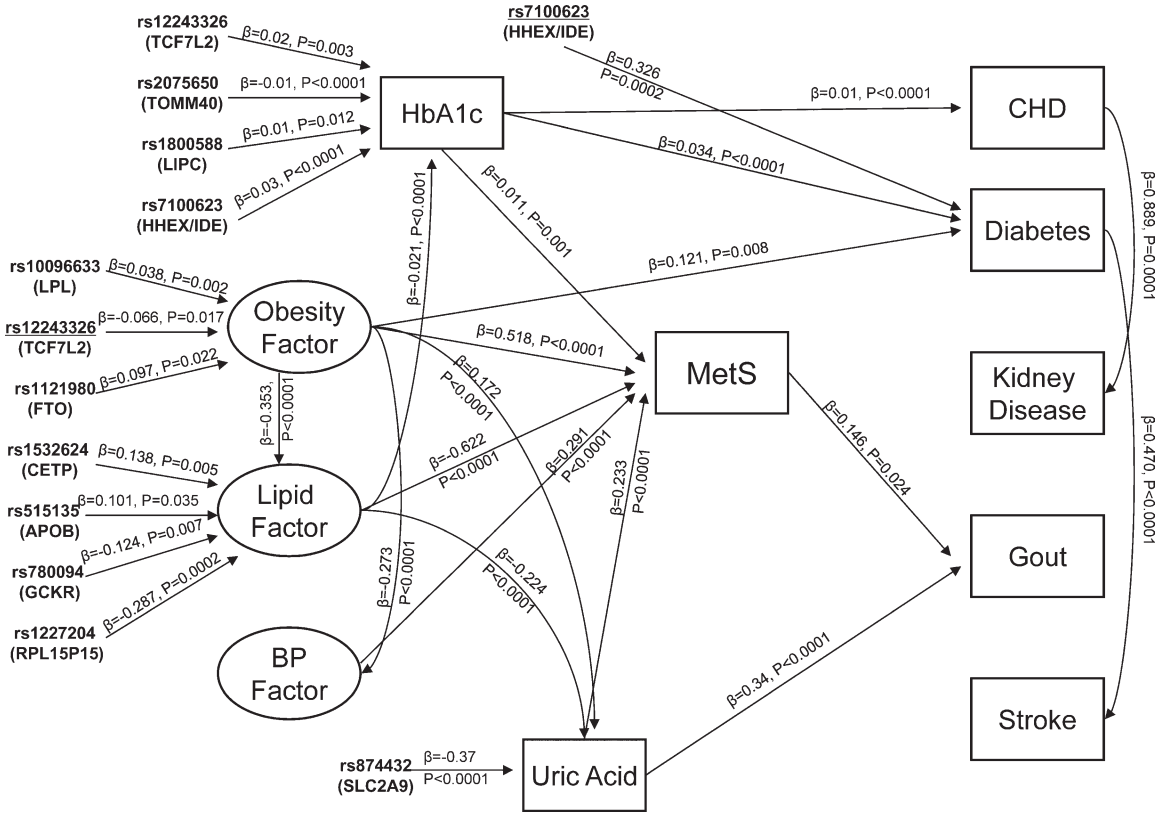

FIGURE 2 Path diagram of the final SEM. Latent variables are depicted in ovals and all measured variables are presented in rectangles. Underlined SNPs are those with nominally significant associations beyond their original GWAS trait.

0.017). However, testing the association of SNPs with traits beyond their original GWAS trait resulted in a discovery-based analysis, which required correction for multiple testing. Following the application of a Bonferroni correction based on the total number of SNPs tested (19), only the association between the HHEX/IDE SNP and type 2 diabetes remained significant ($P = 0.004$).

The final model was well-fit according to CFI (0.951), TLI (0.956), and RMSEA (0.037), and is presented in Figure 2 and Table 5. Our final χ^2 value was 330.45 on 120 degrees of freedom ($P < 0.0001$), but as previously stated, in a model with multiple categorical outcomes this statistic is overly conservative. The final model was re-run using the data without imputed missing values and the overall model and individual associations were not significantly different. Figure 2 provides a path diagram of the final SEM, where latent factors are depicted in ovals and all measureable variables are presented in rectangles. Underlined SNPs are those with nominal associations beyond their original GWAS trait. The path diagram underscores the progression from individual metabolic parameters to MetS, subsequent metabolic diseases, and the association of single-trait GWAS SNPs within the complex network.

Discussion

To our knowledge, this is the first study to assess the composite-network relationships between SNPs derived from single-trait GWAS, latent factors underlying redundant and collinear metabolic parameters, unique manifest variables involved in cardiovascular health, MetS, and downstream metabolic disease. In a sample from an isolated island population with high prevalence of obesity, hypertension, and Mets, 13 measured traits were distilled into three continuous latent factors (obesity, lipid, and blood pressure factors) and two unique manifest variables (HbA1c and UA). Four of these traits encompass those included in MetS diagnosis (obesity, dyslipidemia, hypertension, and glucose intolerance). Of interest, elevated UA (our fifth predictor of MetS) is associated with gout, which was included in the original description of MetS elucidated by Kylin (cf. Ref. 6). As expected, body fatness appears to be a driving element behind most factors and traits involved in the development of MetS and metabolic diseases. These results underscore the far-reaching negative impact of overweight and obesity on metabolic health. We found, however, that the magnitude of effect of lipid control on MetS risk was greater than that of obesity. In addition to its relationship with MetS, the lipid factor was also associated with HbA1c and UA, emphasizing the tight integration of obesity, glucose intolerance, dyslipidemia, and hyperuricemia, and underscoring a potentially key role of dyslipidemia in metabolic health of this population.

In this study, we have observed MetS–metabolic disease associations only between MetS and gout, which has no detected association with other metabolic diseases. In the final model, HbA1c, UA, and the obesity factor were associated with metabolic diseases (HbA1c with coronary heart disease and diabetes, UA with gout, and the

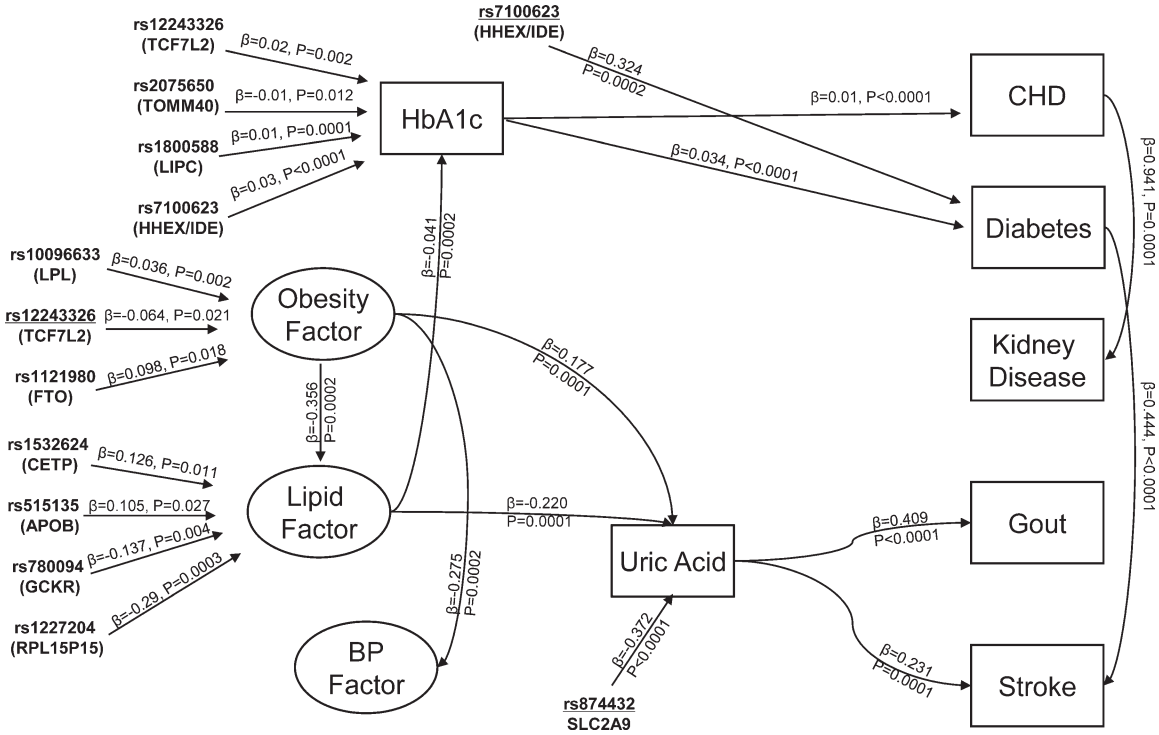

FIGURE 3 Path diagram of the SEM without MetS diagnosis. Latent variables are depicted in ovals and all measured variables are presented in rectangles. Underlined SNPs are those with nominally significant associations beyond their original GWAS trait.

obesity factor with type 2 diabetes). Subsequently, coronary heart disease was related to kidney disease and type 2 diabetes was associated with stroke.

Although we do not see several expected associations (e.g., dyslipidemia with CHD), this arises from the inclusion of MetS as the intermediate between factors and metabolic diseases. The majority of factor variability is absorbed in the MetS diagnosis, which yields insignificant associations between latent factors and metabolic diseases. However, in our model, we also do not find significant associations between MetS and further disease. Therefore, we do not provide support to the utility of MetS as an indicator of elevated metabolic disease risk, as its only association was with gout, and the association of UA with gout was of greater magnitude and significance. In further analysis, we removed MetS from the model and found that its removal led to a comparably fit model ($\chi^2 = 309.64$ on 118 degrees of freedom, CFI and TLI = 0.05, and RMSEA = 0.036), with similar power to predict metabolic diseases (Figure 3). We therefore suggest that MetS may not be a true biological disorder, in accordance with the findings of multiple previous reports (26-28). Our results indicate that it is instead a useful clinical indicator of the co-occurrence of cardiovascular disease risk factors that each individually contribute to risk for metabolic disease. We do, however, demonstrate the domino-effect of the composite MetS traits, particularly obesity, glucose intolerance, and dyslipidemia, and their impact on downstream metabolic diseases.

Eleven genetic variants derived from single-trait GWAS were significant predictors of latent factors, HbA1c, UA, MetS, and metabolic diseases. Nine of these were solely associated with the unique manifest variables and latent factors underlying their original GWAS trait. The effects of two variants, both HbA1c-related SNPs, reached beyond the original replication trait. The concurrent association of the HbA1c-related SNP rs7100623 in HHEX/IDE with type 2 diabetes indicates a significant role of this gene in type 2 diabetes, because the association between the SNP and type 2 diabetes remained after accounting for the SNP impact on HbA1c and the association of HbA1c with type 2 diabetes. The second HbA1c-related SNP (rs12243326) was simultaneously associated with increased HbA1c and a decrease in the obesity factor. Previous GWAS that detected significant signals in TCF7L2 have not included obesity-related traits and do not provide support for this association. Although a recent study reported a variant of TCF7L2 to be protective against obesity in Mexican children (37), this population may be evolutionarily distinct from our study population and these findings have not been replicated. With little evidence to support an association between this well-studied gene and obesity, we suggest that the association may be a model-related phenomenon arising from a spurious association between residual variances arising by chance.

Although we report the associations within the MetS network of traits and diseases, we cannot establish causality because of the cross-sectional design of the study. We are also unable to determine a temporal relationship between individual traits, MetS, and metabolic disease, although MetS is considered a premorbid condition. In addition, translation of this model to clinical practice may be premature as the majority of the model involves unmeasurable and unit-

free latent factors. This study instead improved our understanding of the complete MetS network and downstream metabolic diseases by providing structure and context to these relationships through a quantitative map of the disease network. On the basis of our sampling strategy, our study sample may represent a healthier subset of the population; though by using a population with homogeneous environmental exposures and genetic background, we eliminate potential confounding factors from analysis. However, extrapolation of these results to other populations would require further validation as all metabolic disease diagnoses were based on self-report, and the relative isolation and homogeneity of the population may have led to associations specific to this study (e.g., the high importance of dyslipidemia and uric acid). Finally, we share a limitation of most SEMs, in that our study may be underpowered to detect significant associations, given the high number of variables included in our model.

We provide a quantitative map of metabolic parameters, MetS, and GWAS-extracted SNPs, and give context to the impact and relevance of MetS, particularly on the development of five metabolic diseases. In this complex and highly interrelated system, we underscore the importance of obesity, dyslipidemia, glucose intolerance, and hyperuricemia as driving elements in the development of metabolic disease, and suggest the utility of MetS diagnosis as a clinical rather biological phenomenon. O

Acknowledgments

The authors acknowledge the guidance of Dr. Anil Menon and Dr. M.B. Rao during model interpretation and manuscript preparation. R.D. is the guarantor of this study.

© 2013 The Obesity Society

References

1. Ervin RB. Prevalence of metabolic syndrome among adults 20 years of age and over, by sex, age, race, and ethnicity, and body mass index: United States, 2003-2006. *National Health Statistics Reports*, Number 13.
2. Cowey S, Hardy RW. The metabolic syndrome: a high-risk state for cancer? *Am J Pathol* 2006;169:1505-1522.
3. Laaksonen DE, Lakka HM, Lynch J, et al. The metabolic syndrome and total and cardiovascular disease mortality in middle-aged men. *JAMA* 2002;288:2709-2716.
4. Meigs JB. The metabolic syndrome. *BMJ* 2003;327:61-62.
5. Malik S, Wong ND, Franklin SS, Kamath TV, L'Italien GJ, Pio JR. Impact of the metabolic syndrome on mortality from coronary heart disease, cardiovascular disease, and all causes in United States adults. *Circulation* 2004;10:1245-1250.
6. Eckel RH, Grundy SM, Zimmet PZ. The metabolic syndrome. *Lancet* 2005;365:1415-1428.
7. Galassi A, Reynolds K, He J. Metabolic syndrome and risk of cardiovascular disease: a meta-analysis. *Am J Med* 2006;119:812-819.
8. Lusis AJ, Attie AD, Reue K. Metabolic syndrome: from epidemiology to systems biology. *Nat Rev Genet* 2008;9:819-830.
9. Haffner S, Taegtmeyer H. Epidemic obesity and the metabolic syndrome. *Circulation* 2003;108:1541-1545.
10. Cameron AJ, Shaw JE, Zimmet PZ. The metabolic syndrome: prevalence in worldwide populations. *Endocrinol Metab Clin N Am* 2004;33:351-375.
11. Song J, Kim E, Shin C, et al. Prevalence of the metabolic syndrome among South Korean adults: the Ansan study. *Diabet Med* 2004;21:1154-1155.
12. Gu D, Reynolds K, Wu X, et al. Prevalence of the metabolic syndrome and overweight among adults in China. *Lancet* 2005;365:1398-1405.
13. Jorgensen ME, Moustgaard H, Bjerregaard P, Borch-Johnsen K. Gender differences in the association between westernization and metabolic risk among Greenland Inuit. *Eur J Epidemiol* 2006;21:741-748.
14. Keighley ED, McGarvey ST, Turituri P, Viali S. Farming and adiposity in Samoan adults. *Am J Hum Biol* 2006;18:112-122.
15. Kolcic I, Vorko-Jovic A, Salzer B, Smoljanovic M, Kern J, Vuletic S. Metabolic syndrome in a metapopulation of Croatian island isolates. *Croat Med J* 2006;47:585-592.
16. Deka R, Durakovic Z, Niu W, et al. Prevalence of metabolic syndrome and related metabolic traits in an island population of the Adriatic. *Ann Hum Biol* 2012;39:46-53.
17. Alberti KG, Zimmet PZ. Definition, diagnosis and classification of diabetes mellitus and its complications. I. Diagnosis and classification of diabetes mellitus provisional report of a WHO consultation. *Diabet Med* 1998;15:539-553.
18. National Institutes of Health. Third Report of the National Cholesterol Education Program Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III). National Institutes of Health, NIH Publication No 01-3670, Bethesda, MD, 2001.
19. Grundy SM, Cleeman JI, Daniels SR, et al. Diagnosis and management of the metabolic syndrome: an American Heart Association/National Heart, Lung, and Blood Institute Scientific Statement. *Circulation* 2005;112:2735-2752.
20. International Diabetes Federation. The IDF consensus worldwide definition of the metabolic syndrome. Available at: http://www.idf.org/webdata/docs/MetSyndrome_Final.pdf (Accessed April 14, 2005).
21. Chan JC, Cheung JC, Lau EM, et al. The metabolic syndrome in Hong Kong Chinese. The interrelationships among its components analyzed by structural equation modeling. *Diabetes Care* 1996;19:953-959.
22. Zhang S, Liu X, Yu Y, et al. Genetic and environmental contributions to phenotypic components of metabolic syndrome: a population-based twin study. *Obesity (Silver Spring)* 2009;17:1581-1587.
23. Klaus JR, Hurwitz BE, Llabre MM, et al. Central obesity and insulin resistance in the cardiometabolic syndrome: pathways to preclinical cardiovascular structure and function. *J Cardiometab Syndr* 2009;4:63-71.
24. Nock NL, Wang X, Thompson CL, et al. Defining genetic determinants of the Metabolic Syndrome in the Framingham Heart Study using association and structural equation modeling methods. *BMC Proc* 2009;15 (Suppl 7):S50.
25. Shen BJ, Todaro JF, Niaura R, et al. Are metabolic risk factors one unified syndrome? Modeling the structure of the metabolic syndrome x. *Am J Epidemiol* 2003;157:701-711.
26. Eckel RH, Alberti KG, Grundy SM, Zimmet PZ. The metabolic syndrome. *Lancet* 2010;275:181-183.
27. Simmons RK, Alberti KG, Gale EA, et al. The metabolic syndrome: useful concept or clinical tool? Report of a WHO Expert Consultation. *Diabetologia* 2010;53:600-605.
28. Kahn R, Buse J, Ferrannini E, Sterm M. The metabolic sydnrome: time for a critical appraisal. Joint statement from the American Diabetes Association and the European Association for the Study of Diabetes. *Diabetologia* 2005;48:1684-1699.
29. Teran-Garcia M, Bouchard C. Genetics of the metabolic syndrome. *Appl Physiol Nutr Metab* 2007;32:89-114.
30. Rudan I, Rudan D, Campbell H, et al. Inbreeding and risk of late onset complex disease. *J Med Genet* 2003;40:925-932.
31. Pucarin-Cvetković J, Mustajbegović J, Doko Jelinić J, et al. Body mass index and nutrition as determinants of health and disease in population of Croatian Adriatic islands. *Croat Med J* 2006;47:619-626.
32. Rudan P, Sujoldzic, Simic D, Bennet LA, Roberts DF. *Population structure in eastern Adriatic: the influence of historical processes, migration patterns, isolation and ecological pressures, and their interaction*. In: *Isolation, Migration and Health* Roberts DF, Fujiki N, Torizuka K, eds. Cambridge (SSHB): Cambridge University Press; 1992, pp 204-218.
33. Box GP, Cox DR. An analysis of transformations (with discussion). *J R Stat Soc B* 1964;26:211-252.
34. Little RJA, Rubin DB. *Statistical Analysis with Missing Data*. New York: Wiley; 1987.
35. Hindorff LA, Junkins HA, Hall PN, Mehta JP, Manolio TA. A catalog of published genome-wide association studies, 2011. Available at: www.genome.gov/gwastudies.
36. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct Equat Model* 1999;6:1-55.
37. Klünder-Klünder M, Mejía-Benitez MA, Flores-Huerta S, Burguete-García AI, García-Mena J, Cruz M. rs1225537 variant of TCF7L2 gene is protective for obesity in Mexican children. *Arch Med Res* 2011;42:495-501.