G&I Genomics & Informatics

eISSN 2234-0742
Genomics Inform 2014;12(4):187-194
http://dx.doi.org/10.5808/GI.2014.12.4.187

ORIGINAL ARTICLE

Genome-Wide Association Study of Metabolic Syndrome in Koreans

Seok Won Jeong, Myungguen Chung, Soo-Jung Park, Seong Beom Cho, Kyung-Won Hong*

Division of Bio-Medical Informatics, Center for Genome Science, National Institute of Health, KCDC, Cheongju 363-951, Korea

Metabolic syndrome (METS) is a disorder of energy utilization and storage and increases the risk of developing cardiovascular disease and diabetes. To identify the genetic risk factors of METS, we carried out a genome-wide association study (GWAS) for 2,657 cases and 5,917 controls in Korean populations. As a result, we could identify 2 single nucleotide polymorphisms (SNPs) with genome-wide significance level p-values ($< 5 \times 10^{-8}$), 8 SNPs with genome-wide suggestive p-values ($5 \times 10^{-8} \leq p < 1 \times 10^{-5}$), and 2 SNPs of more functional variants with borderline p-values ($5 \times 10^{-5} \leq p < 1 \times 10^{-4}$). On the other hand, the multiple correction criteria of conventional GWASs exclude false-positive loci, but simultaneously, they discard many true-positive loci. To reconsider the discarded true-positive loci, we attempted to include the functional variants (nonsynonymous SNPs [nsSNPs] and expression quantitative trait loci [eQTL]) among the top 5,000 SNPs based on the proportion of phenotypic variance explained by genotypic variance. In total, 159 eQTLs and 18 nsSNPs were presented in the top 5,000 SNPs. Although they should be replicated in other independent populations, 6 eQTLs and 2 nsSNP loci were located in the molecular pathways of *LPL, APOA5*, and *CHRM2*, which were the significant or suggestive loci in the METS GWAS. Conclusively, our approach using the conventional GWAS, reconsidering functional variants and pathway-based interpretation, suggests a useful method to understand the GWAS results of complex traits and can be expanded in other genomewide association studies.

Keywords: expression quantitative trait loci, genome-wide association study, metabolic networks and pathways, single nucleotide polymorphism

Introduction

Metabolic syndrome (METS) is a disorder of energy utilization and storage and increases the risk of developing cardiovascular disease and diabetes. METS includes multiple clinical traits, as follows: increased plasma glucose, abdominal obesity, dyslipidemia, and high blood pressure [1]. METS is a great concern in developing countries, because the prevalence of METS is gradually increasing, especially in countries where it follows obesity trends, sedentary lifestyle, and high consumption of calories [2-4]. A Korean twin study showed that the METS has 51%–60% heritability, indicating a significant role of genetic factors in the development of METS [5]. Therefore, understanding the genetic factors underlying the syndrome and their correlation is clinically important.

Recent advances in high-throughput genomics technologies have allowed massive testing of genetic variants in minimal time [6, 7]. The reductions in cost and time have made it feasible to conduct large-scale genome-wide association studies (GWASs) that genotype many thousands of single nucleotide polymorphisms (SNPs) in thousands of individuals. So far, approximately 1,900 reports with 13,000 SNPs (GWAS catalog; Apr 14, 2014) have been published to identify the gene-disease or -non-disease trait associations. The quantitative traits related to METS have already been studied by the conventional GWAS in the Korean population [8], but there is no Korean GWAS report about METS cases and controls.

On the other hand, relatively little trait heritability can be explained by the conventional GWAS [9]. These pheno-

Received July 8, 2014; Revised September 16, 2014; Accepted September 29, 2014

***Corresponding author:** Tel: +82-43-719-8858, Fax: +82-43-719-8869, E-mail: zknightm@gmail.com

This is 2014 KNIH KARE best paper awarded.

mena, called missing heritability problems, are hard to solve by conventional GWAS, in part because of the extensive multiple testing correction in GWAS analysis, low effect size of common variants, and the difficulty of detecting low-frequency or rare variants in conventional GWAS. Multiple testing correction is necessary to exclude false-positive loci, but simultaneously, it discards many true-positive loci [10]. Also, most SNPs in these GWASs lie in intergenic and intron regions and do not appear to affect protein sequence. Thus, these SNPs are likely functionally neutral or just proxies of causal variants located in the same linkage disequilibrium (LD). To understand the amount of true-positive signals in the discarded association results, we computed the proportion [V(G)/V(P)] of the phenotypic variance [V(P)] that is explained by the genotype variance [V(G)] using the significant and discarded SNP results [11].

Recently, Fransen *et al*. [10] reported a GWAS using expression quantitative trait loci (eQTL) information. They selected eQTL among the GWAS results for Crohn's disease and conducted follow-up replication studies [10]. They showed that eQTL-based pre-selection for follow-up is a useful approach for identifying risk loci from a moderately sized GWAS.

Based on previous knowledge, we applied an alternative analysis strategy to understand the genetic components of METS. First of all, we conducted a conventional GWAS for METS cases and healthy controls to discover the top significant signals. Thereafter, we tried to uncover the functional variants, such as nonsynonymous SNPs (nsSNPs) and eQTLs, among the SNPs to be discarded using the stringent criteria of the conventional GWAS. Finally, we drew a pathway of the significantly associated GWAS SNPs and the remaining less significantly associated functional SNPs. The overall study design is schematically described in Supplementary Fig. 1.

Methods

Study subjects

The study subjects were originally derived from a part of the Korean Genome and Epidemiology Study (KoGES) project, which was the national project to establish genome epidemiology cohorts of Korean dwellers or immigrants/emigrants [12]. Among the KoGES cohorts, the Korean Association Resource Consortium (KARE) has established a public GWAS dataset by using the Ansan-Anseong cohort, which is an ongoing biennially followed-up cohort in the KoGES [13]. The KARE dataset consists of the individual SNP chip genotypes and the epidemiological/clinical phenotypes for studying the genetic components of Korean public health. Written informed consent was obtained from all participants at the KoGES, and this research project was approved by the Institutional Review Board of Korea National Institute of Health (KNIH). The obtained KARE dataset passed the quality control criteria and was reported in previous GWAS reports [8, 13]. Briefly, the subjects with genotype accuracies below 98% and high missing genotype call rates ($\geq$5%), high heterozygosity ($>$30%), or inconsistency in sex were excluded from subsequent analysis. Individuals who had a tumor were excluded, as were related individuals whose estimated identity-by-state values were high ($>$0.80). Based on these criteria, 8,842 samples were selected; these quality control steps have been described in a previous GWAS [13].

Study phenotypes and covariates

We used the general information on resident areas (Anseong or Ansan), sex, and age as the covariates and past disease history of diabetes, hypertension, and lipidemia as exclusion criteria for non-METS healthy controls. The height and body weight were used to calculate the body mass index (BMI) as another covariate, and waist circumference (WC), systolic and diastolic blood pressures (SBP and DBP), fasting plasma glucose levels (GLU0), high-density lipoprotein (HDL) cholesterol, and triglyceride (TG) were used to diagnose METS. METS was defined by the presence of three or more of the following five components according to the NCEP-ATPIII criteria using WC for Asians [14, 15]: WC ($\geq$90 cm for men and $\geq$80 cm for women), HDL ($<$40 mg/dL for men, $<$50 mg/dL for women), TG ($\geq$150 mg/dL), SBP ($\geq$130 mm Hg) and/or DBP ($\geq$85 mm Hg), and GLU0 ($\geq$100 mg/dL).

Study genotypes

The genotyping of the cohort population was previously described for the KARE study [16]. Most DNA samples were isolated from the peripheral blood of participants and genotyped using Affymetrix Genomewide Human SNP array 5.0 (Affymetrix, Inc., Santa Clara, CA, USA). The quality control steps of the genotypes have been described elsewhere [13]. Briefly, the calling of the genotyping was determined by Bayesian Robust Linear Modeling using the Mahalanobis Distance genotyping algorithm [17]. Consequently, 352,227 SNPs had a missing genotype call rate below 0.1, a minor allele frequency greater than 0.01, and no deviation from Hardy-Weinberg equilibrium ($p > 1 \times 10^{-6}$). Additionally, the previous GWAS reported no population stratification between the Anseong and Ansan cohorts [13].

Statistical analysis

The GWAS for METS cases and controls was conducted by logistic regression analysis, adjusting for residential area,

sex, age, and BMI as covariates, implemented in PLINK version 1.07 [18]. The significant associations were defined by genomewide significance level p-values ($< 5 \times 10^{-8}$) and genome-wide suggestive p-values ($5 \times 10^{-8} \leq p < 1 \times 10^{-5}$).

The LD between the previously reported GWAS SNPs and the SNPs of the current GWAS was investigated with SNAP web-based software (http://www.broadinstitute.org/mpg/snap) and GWAS catalog (http://www.genome.gov/gwastudies/). For example, we entered our top significant SNPs in the SNAP input panel and found high LD SNPs with $r^2 > 0.9$ and D' = 1 around 1 Mbp. The high-LD SNPs were investigated in the GWAS catalog (http://www.genome.gov/gwasstudies) as to whether they were previously reported or not.

To maximize the candidate risk factors of METS, we selected additional functional SNPs in the eQTLs or nsSNP loci ($5 \times 10^{-5} \leq p < 1 \times 10^{-4}$). Among the Affymetrix 5.0 SNPs, we investigated the eQTL SNPs from regulomeDB (http://regulomedb.org) and the nsSNPs from BioMart (http://www.biomart.org).

The genetic variances of the top association SNPs were estimated by GCTA v1.24 [16], which is a tool for estimating the proportion [V(G)/V(P)] of phenotypic variance [V(P)] explained by SNPs [V(G)] for complex traits. We selected the SNP sets based on the GWAS p-values from 100 to 1,000 SNPs with 10-SNP intervals and from 1,000 to 5,000 SNPs with 1,000-SNP intervals. We decided the number of SNPs in the maximum set based on the genetic variance approximated the METS heritability reported from the Korean twin study [5]. The pair-wise genetic relationships were estimated using the make-grm option, and the proportion of the phenotypic variance explained by the associated SNPs was estimated by the grm-test option with the restricted maximum likelihood [11].

In silico analysis

The functional relevance of the associated SNP sites was analyzed by overlapping the gene-coding sequence or the Encyclopedia of DNA Element (ENCODE) regulatory element positions in the University of California Santa Cruz (UCSC) genome browser (http://genome.ucsc.edu). Thereafter, regulomeDB (http://regulome.stanford.edu/) was utilized to extract eQTL information. In addition, Pathway Studio version 9.0 software (Ariadne Genomics, Rockville, MD, USA) was utilized to analyze the functional interactions and possible pathways among genes/proteins in our data. It provides an interpretation of the biological implications from gene/protein expression data, the establishment of molecular pathways, and an identification of protein interaction maps and their association to cellular process [19].

Results and Discussion

Genome-wide association study

Table 1 describes the clinical characteristics of Ansung and Ansan regarding the METS criteria: BMI, WC, SBP, DBP, GLU0, HDL, and TG. Based on the NCEP-ATPIII METS criteria for Asians [14], 2,657 KARE subjects were included in the METS cases. The SNPs showing strong and moderate

Table 1. Clinical characteristics of metabolic syndrome-related traits

	Total	Anseong	Ansan	p-value
No. of individuals	8,842	4,205	4,637	
Gender (men:women)	4,183 (47.3):4,659 (52.7)	1,809 (43.0):2,396 (57.0)	2,374 (51.2):2,263 (48.8)	
Age (y)	52.2 ± 8.9	55.7 ± 8.7	49.1 ± 7.9	<0.01
Height (cm)	160.0 ± 8.7	158.3 ± 8.6	161.6 ± 8.4	<0.01
Body mass index (kg/m^2)	24.6 ± 3.1	24.5 ± 3.3	24.7 ± 3.0	<0.01
Fasting glucose (mg/dL)	87.7 ± 21.9	85.9 ± 18.3	89.1 ± 24.4	<0.01
DBP (mm Hg)	80.3 ± 11.5	82.5 ± 10.9	78.2 ± 11.6	<0.01
SBP (mm Hg)	121.7 ± 18.6	126.6 ± 18.8	117.2 ± 17.3	<0.01
Waist circumference (cm)	82.7 ± 8.8	84.6 ± 8.7	81.0 ± 8.5	<0.01
HDL cholesterol (mg/dL)	44.7 ± 10.1	44.6 ± 10.3	44.7 ± 9.9	0.62
TG (mg/dL)	162.9 ± 105.7	165.0 ± 107.2	161.0 ± 104.3	0.07
METS[a]				
METs case/control	2,657/5,917	1,490/2,454	1,167/3,463	<0.01

Values are presented as number (%) or mean ± SD.
DBP, diastolic blood pressure; SBP, systolic blood pressure; HDL, high-density lipoprotein; TG, triglyceride; MET, metabolic syndrome.
[a]METs status: three or more of the component as follows: waist circumference (≥90 cm for men, ≥80 cm for women), HDL (<40 mg/dL for men, <50 mg/dL for women), TG (≥150 mg/dL), blood pressure (SBP ≥ 130 mm Hg or DBP ≥ 85 mm Hg), fasting glucose (≥100 mg/dL).

evidence of association ($p < 1 \times 10^{-5}$) are indicated in the Manhattan plot of the GWAS (Fig. 1). In addition to these SNPs, we identified several functional SNPs with suggestive evidence of association ($5 \times 10^{-8} \leq p$). In this study, we selected 12 SNPs, of which 2 had genome-wide significant associations ($p < 5 \times 10^{-8}$), 8 had suggestive associations ($5 \times 10^{-8} \leq p < 1 \times 10^{-5}$), and 2 had functional variants ($1 \times 10^{-5} \leq p < 1 \times 10^{-4}$) (Table 2) [8, 20-22]. The top SNP (rs11216126) and 3 suggestive SNPs (rs6589566, rs17482753, and rs10503669) were previously reported as being associated with METS-related traits, such as serum cholesterol levels or TG levels [8, 20, 21].

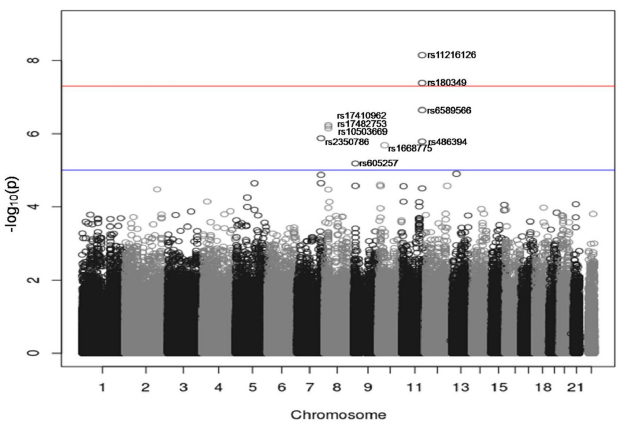

Fig. 1. Manhattan plot of metabolic syndrome genome-wide association study $-\log_{10}$(p-values). All black and grey circles indicate the individual single nucleotide polymorphisms (SNPs). The red horizontal line is the genome-wide significance level ($p = 5 \times 10^{-8}$), and the blue horizontal line is the genome-wide suggestive level ($5 \times 10^{-8} \leq p < 1 \times 10^{-5}$). The top significant SNPs are depicted on the right site of the SNP.

LD analysis using 10 SNPs was conducted with the previously reported GWAS SNPs. As a result, 5 SNPs had strong LD with the 15 highly linked GWAS SNPs (Table 3) [23-34]. Among the 6 remaining SNPs associated with METS in our GWAS results, 2 SNPs (rs180349 and rs17410962) showed high LD with the previously reported SNPs ($r^2 > 0.9$ and $D' = 1$) even though the 2 SNPs have not been reported regarding metabolic traits.

Therefore, we discovered 10 significant or suggestive associated SNPs in the METS GWAS, but 6 of them were already reported or linked to the reported SNPs. The remaining 4 suggestive signals and 2 functional variants have been first reported in the current study, and a replication study should be performed in other independent populations.

In silico annotation of the linked genes and functional relevance

The 10 associated SNPs and the LD SNPs were located in six functional gene regions, and one SNP was located in the intergenic region. The top signals were located downstream of a functional spliceosome-associated protein, named BUD13, a homolog of yeast (*BUD13*) gene chromosome 11 and near the *BUD13* gene. *BUD13* has been reported to be

Table 2. Genome-wide association results for METs case-control study in the Korean population

SNP ID	CHR	BP	Effect allele/other	EAF	Previous GWAS reports for the METS SNPs	OR	95% Confidence interval		p-value
							Lower	Upper	
Genome-wide significant levels ($p < 5 \times 10^{-8}$)									
rs11216126	11	116122450	A/C	0.798	Decrease HDL [8]	1.33	1.21	1.46	7.15×10^{-9}
rs180349	11	116117037	A/T	0.227	-	1.28	1.17	1.4	4.12×10^{-8}
Genome-wide suggestive levels ($5 \times 10^{-8} \leq p < 1 \times 10^{-5}$)									
rs6589566	11	116157633	C/T	0.218	Increase triglycerides [20]	1.27	1.16	1.39	2.26×10^{-7}
rs17410962	8	19892360	G/A	0.876	-	1.34	1.2	1.5	5.75×10^{-7}
rs17482753	8	19876926	G/T	0.876	Increase triglycerides [21]	1.34	1.2	1.51	6.34×10^{-7}
rs10503669	8	19891970	C/A	0.879	Increase triglycerides and decrease HDL cholesterol [8]	1.34	1.2	1.51	7.25×10^{-7}
rs2350786	7	136327110	A/G	0.637	-	1.21	1.12	1.31	1.33×10^{-6}
rs486394	11	116031532	C/A	0.122	-	1.32	1.18	1.47	1.64×10^{-6}
rs1668775	10	36639540	T/C	0.211	-	1.24	1.14	1.36	2.07×10^{-6}
rs605257	9	10300942	T/A	0.770	-	1.22	1.12	1.33	6.48×10^{-6}
Expression quantitative trait loci ($1 \times 10^{-5} \leq p < 1 \times 10^{-4}$)									
rs1996794	11	9779172	C/A	0.411	eQTL of SWAP70 [22]	1.17	1.09	1.27	2.73×10^{-5}
rs1032550	11	9769884	C/T	0.410	eQTL of SWAP70 [22]	1.16	1.08	1.25	2.73×10^{-5}

METS, metabolic syndrome; SNP, single nucleotide polymorphism; CHR, chromosome; BP, base pair based on the human reference genome (hg18); EAF, effect allele frequency; GWAS, genome-wide association study; OR, odds ratio; HDL, high-density lipoprotein; eQTL, expression quantitative trait loci.

Table 3. Reported GWAS SNPs that show LD with our study SNPs

SNP		LD states		Reported traits	References
This study	Reported GWAS	r^2	D'		
Genome-wide significant levels ($p < 5 \times 10^{-8}$)					
rs180349	rs10790162	0.912	0.955	HDL cholesterol and triglycerides	[23]
	rs1558861	0.956	1	Triglycerides	[24]
Genome-wide suggestive levels ($5 \times 10^{-8} \leq p < 1 \times 10^{-5}$)					
rs6589566	rs10790162	0.912	0.955	HDL cholesterol and triglycerides	[23]
	rs2075290	0.956	1	HDL cholesterol and triglycerides	[23]
	rs2160669	1	1	Obesity-related traits	[25]
	rs2266788	1	1	HDL cholesterol and triglycerides	[23]
	rs651821	1	1	Triglycerides	[26]
	rs964184	0.957	1	HDL cholesterol	[27]
rs17410962	rs10096633	1	1	Metabolic traits	[28]
rs17482753	rs1059611	0.925	1	Lipid metabolism phenotypes	[29]
rs10503669	rs12678919	1	1	HDL cholesterol	[30]
	rs17091905	1	1	Cardiovascular disease risk factors	[31]
	rs328	1	1	Triglycerides	[32]
	rs7016880	0.925	1	Hypertriglyceridemia	[33]
	rs7841189	1	1	Metabolic syndrome	[34]

The underlined SNPs indicate that the lead SNPs have not been reported, but there were highly linked GWAS SNPs.
GWAS, genome-wide association study; SNP, single nucleotide polymorphism; LD, linkage disequilibrium; HDL, high-density lipoprotein.

Table 4. GWAS results of eQTLs and nonsynonymous SNPs consisting of the LPL and APOA5 pathways

CHR	SNP	BP	A1	MAF	95% Confidence interval			p-value	Gene	Amino acid substitution	Description
					OR	Lower	Upper				
eQTLs among the METS GWAS											
21	rs2236472	45727840	A	0.124	0.83	0.74	0.93	1.6E-03	*COL18A1*	-	Collagen, type XVIII, alpha 1
6	rs4713671	33807877	A	0.191	1.13	1.04	1.24	7.2E-03	*ITPR3*	-	Inositol 1,4,5-triphosphate receptor, type 3
19	rs344802	50496147	T	0.204	0.89	0.81	0.97	1.0E-02	*CKM*	-	Creatine kinase, muscle
18	rs4998986	55282713	A	0.369	0.88	0.82	0.96	2.0E-03	*MALT1*	-	Mucosa-associated lymphoid tissue lymphoma translocation gene 1
18	rs4998985	55282774	A	0.298	0.86	0.79	0.93	3.3E-04	*MALT1*	-	Mucosa-associated lymphoid tissue lymphoma translocation gene 1
10	rs871026	1.31E+08	G	0.384	0.88	0.81	0.95	9.8E-04	*MGMT*	-	O-6-Methylguanine-DNA methyltransferase
Nonsynonymous substitution SNPs among the METS GWAS											
9	rs2296871	1.34E+08	A	0.263	0.89	0.82	0.97	8.6E-03	*MAPK7*	E79G	Mitogen-activated protein kinase 7
1	rs11802875	1.62E+08	A	0.0333	1.35	1.10	1.65	4.0E-03	*NUF2*	S229L	NUF2, NDC80 kinetochore complex component, homolog (*Saccharomyces cerevisiae*)

GWAS, genome-wide association study; eQTL, expression quantitative trait loci; SNP, single-nucleotide polymorphism; LPL, lipoprotein lipase; APOA5, apolipoprotein A-V; CHR, chromosome; BP, base pair; A1, minor allele; MAF, minor allele frequency; OR, odds ratio; METS, metabolic syndrome.

associated with lipid, metabolic syndrome X [23], TG [32], and metabolic traits in East Asians [8], demonstrating that it is putatively functionally associated with METS in the Korean population (Supplementary Table 1).

The second significant SNP was rs6589566, which has 6 high-LD SNPs. Notably, *in silico* annotation of the SNP's

function showed that rs651821 was located in the 5' untranslated region (UTR) of the *APOA5* gene, and also, the SNP was reported as an eQTL of the transgelin (*TAGLN*) gene (Table 4). The results indicate that the remaining SNPs are surrogate markers of rs6589566. Among them, rs651821 and rs964184 are associated with TG level, as one component of METS evaluation, in Chinese populations [26] and in Mexicans [35]. Both SNPs exhibit eQTL of the *TAGLN* gene. *TAGLN* has been documented as a repressive regulator of matrix metalloproteinase 9 (*MMP9*) gene expression [36] and is considered a putative tumor suppressor due to suppression of MMP-9, which harbors tumor metastatic properties [37]. However, MMP-9 is also involved in the progression of METS via chymase activity [38] and has been suggested to be used a diagnostic marker of METS [39]. Thus, it can be explained that the functional eQTL *TAGLN* of rs651821 and rs964184 could be a novel marker for the evaluation of METS in terms of strong regulation of MMP-9.

The third most significant signals were located in the lipoprotein lipase (*LPL*) gene region. Three SNPs (rs10503669, rs17482753, and rs17410962) located in chr 8 were eQTL that contributed to LPL expression in monocytes. LPL is a critical protein of lipid metabolism and is significantly associated with METS in Asian Indians [40], indicating that the functional eQTL-SNP of LPL expression could be a marker for the evaluation of METS in Korean populations. Those three SNPs were in strong LD with rs328, which is a stop-gain mutation of the LPL coding sequence [41] (Supplementary Table 1). Although the remaining 14 SNPs were non-eQTL-SNPs, those SNPs have been reported as being in association with HDL cholesterol, low-density lipoprotein cholesterol, TG, and obesity, indicating that they are putative candidate markers for the evaluation of METS.

Although the remaining 6 SNPs and their nearest genes have not been functionally studied regarding METS-associated traits, further studies are required to elucidate for their role in METS.

Pathway network analysis

The results of the V(G)/V(P) for 100 to 5,000 SNPs were plotted in Fig. 2. When we used 5,000 SNPs, V(G)/V(P) approximated 50%, and we extracted functional SNPs, such as eQTLs and nsSNPs, from the 5,000 SNPs. We could extract 159 eQTLs and 18 nsSNPs among the 5,000 SNPs (Supplementary Table 2). Notably, 6 eQTL genes and 2 nsSNP genes consisted of *LPL* and the apolipoprotein A-V (*APOA5*) pathway through the interaction of a number of mediated genes (Table 4). Among them, muscle creatinine kinase (*CKM*) has been documented to regulate LPL activity [42], demonstrating that it is putatively associated with METS. Those additionally identified genes might be candidate targets of METS for further study (Fig. 3).

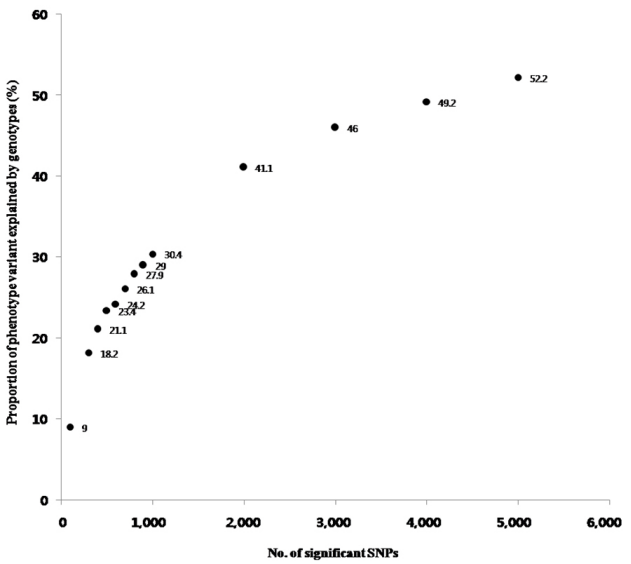

Fig. 2. The proportion of phenotypic variance [V(P)] explained by the genotypic variance [V(G)]. The horizontal axis denotes the number of single nucleotide polymorphisms (SNPs). Approximately 50% of phenotypic variance could be explained by the top 5,000 SNPs.

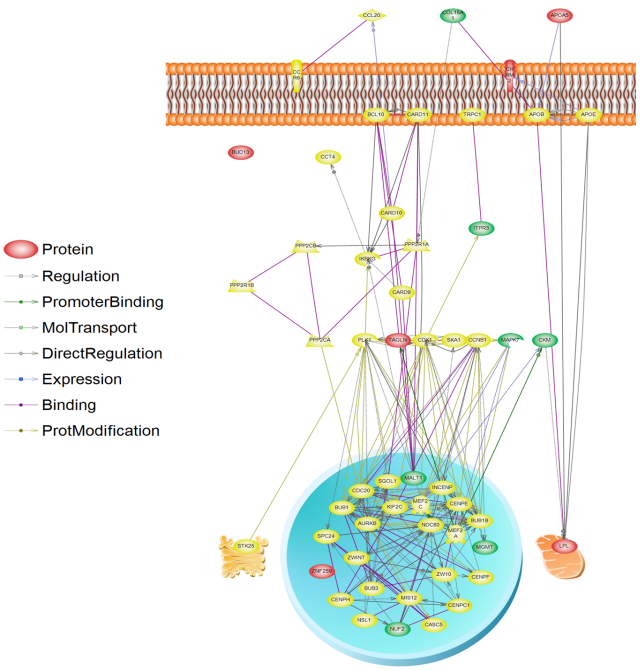

Fig. 3. Illustration of molecular pathway for significantly associated single nucleotide polymorphism (SNP) loci in the metabolic syndrome genome-wide association study (GWAS). The molecules depicted by the significant GWAS loci (reds), functional SNPs (expression quantitative trait loci or nonsynonymous SNPs) loci (green), and the other intermediate molecules (yellow) are illustrated on the cell organelles.

eQTLs and nsSNPs provide insights into the regulation of transcription and aid in the interpretation of GWASs [22]. Most of the eQTL resources are available in online databases, such as RegulomeDB (http://regulome.stanford.edu/), including several published resources in various cell types, such as monocytes [43], human brain [44], lymphoblastoid cell lines [45, 46], and human liver [47]. Probably, RegulomeDB is one of the most useful eQTL databases, because it contains rich information about the products of the ENCODE project, such as transcription factor binding sites, chromatin structure, histone modification, and eQTLs. Our pathway results suggest an internal mechanism of *LPL*, *APOA5*, and muscarinic acetylcholine receptor M2 (*CHRM2*) functions in METS. Therefore, we suggest that 6 eQTLs and 2 nsSNP loci might be additional targets for further association studies and functional analysis.

Conclusively, our approach using the conventional GWAS, reconsidering functional variants and the pathway-based interpretation, suggests a useful method to understand the GWAS results of complex traits and can be expanded in other GWASs.

Supplementary materials

Supplementary data including two tables and one figure can be found with this article online at http://www.genominfo.org/src/sm/gni-12-187-s001.pdf.

Acknowledgments

This work was supported by grants from the Korea Centers for Disease Control and Prevention (KCDC), Republic of Korea (4845-301, 4851-302, 4851-307). This study was also supported by an internal project, "Construction of databases and an analysis system for Korean reference genomes for disease researches" (2013-NG72001-00), of the Korea National Institute of Health, KCDC.

References

1. Alberti KG, Zimmet P, Shaw J; IDF Epidemiology Task Force Consensus Group. The metabolic syndrome: a new worldwide definition. *Lancet* 2005;366:1059-1062.
2. Haffner S, Taegtmeyer H. Epidemic obesity and the metabolic syndrome. *Circulation* 2003;108:1541-1545.
3. Cameron AJ, Shaw JE, Zimmet PZ. The metabolic syndrome: prevalence in worldwide populations. *Endocrinol Metab Clin North Am* 2004;33:351-375.
4. Song J, Kim E, Shin C, Kim SS, Lee HK, Jung M, *et al*. Prevalence of the metabolic syndrome among South Korean adults: the Ansan study. *Diabet Med* 2004;21:1154-1155.
5. Sung J, Lee K, Song YM. Heritabilities of the metabolic syndrome phenotypes and related factors in Korean twins. *J Clin Endocrinol Metab* 2009;94:4946-4952.
6. Pennisi E. Breakthrough of the year. Human genetic variation. *Science* 2007;318:1842-1843.
7. Naidoo N, Pawitan Y, Soong R, Cooper DN, Ku CS. Human genetics and genomics a decade after the release of the draft sequence of the human genome. *Hum Genomics* 2011;5:577-622.
8. Kim YJ, Go MJ, Hu C, Hong CB, Kim YK, Lee JY, *et al*. Large-scale genome-wide association studies in East Asians identify new genetic loci influencing metabolic traits. *Nat Genet* 2011;43:990-995.
9. Eichler EE, Flint J, Gibson G, Kong A, Leal SM, Moore JH, *et al*. Missing heritability and strategies for finding the underlying causes of complex disease. *Nat Rev Genet* 2010;11:446-450.
10. Fransen K, Visschedijk MC, van Sommeren S, Fu JY, Franke L, Festen EA, *et al*. Analysis of SNPs with an effect on gene expression identifies *UBE2L3* and *BCL3* as potential new risk genes for Crohn's disease. *Hum Mol Genet* 2010;19:3482-3488.
11. Yang J, Benyamin B, McEvoy BP, Gordon S, Henders AK, Nyholt DR, *et al*. Common SNPs explain a large proportion of the heritability for human height. *Nat Genet* 2010;42:565-569.
12. Ko KP, Min H, Ahn Y, Park SJ, Kim CS, Park JK, *et al*. A prospective study investigating the association between environmental tobacco smoke exposure and the incidence of type 2 diabetes in never smokers. *Ann Epidemiol* 2011;21:42-47.
13. Cho YS, Go MJ, Kim YJ, Heo JY, Oh JH, Ban HJ, *et al*. A large-scale genome-wide association study of Asian populations uncovers genetic factors influencing eight quantitative traits. *Nat Genet* 2009;41:527-534.
14. Ahn Y, Park SJ, Kwack HK, Kim MK, Ko KP, Kim SS. Rice-eating pattern and the risk of metabolic syndrome especially waist circumference in Korean Genome and Epidemiology Study (KoGES). *BMC Public Health* 2013;13:61.
15. National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III). Third Report of the National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III) final report. *Circulation* 2002;106:3143-3421.
16. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. *Am J Hum Genet* 2011;88:76-82.
17. Rabbee N, Speed TP. A genotype calling algorithm for affymetrix SNP arrays. *Bioinformatics* 2006;22:7-12.
18. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, *et al*. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007;81:559-575.
19. Nikitin A, Egorov S, Daraselia N, Mazo I. Pathway studio: the analysis and navigation of molecular networks. *Bioinformatics* 2003;19:2155-2157.
20. Wallace C, Newhouse SJ, Braund P, Zhang F, Tobin M, Falchi M, *et al*. Genome-wide association study identifies genes for biomarkers of cardiovascular disease: serum urate and dyslipidemia. *Am J Hum Genet* 2008;82:139-149.
21. Heid IM, Boes E, Müller M, Kollerits B, Lamina C, Coassin S,

et al. Genome-wide association analysis of high-density lipoprotein cholesterol in the population-based KORA study sheds new light on intergenic regions. *Circ Cardiovasc Genet* 2008;1:10-20.
22. Liang L, Morar N, Dixon AL, Lathrop GM, Abecasis GR, Moffatt MF, *et al*. A cross-platform analysis of 14,177 expression quantitative trait loci derived from lymphoblastoid cell lines. *Genome Res* 2013;23:716-726.
23. Kraja AT, Vaidya D, Pankow JS, Goodarzi MO, Assimes TL, Kullo IJ, *et al*. A bivariate genome-wide approach to metabolic syndrome: STAMPEED consortium. *Diabetes* 2011;60:1329-1339.
24. Kooner JS, Chambers JC, Aguilar-Salinas CA, Hinds DA, Hyde CL, Warnes GR, *et al*. Genome-wide scan identifies variation in MLXIPL associated with plasma triglycerides. *Nat Genet* 2008;40:149-151.
25. Comuzzie AG, Cole SA, Laston SL, Voruganti VS, Haack K, Gibbs RA, *et al*. Novel genetic loci identified for the pathophysiology of childhood obesity in the Hispanic population. *PLoS One* 2012;7:e51954.
26. Tan A, Sun J, Xia N, Qin X, Hu Y, Zhang S, *et al*. A genome-wide association and gene-environment interaction study for serum triglycerides levels in a healthy Chinese male population. *Hum Mol Genet* 2012;21:1658-1664.
27. Coram MA, Duan Q, Hoffmann TJ, Thornton T, Knowles JW, Johnson NA, *et al*. Genome-wide characterization of shared and distinct genetic components that influence blood lipid levels in ethnically diverse human populations. *Am J Hum Genet* 2013;92:904-916.
28. Sabatti C, Service SK, Hartikainen AL, Pouta A, Ripatti S, Brodsky J, *et al*. Genome-wide association analysis of metabolic traits in a birth cohort from a founder population. *Nat Genet* 2009;41:35-46.
29. Chasman DI, Paré G, Mora S, Hopewell JC, Peloso G, Clarke R, *et al*. Forty-three loci associated with plasma lipoprotein size, concentration, and cholesterol content in genome-wide analysis. *PLoS Genet* 2009;5:e1000730.
30. Teslovich TM, Musunuru K, Smith AV, Edmondson AC, Stylianou IM, Koseki M, *et al*. Biological, clinical and population relevance of 95 loci for blood lipids. *Nature* 2010;466:707-713.
31. Middelberg RP, Ferreira MA, Henders AK, Heath AC, Madden PA, Montgomery GW, *et al*. Genetic variants in *LPL, OASL* and *TOMM40/APOE-C1-C2-C4* genes are associated with multiple cardiovascular-related traits. *BMC Med Genet* 2011;12:123.
32. Kathiresan S, Melander O, Guiducci C, Surti A, Burtt NP, Rieder MJ, *et al*. Six new loci associated with blood low-density lipoprotein cholesterol, high-density lipoprotein cholesterol or triglycerides in humans. *Nat Genet* 2008;40:189-197.
33. Johansen CT, Wang J, Lanktree MB, Cao H, McIntyre AD, Ban MR, *et al*. Excess of rare variants in genes identified by genome-wide association study of hypertriglyceridemia. *Nat Genet* 2010;42:684-687.
34. Kristiansson K, Perola M, Tikkanen E, Kettunen J, Surakka I, Havulinna AS, *et al*. Genome-wide screen for metabolic syndrome susceptibility Loci reveals strong lipid gene contribution but no evidence for common genetic basis for clustering of metabolic syndrome traits. *Circ Cardiovasc Genet* 2012;5:242-249.
35. Weissglas-Volkov D, Aguilar-Salinas CA, Nikkola E, Deere KA, Cruz-Bautista I, Arellano-Campos O, *et al*. Genomic study in Mexicans identifies a new locus for triglycerides and refines European lipid loci. *J Med Genet* 2013;50:298-308.
36. Nair RR, Solway J, Boyd DD. Expression cloning identifies transgelin (SM22) as a novel repressor of 92-kDa type IV collagenase (MMP-9) expression. *J Biol Chem* 2006;281:26424-26436.
37. Assinder SJ, Stanton JA, Prasad PD. Transgelin: an actin-binding protein and tumour suppressor. *Int J Biochem Cell Biol* 2009;41:482-486.
38. Takai S, Jin D, Miyazaki M. Chymase as an important target for preventing complications of metabolic syndrome. *Curr Med Chem* 2010;17:3223-3229.
39. Cicero AF, Derosa G, Manca M, Bove M, Borghi C, Gaddi AV. Vascular remodeling and prothrombotic markers in subjects affected by familial combined hyperlipidemia and/or metabolic syndrome in primary prevention for cardiovascular disease. *Endothelium* 2007;14:193-198.
40. Ranjith N, Pegoraro RJ, Rom L. Lipid profiles and associated gene polymorphisms in young Asian Indian patients with acute myocardial infarction and the metabolic syndrome. *Metab Syndr Relat Disord* 2009;7:571-578.
41. Deo RC, Reich D, Tandon A, Akylbekova E, Patterson N, Waliszewska A, *et al*. Genetic differences between the determinants of lipid profile phenotypes in African and European Americans: the Jackson Heart Study. *PLoS Genet* 2009;5:e1000342.
42. Sattler W, Levak-Frank S, Radner H, Kostner GM, Zechner R. Muscle-specific overexpression of lipoprotein lipase in transgenic mice results in increased alpha-tocopherol levels in skeletal muscle. *Biochem J* 1996;318(Pt 1):15-19.
43. Zeller T, Wild P, Szymczak S, Rotival M, Schillert A, Castagne R, *et al*. Genetics and beyond: the transcriptome of human monocytes and disease susceptibility. *PLoS One* 2010;5:e10693.
44. Gibbs JR, van der Brug MP, Hernandez DG, Traynor BJ, Nalls MA, Lai SL, *et al*. Abundant quantitative trait loci exist for DNA methylation and gene expression in human brain. *PLoS Genet* 2010;6:e1000952.
45. Veyrieras JB, Kudaravalli S, Kim SY, Dermitzakis ET, Gilad Y, Stephens M, *et al*. High-resolution mapping of expression-QTLs yields insight into human gene regulation. *PLoS Genet* 2008;4:e1000214.
46. Montgomery SB, Sammeth M, Gutierrez-Arcelus M, Lach RP, Ingle C, Nisbett J, *et al*. Transcriptome genetics using second generation sequencing in a Caucasian population. *Nature* 2010;464:773-777.
47. Schadt EE, Molony C, Chudin E, Hao K, Yang X, Lum PY, *et al*. Mapping the genetic architecture of gene expression in human liver. *PLoS Biol* 2008;6:e107.