

Combined Genome-Wide Linkage and Association Analyses of Fasting Glucose Level in Healthy Twins and Families of Korea

Young Ju Suh,¹ SungHwan Kim,²
So Hun Kim,³ Jia Park,⁴ Hyun Ae Lim,⁴
Hyun Ju Park,⁴ Hangseok Choi,⁵
Daniel Ng,⁶ Mi Kyeong Lee,⁷
and Moonsuk Nam^{3,4}

¹Institute of Clinical Research, Inha University School of Medicine, Incheon, Korea; ²Department of Biostatistics, University of Pittsburg, PA, USA; ³Department of Internal Medicine, Inha University School of Medicine, Incheon; ⁴Clinical Trial Center, Inha University Hospital, Incheon; ⁵National Health Insurance Corporation, Seoul, Korea; ⁶Department of Epidemiology and Public Health, Yong Loo Lin School of Medicine, National University of Singapore, Singapore; ⁷Department of Epidemiology and Institute of Environment and Health, Seoul National University School of Public Health, Seoul, Korea

Received: 6 July 2012
Accepted: 4 January 2013

Address for Correspondence:
Moonsuk Nam, MD

Division of Endocrinology and Metabolism, Department of Internal Medicine, Inha University School of Medicine, 27 Inhang-ro, Jung-gu, Incheon 400-711, Korea
Tel: +82.32-890-1101, Fax: +82.32-890-1120
E-mail: namms@inha.ac.kr

This work was supported by the Inha University Research grant (INHA-42825), by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (grant no. 43466), by the Korea Healthcare Technology R&D Project Ministry of Health and Welfare (A070001 and A102065), by Korea Centers for Disease Control and Prevention (4845-301, 4851-302), and by an intramural grant from the Korea National Institute of Health (2009-E73001-00), the Republic of Korea.

INTRODUCTION

Type 2 diabetes is a chronic disease that has a major impact on public health and is a major cause of morbidity and mortality. However, the etiology of type 2 diabetes is not wholly understood, because type 2 diabetes is a complex disease that is influenced by many risk factors (1, 2). Fasting glucose (FG) level is an important type 2 diabetes related trait. A chronically elevated FG level reflects a damaged glucose metabolism and is a crucial risk marker for type 2 diabetes (3).

Furthermore, the heritability of FG level has been reported to be in 17%-51% in the general population (4). Therefore, knowledge of genetic susceptibility for FG level would improve our

This study was undertaken to identify genetic polymorphisms that are associated with the risk of an elevated fasting glucose (FG) level using genome-wide analyses. We explored a quantitative trait locus (QTL) for FG level in a genome-wide study from a Korean twin-family cohort (the Healthy Twin Study) using a combined linkage and family-based association analysis approach. We investigated 1,754 individuals, which included 432 families and 219 pairs of monozygotic twins. Regions of chromosomes 2q23.3-2q31.1, 15q26.1-15q26.3, 16p12.1, and 20p13-20p12.2, were found to show evidence of linkage with FG level, and several markers in these regions were found to be significantly associated with FG level using family-based or general association tests. In particular, a single-nucleotide polymorphism (rs6138953) on the *PTPRA* gene in the 20p13 region (combined $P = 1.8 \times 10^{-6}$) was found to be associated with FG level, and the *PRKCB1* gene (in 16p12.1) to be possibly associated with FG level. In conclusion, multiple regions of chromosomes 2q23.3-2q31.1, 15q26.1-15q26.3, 16p12.1, and 20p13-20p12.2 are associated with FG level in our Korean twin-family cohort. The combined approach of genome-wide linkage and family-based association analysis is useful to identify novel or known genetic regions concerning FG level in a family cohort study.

Key Words: Fasting Glucose Level; Genome-Wide; Genetic Linkage; *PRKCB1*; *PTPRA*

understanding of the risks of an impaired FG level and of type 2 diabetes.

Linkage and association analyses are the methods generally used to map genes that confer to susceptibility to complex diseases, such as, type 2 diabetes (5). Linkage analysis is performed on the joint distributions of trait and marker alleles in family members to determine whether a major gene associated with a familial trait is linked to a well-defined marker. Conversely, association analysis is useful for verifying susceptibility genes, in which the distributions of alleles at polymorphic markers are compared. Association studies are valuable for localizing regions of interest identified by linkage analysis. Recently, many researchers have performed linkage and association studies for

complex disease using high throughput screening tools, and single-nucleotide polymorphism (SNP) arrays provide one of the most powerful high throughput screening approaches to genome-wide studies (6).

Adding twins can enrich a family study. Dizygotic (DZ) twins are valuable to gene mapping studies as ideal sib-pairs (7). Although monozygotic (MZ) twins do not contribute to gene mapping per se, the presence of them provides more precise discrimination between polygenic and shared environmental variance, which can potentially increase power of family-based regression methods (8).

Aiming at identifying novel or known genetic variations concerning FG level, we analyzed associations between genetic polymorphisms and FG level using a genome-wide SNP array chip in this study. In addition, we explored the trait of FG level in a healthy, family cohort containing MZ twins using a combined linkage/family-based association analysis approach.

MATERIALS AND METHODS

Population and data collection

In the present study, we used the Korean Healthy Twin study cohort data obtained from the Center for Genome Science at the Korean National Institute of Health. The aim of the Korean Healthy Twin cohort study was to explore genetic variation related to complex traits and to study the effects of environment on the etiologies of complex diseases, in adult same-sex twin pairs aged 30-yr or older and their first-degree family members regardless of family size. Participants were recruited by placing advertisements in daily newspaper and by using a poster campaign in about 300 hospitals and governmental agencies. The study design used in the Korean Healthy Twin study cohort has been previously described in detail (9). Of the 2,124 subjects in the cohort, we investigated 1,754 individuals, which included 432 families and 219 pairs of MZ twins (female and male pairs) after excluding individuals without an FG level. Participants in the Healthy Twin Study that signed a consent form before enrollment were interviewed by trained interviewers at each institution, and general epidemiological and clinical data were recorded. Subject weights and heights were also recorded, and blood samples were taken using standard protocols. Plasma glucose was assayed using a hexokinase enzymatic method.

Genotyping analysis

Genotyped data were obtained from the Center for Genome Science in the Korean Centers for Disease Control and Prevention. Sixteen short tandem repeat markers (15 autosomal short tandem repeat markers plus one sex-determining marker) were used to confirm sample identities for the twins (AmpFISTR Identifier Kit; PerkinElmer, Norwalk, CT, USA). Only if at least 15 of the 16 markers corresponded in each twin pair, was the pair

considered MZ (9). Microarray-based SNP genotyping was performed using the genome-wide human SNP array 6.0 chip (Affymetrix, Santa Clara, CA, USA). All genomic DNA samples were normalized to a concentration of 50 ng/ μ L. The Affymetrix standard protocol was followed. Genomic DNA was digested using restriction enzymes such as *Nsp*I and *Sty*I and amplified by PCR. We labeled the amplified genomic DNA using Streptavidin, fragmented it to be 50-200 bp, and hybridized it in the chip. Scanning was performed using a high-resolution Affymetrix GeneChip scanner 3000 7G. Image files were used to transfer the data into GCOS 1.4 for subsequent processing using G-TYPE 4.0 software. Using V3 annotation for the genome-wide human SNP array 6.0 chip, a total of 906,600 SNPs and 946,000 copy number probes were genotyped per sample. Genotype calls were generated by a proprietary Birdseed 2.0 algorithm. Genotyped SNPs with a call rate of less than 95% were dropped. Monomorphic SNPs, SNPs with a minor allele frequency of < 0.01 , or SNPs out of Hardy-Weinberg Equilibrium ($P < 0.001$) were filtered out. Finally, 520,484 SNPs were subjected to further analyses.

Statistical analysis

We examined the relationships between FG levels and covariates, such as, age, sex, and body mass index (BMI), using Spearman correlation coefficient test or Wilcoxon rank sum test (Mann-Whitney U test). Statistical analysis was performed using SAS version 9.1 (SAS Institute, Cary, NC, USA). Familial correlations of FG level between possible pairs in the pedigree and heritability of FG level were obtained using S.A.G.E. software, version 6.0.1 (<http://darwin.cwru.edu/>).

Combined genome-wide linkage analysis with peak wise association tests was performed to identify genetic markers of FG level. During the genome-wide linkage scan, multivariate regression-based quantitative trait loci (QTL) analysis of the number of alleles identical by descent (IBD) at a given marker was performed on the squared sum and squared difference of FG level after adjusting for confounders, such as, age, sex, and BMI (10). For adjusted FG level, we used regression residuals of $\log(\text{FG}_{mi})$ on age, sex, and BMI in family members i ($i = 1, 2, \dots$) within the m th family. We considered $S = [S_{ij} = (X_i + X_j)^2]$ and $D = [D_{ij} = (X_i - X_j)^2]$ as predictor variables, and $\Pi = [(\hat{\pi}_{ij})]$ for the estimated proportion of alleles IBD $\hat{\pi}_{ij}$ ($i \neq j$) as dependent variables in the multivariate regression model. MERLIN-REGRESS software (8) was used for extended regression based QTL analysis. For this analysis, we controlled overestimated genetic variance in the model caused by MZ twins sharing identical genotypes within pairs.

We next performed association studies for SNPs in linkage regions using a LOD score ≥ 1.3 in the genome-wide linkage analysis. We considered the T statistic of 2.45, which corresponds to a LOD score of 1.3, since $4.6 \times \text{LOD score}$ is distributed as a 50-50 mixture of χ^2 with 1 df and a point mass at 0 under the

null hypotheses of no linkage. We used two different association analysis approaches, that is, the family-based association test (FBAT) in all families and population-based tests of association in founders of each family, after adjusting for age, sex, and BMI to identify specific genetic loci associated with the risk of an increased FG level. Using FBAT, we merged each MZ twin's pair as one subject by averaging traits, such as, BMI and FG level, of each MZ twin pair to adjust for possible overestimation of genetic variance in the model. FBAT under the additive model was used to allow both variations in a quantitative trait $\log(\text{FG})$ and in covariates into a general score test of association in 1,535 subjects using FBAT software, version 2.0.3 (<http://www.biostat.harvard.edu/~fbat/>).

Population-based association tests under various modes of inheritance, including additive, dominant and recessive models, and regression analyses of the $\log(\text{FG})$ on interactive effects of significant SNPs were performed in 467 unrelated founders of each family using PLINK software, version 1.07 (<http://pngu.mgh.harvard.edu/purcell/plink/>). Of the additive, dominant, and recessive models for each SNP-trait association, we selected the best genetic model based on Akaike's information criterion (AIC). To combine two independent P values derived by FBAT and association analysis, we used Fisher's method for combining P values (11).

Ethics statement

The study protocol of the healthy twin study was approved by the institutional review board of the Seoul National University (IRB No. 144-2011-07-11). Informed consent was submitted by all participants.

RESULTS

Distribution of subject characteristics

Table 1 shows the distribution of subject characteristics for all families and founders of each family. Founders tended to be older and had a higher FG level than all families. Comparisons of FG levels in unrelated population of founders showed they were different by sex ($P = 0.028$ by Wilcoxon's rank sum test), and significantly correlated with age and BMI ($r = 0.132$, $P = 0.004$ for age and $r = 0.156$, $P < 0.001$ for BMI by Spearman correlation coefficient analysis). The proportions of study population with a history of type 2 diabetes were 5.3% for all families and 12.6% for founders.

Familial correlations of FG level between two random variables within nuclear family are shown in Table 2. Correlation estimate between MZ twins was moderately significant with $r = 0.528$ ($SE = 0.049$, $P < 0.001$), and was higher than that between pairs of full siblings ($r = 0.257$, $SE = 0.034$). No large differences were found between correlation estimates for all family members and families without one twin member in each pair of MZ twins. Heritability estimates for FG level were $h^2 = 0.27$ ($SE = 0.07$, $P < 0.001$) and adjusted $h^2 = 0.47$ ($SE = 0.06$, $P < 0.001$) (adjusting for age, sex and BMI). Without one member of twins in each pair of MZ twins, we obtained heritability estimates for FG level of $h^2 = 0.31$ ($SE = 0.07$, $P < 0.001$) and adjusted $h^2 = 0.49$ ($SE = 0.06$, $P < 0.001$) (adjusted for age, sex, and BMI).

Genome-wide linkage analysis

Initially, we performed an autosomal genome-wide linkage study for $\log(\text{FG})$, after adjusting for age, sex, and BMI, using multi-

Table 1. Distribution of study population characteristics

Parameters	All families		Founders only		P value
	No. (%) or Median (IQR)*	FG level [†] Median (IQR)	No. (%) or Median (IQR)*	FG level [†] Median (IQR)	
Age (yr)	40 (35-55)		60 (60-65)		0.004 [‡]
Sex					
Male	683 (38.9%)	94 (88-101)	176 (37.7%)	95 (89-106)	0.028 [§]
Female	1071 (61.1%)	89 (84-94)	291 (62.3%)	93 (87-100)	
BMI (kg/m ²)	23.6 (21.4-25.8)		24.7 (22.5-26.7)		< 0.001 [‡]
History of type 2 diabetes					
No	1659 (94.8%)	90 (84-95)	408 (87.4%)	92.5 (87-99)	< 0.001 [§]
Yes	92 (5.3%)	123 (104.5-145)	59 (12.6%)	115 (104-140)	
Total (No.)	1,754	90 (85-96)	467	94 (88-101)	

*No. (%) for categorical variables and Median (interquartile range) for continuous variables; [†]FG level: fasting glucose level (mg/dL); [‡] P value as tested for Spearman correlation coefficient between the FG level and each variable; [§] P value as obtained by Wilcoxon rank sum test (Mann-Whitney U test) to test a difference of the FG level between two groups.

Table 2. Familial correlations of FG level between two random variables regarding each component of family members

Relationships	All family members			Family members without one twin member in each MZ twins pair		
	No. of pairs	Correlation \pm SE	P value	No. of pairs	Correlation \pm SE	P value
Parent-offspring	1,327	0.193 \pm 0.034	< 0.001	1,068	0.200 \pm 0.035	< 0.001
Full siblings	1,779	0.257 \pm 0.034	< 0.001	1,217	0.219 \pm 0.039	< 0.001
MZ Twins	219	0.528 \pm 0.049	< 0.001			

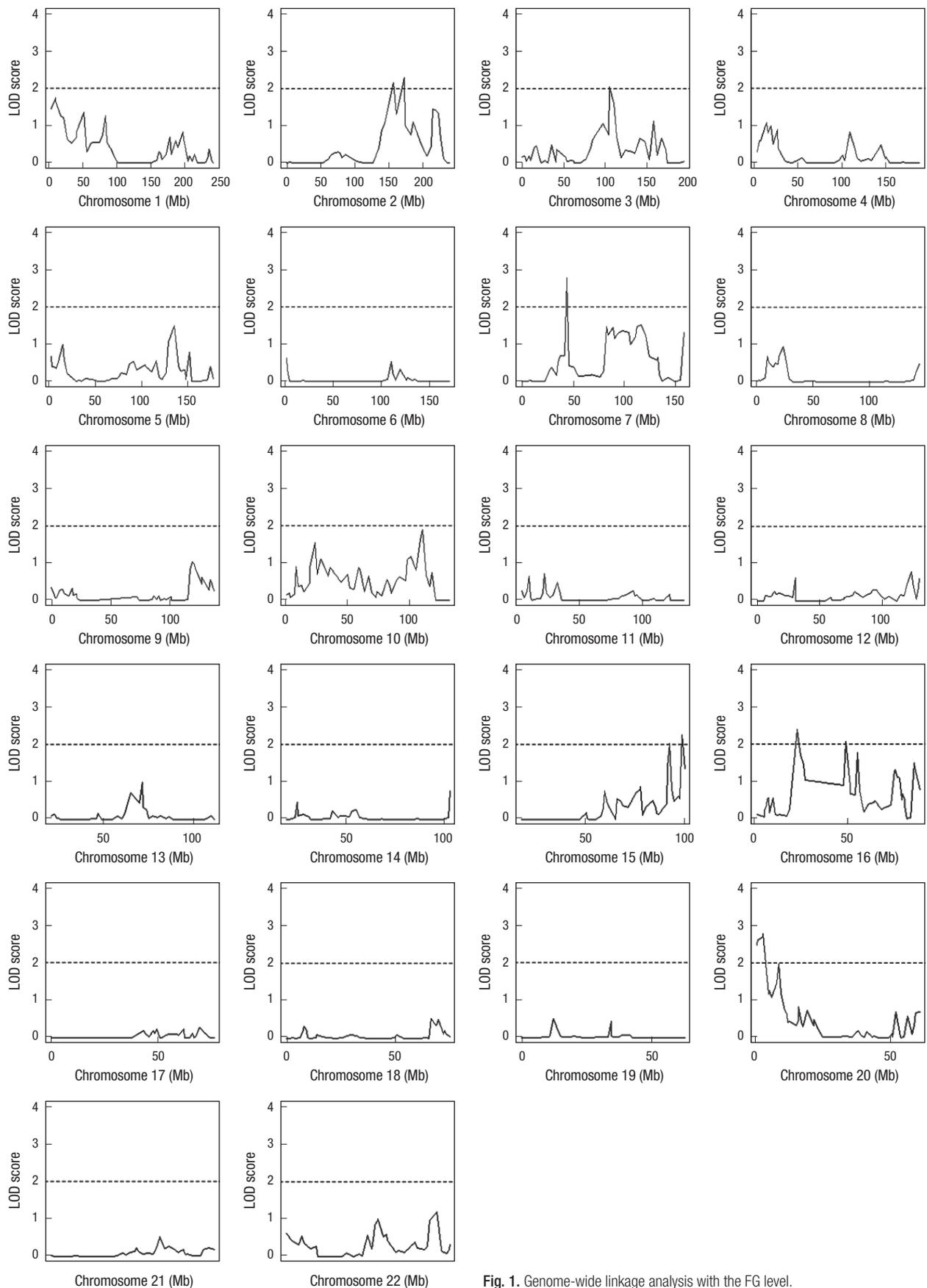


Fig. 1. Genome-wide linkage analysis with the FG level.

Table 3. The regions for evidence of linkage for the FG level

Cytoband	Region (Mb)	Max LOD*	Position (bp)	Near SNP	Near gene [†]
2q23.3-2q31.1	151.8-172.1	2.31	172,072,250	rs17287624	<i>C2orf37</i>
3q13.11-3q13.13	105.2-110.6	2.05	105,215,684	rs1900677	
7p13	43.5	2.77	43,515,792	rs11760314	<i>HECW1</i>
15q26.1-15q26.3	91.1-99.9	2.26	98,347,739	rs12916269	<i>ADAMTS17</i>
16p12.1	22.8-26.5	2.39	22,838,204	rs13329797	<i>USP31</i>
16q12.1-16q13	48.8-54.9	2.08	48,775,110	rs7192973	<i>PAPD5</i>
20p13-20p12.2	0.0-9.2	2.77	2,970,618	rs6138990	<i>PTPRA</i>

*The peak LOD score within the corresponding linkage region; [†]The nearest gene to the corresponding marker.

Table 4. Linkage regions which include known markers with significant evidence of association or linkage

Significant linkage region			Known markers associated with FG level or T2D					
Cytoband	Region (Mb)	Max	Marker*	Position (bp)	Near gene	Trait [†]	Test [‡]	Reference
2q23.3-2q31.1	151.8-172.1	2.33	rs10181181	160,795,657		T2D	GWAS	(33)
			rs2925757	160,809,415		T2D	GWAS	(33)
			rs12692588	160,832,428		T2D	GWAS	(33)
			rs7572970	160,844,902	<i>RBMS1</i>	T2D	GWAS	(33)
			rs1020731	160,852,301	<i>RBMS1</i>	T2D	GWAS	(33)
			rs1020732	160,852,485	<i>RBMS1</i>	T2D	GWAS	(33)
			rs12692592	160,871,627	<i>RBMS1</i>	T2D	GWAS	(33)
			rs4077463	160,874,480	<i>RBMS1</i>	T2D	GWAS	(33)
			rs7593730	160,879,700	<i>RBMS1</i>	T2D	GWAS	(33)
			rs4589705	160,884,382	<i>RBMS1</i>	T2D	GWAS	(33)
			rs1882347	164,116,560		FG level	Linkage (LOD = 2.2)	(12)
			D2S2330	166,405,837		T2D	Linkage (LOD = 1.2)	(16)
			D2S2345	168,428,635	<i>B3GALT1</i>	T2D	Linkage (LOD = 3.0)	(5, 13)
			rs560887	169,471,394	<i>G6PC2</i>	FG level	GWAS/Rep	(34-36)
			rs563694	169,482,317	<i>G6PC2/ABCB11</i>	FG level	GWAS	(37)
			rs575671	169,489,064	<i>ABCB11</i>	FG level	GWAS/Rep	(36)
rs483234	169,490,820	<i>ABCB11</i>	FG level	Rep.Assoc	(34)			
rs3755157	169,500,417	<i>ABCB11</i>	FG level	Rep.Assoc	(34)			
rs853787	169,510,498	<i>ABCB11</i>	FG level	GWAS	(35)			
rs853778	169,519,470	<i>ABCB11</i>	FG level	Rep.Assoc	(34)			
15q26.1-15q26.3	91.1-99.9	2.26	D15S130	92,512,350		Glucose 120	Linkage (LOD = 1.3)	(14)
			rs11073418	93,448,799		T2D	GWAS	(38)
			D15S1014	95,803,646		Glucose 120	Linkage (LOD = 1.3)	(14)
16p12.1	22.8-26.5	2.18	rs3760106	23,753,297	<i>PRKCB1</i>	T2D	Rep.Assoc	(39)
			rs2575390	23,754,255	<i>PRKCB1</i>	T2D	Rep.Assoc	(39)
			5'-upstream		<i>PRKCB1</i>	T2D	Rep.Assoc	(40)
			D16S420	24,143,752		T2D	Linkage (NPL = 1.7)	(15)
			D16S3068	25,468,276		T2D	Linkage (LOD = 1.4)	(16)
16q12.1-16q13	48.8-54.9	2.08	rs8050136	52,373,776	<i>FTO</i>	T2D	GWAS	(6, 19, 20)
			rs9939609	52,378,028		T2D	GWAS	(19)
20p13-20p12.2	0.0-9.2	2.78	rs7361808	1,784,935	<i>SIRPA</i>	T2D	GWAS	(19)
			D20S905	5,811,664		T2D	Linkage (LOD = 3.1)	(17, 18)
			rs958435	6,524,785		T2D	Assoc	(20)
			D20S892	6,698,245	<i>BMP2</i>	T2D	Assoc	(18)
			D20S115	7,608,052		T2D	Linkage (LOD = 1.7)	(17), (16)
			rs6140410	7,685,540		T2D	Assoc	(20)

*The peak LOD score within the corresponding linkage region; [†]Trait tested for the corresponding study. T2D, Type 2 diabetes; FG level, fasting glucose level; glucose 120, glucose level obtained at 120 min interval; [‡]Statistical method to map genes. GWAS, genome-wide association study; Linkage, linkage analysis; Rep.assoc, replicated association study.

variate regression-based QTL analysis (10). We detected overall 7 peak regions in chromosomes 2, 3, 7, 15, 16, and 20 with significant evidence of linkage with FG level by genome-wide linkage analysis (Fig. 1). Top linkage peaks with LODs of > 2 in each region were located in chromosome 2q31.1 (172.1 Mb), 3q13.11 (105.2 Mb), 7p13 (43.5 Mb), 15q26.3 (98.3 Mb), 16p12.1 (22.8 Mb), 16q12.1 (48.8 Mb), and 20p13 (3.0 Mb), as shown in Table

3. Subsequently, we focused on SNPs in the 6 linkage regions of chromosomes 2q23.3-2q31.1, 3q13.11-3q13.13, 15q26.1-15q26.3, 16p12.1, 16q12.1-16q13, and 20p13-20p12.2 with relatively wide regions showing linkage (except for 7p13 in which only a small region showed linkage). Among these regions, we were able to find several well-known markers with significant evidence of linkage or association with FG level or risk of diabetes in the fol-

Table 5. Significant SNPs* detected by FBAT or association analyses within the region for evidence of linkage

Cytoband	SNPs in the linkage region					FBAT		Association with founders-only			Combined P ^s
	Position(bp)	SNP rs*	Near gene	LOD [†]	MAF	U [‡]	P value [§]	β	P value [§]	Best model	
2q24.3	167,709,716	rs16853036	<i>XIRP2</i>	1.789	0.052	0.140	0.821	0.713	< 0.001	rec	< 0.001
2q31.1	171,404,718	rs7561581	<i>GAD1</i>	1.592	0.455	1.172	0.329	0.064	< 0.001	dom	0.002
15q26.3	98,393,341	rs2727184	<i>ADAMTS17</i>	1.809	0.131	0.640	0.557	0.204	< 0.001	rec	0.001
16p12.1	22,846,113	rs6497610		2.242	0.386	0.251	0.862	-0.071	< 0.001	dom	< 0.001
	23,982,020	rs11640219	<i>PRKCB1</i>	1.757	0.047	0.088	0.912	0.578	< 0.001	rec	0.005
	23,994,362	rs9923820	<i>PRKCB1</i>	1.745	0.048	0.140	0.859	0.578	< 0.001	rec	0.005
20p13	2,252,832	rs2014045	<i>TGM3</i>	1.956	0.138	0.434	0.617	0.209	< 0.001	rec	0.002
	2,532,331	rs16987508	<i>TMC2</i>	2.233	0.314	1.053	0.426	0.093	< 0.001	rec	0.002
	2,774,321	rs6132978	<i>VPS16</i>	2.644	0.078	2.235	< 0.001	0.233	0.018	rec	< 0.001
20p12.3	2,858,880	rs6138953	<i>PTPRA</i>	2.669	0.071	1.938	0.003	0.712	< 0.001	rec	1.8E-06
	8,868,528	rs11906576		1.511	0.123	-3.537	< 0.001	0.002	0.984	rec	0.005
	8,893,635	rs6516421		1.519	0.053	-2.718	< 0.001	-0.015	0.629	dom	0.002

*Significant SNPs ($P < 0.001$) were selected from the FBAT or the association study among the SNPs with $\text{LOD} \geq 1.3$ within the regions with evidence of linkage; [†]The LOD score for the corresponding marker; [‡]The general FBAT statistic; [§] P values were calculated after adjusting for age, sex and BMI. Combined P value for each SNP was derived using Fisher's method; ^{||}The best genetic model with the lowest AIC was selected among the dominant (dom), recessive (rec), and additive (add) models.

lowing regions of chromosomes 2q23.3-2q31.1, 15q26.1-15q26.3, 16p12.1, 16q12.1-16q13, and 20p13-20p12.2 (Table 4).

Association analyses for SNPs in the linkage peaks

We then performed two different association studies on SNP sites with an LOD score ≥ 1.3 within the significant linkage regions described above. Using FBAT in all families, we identified 3 significant SNPs within 6 linkage regions associated with FG level after adjusting for age, sex, and BMI ($P < 0.001$, Table 5). We also performed population-based association analyses in founders only for each SNP with respect to FG level after adjusting for age, sex, and BMI for the various modes of inheritance. We confirmed 9 significant SNP sites using the best genetic model for each SNP within 6 linkage regions ($P < 0.001$, Table 5). By combining the two different P values obtained by FBAT and association analysis, we were able to identify 4 significant SNPs, namely, rs16853036 located in the *XIRP2* gene (xin actin-binding repeat containing 2 in chromosome 2q24.3), rs6497610 (in chromosome 16p12.1), rs6132978 in the *VPS16* gene (vacuolar protein sorting 16 homolog in the chromosome 20p13), and rs6132978 in the *PTPRA* gene (protein tyrosine phosphatase, receptor type, A in the chromosome 20p13) that were significantly associated with FG level with a combined $P < 0.001$ (Table 5). In particular, rs6138953 located in the *PTPRA* gene was found to be highly associated with FG level (combined $P = 1.8 \times 10^{-6}$) under the recessive model.

DISCUSSION

FG level is a typical complex trait related to type 2 diabetes that is affected by multiple genes. FG level does not follow any mode of Mendelian inheritance and its phenotype is widely variable. In this study, we used two ways to map genes that confer susceptibility to an elevated FG level, that is, genome-wide linkage

and association analyses. Initially, we performed genome-wide linkage scan analysis using the Affymetrix SNP 6.0 chip with the objective of identifying those markers which are linked to major genes for FG level in all family members. We then used two different types of association analyses to identify susceptibility genes with the FG level and to localize linkage regions of concern in genes identified by genome-wide linkage scan analysis.

We found multiple regions with evidence of linkage that were associated with the risk of an elevated FG level in our healthy twin-family cohort. Based on our combined genome-wide linkage and association analyses, we were able to confirm several regions that have evidence of linkage with FG level and type 2 diabetes in different populations. In particular, regions of chromosomes 2q23.3-2q31.1, 15q26.1-15q26.3, 16p12.1, and 20p13-20p12.2 were consistently found to show evidence of linkage with FG level in different populations (Table 4). Meigs et al. (12) found that chromosome 2q24.3 (rs1882347) have evidence of linkage ($\text{maxLOD} = 2.2$) with FG level in the Framingham Heart Study cohort. Furthermore, a region of D2S2345 located in 2q24.3 showed evidence of linkage with an $\text{LOD} = 3.0$ in French Europeans in studies by Busfield et al. (13) and Huang et al. (5). However, chromosome 15q26.2 region showed relatively weak evidence of linkage with glucose level ($\text{LOD} = 1.3$) in Amish families (14). Chromosome 16p12.1 also showed weak evidence of linkage with type 2 diabetes ($\text{LOD} < 2.0$) in Finnish families (15) and French Europeans (16). The 20p13-20p12.2 region, which showed evidence of linkage and association with FG level in the present study, has been previously reported to show evidence of linkage (16-18) and association (18-20) with type 2 diabetes in different populations. Furthermore, D20S905 region (located in 20p12.3) showed peak evidence of linkage with type 2 diabetes with an $\text{LOD} > 3$ in the FUSION (non-insulin-dependent diabetes mellitus genetics) study in Finnish families (17, 18).

The findings of this study were mostly novel, however, of the

identified genes (Table 5), *PTPRA* and *PRKCB1* are likely to have direct/indirect impacts on insulin signaling and FG levels. The rs6138953 SNP in the intron region of *PTPRA*, located in 20p13, was found to be significantly associated with FG level under the recessive model in the present study. The protein encoded by the *PTPRA* gene is protein tyrosine phosphatase receptor type A (*PTPRA*), which is a member of the protein tyrosine phosphatase (PTP) family. *PTPRA* may play a role as a negative regulator of insulin signaling, and has been shown to dephosphorylate the insulin receptor, inhibit insulin-mediated effects, and modify insulin secretion in vitro (21). The expression of *PTPRA* is up-regulated in the skeletal muscles of obese humans (22), but unlike protein tyrosine phosphatase-1B (*PTP1B*) or leukocyte antigen related (*LAR*), there is insufficient evidence to prove its importance in insulin signaling. The present finding may support the suggestion that *PTPRA* participates in the pathogenesis of impaired glucose homeostasis in humans due to altered insulin signaling. In addition to its association with insulin signaling, *PTPRA* has been actively studied in relation to cancer. Protein tyrosine phosphorylation, which is controlled by protein tyrosine kinase (PTKs) and PTPs, play an important role in various cellular events, such as, cell growth, differentiation, and oncogenesis (23), and the oncogenic function of various PTPs (*PTPN11*, *PTPRH*, *PTPRF*, and other) have been well demonstrated (24). However, concerning *PTPRA*, the function in human tumor biology has not been clearly verified, because in vitro (25) and clinical research (26) results in various cancers (colon, breast, gastric cancer, and other) are inconsistent. Therefore, further studies on the potential oncogenic role of *PTPRA* needed to be implemented.

The mutant types of rs11640219 and rs9923820, which are located in the intron region of the *PRKCB1* gene (in 16p12.1), showed weak associations with FG level elevation according to the recessive model in the founder population in the present study. The protein encoded by *PRKCB1* is protein kinase C (PKC) β ($\beta 1$ and $\beta 2$), which has been reported to be involved in glucose homeostasis and to play an important role in the pathogenesis of diabetic complications (27). Furthermore, the overexpressions of *PKC* $\beta 1$ and $\beta 2$ was found to inhibit the tyrosine kinase activity of the insulin receptor in vitro (28), and *PKC* β knockout mice showed mildly improved glucose homeostasis and increased glucose transport in muscle and adipose tissue (29). In another study, *PKC* $\beta 2$ overexpression was found to induce c-myc expression and to suppress insulin gene transcription in rat pancreatic islets (30). On the other hand, *PKC* $\beta 2$ was required to ensure adequate insulin-dependent glucose uptake in rat skeletal muscle (31). In another study by (32) a functional *PKC* β promoter polymorphism at position -546, which reduced promoter activity, was found to be associated with insulin resistance in human subjects.

This study has several strengths. First, our data included vari-

ous familial relationships in a Korean twin-family cohort, which enable us to identify evidence of genetic signals more accurately. Second, to the best of our knowledge, this is the first study to report linkages and associations of polymorphisms in chromosomes 16p12.1 and 20p13-20p12.2 with respect to FG level elevation. Concerning the 16p12.1 and 20p13-20p12.2 regions, little linkage and association analyses has been performed with respect to fasting glucose, insulin, and insulin resistance, because most research has been focused on type 2 diabetes and its complications. However, we found that the *PTPRA* and *PRKCB1* genes among 11 selected genes (Table 5) are involved in glucose metabolism and insulin secretion directly or indirectly, and thus, may affect FG levels.

However, this study is limited by a lack of extensive fine-mapping and reliable replications of FG levels in individuals. Unclear biological understanding stayed behind for the other significantly detected polymorphisms, and thus, further studies are required on these polymorphisms. This study also has a limited power to identify well-known genetic variants involved in FG level, due to small sample size. Also, the absence of information on specific diabetic medications for the subjects who had diabetes is a limitation of our study.

In addition, meta-analyses of linkage regions and more extensive fine-mapping experiments that more broadly probe regions believed to contain additional markers associated with FG levels are needed in our future studies. Furthermore, replication studies are needed to explore the association between confirmed susceptibility loci and environmental factors of the trait.

Summarizing, we found multiple regions that showed evidence of linkage and an association with FG level in our Korean twin-family cohort. This study identifies that chromosomes 2q23.3-2q31.1, 15q26.1-15q26.3, 16p12.1, and 20p13-20p12.2 show evidence of linkage with FG level. In particular, SNP (rs-6138953) on *PTPRA* in the 20p13 region is associated with FG level, and the *PRKCB1* gene (in 16p12.1) is possibly associated with FG level. These findings contribute to our understanding of the pathogenesis of type 2 diabetes in the Korean population. The combined approach of genome-wide linkage and family-based association analysis is useful to identify novel or known genetic regions concerning FG level in a family cohort study.

ACKNOWLEDGMENTS

The authors have no conflicts of interest to disclose.

REFERENCES

- Ohlson LO, Larsson B, Björntorp P, Eriksson H, Svärdsudd K, Welin L, Tibblin G, Wilhelmsen L. Risk factors for type 2 (non-insulin-dependent) diabetes mellitus: thirteen and one-half years of follow-up of the participants in a study of Swedish men born in 1913. *Diabetologia* 1988; 31:

- 798-805.
2. Manson JE, Rimm EB, Stampfer MJ, Colditz GA, Willett WC, Krolewski AS, Rosner B, Hennekens CH, Speizer FE. *Physical activity and incidence of non-insulin-dependent diabetes mellitus in women. Lancet* 1991; 338: 774-8.
 3. Kim DJ, Cho NH, Noh JH, Kim HJ, Choi YH, Jung JH, Min YK, Lee MS, Lee MK, Kim KW. *Fasting plasma glucose cutoff value for the prediction of future diabetes development: a study of Middle-Aged Koreans in a Health Promotion Center. J Korean Med Sci* 2005; 20: 562-5.
 4. Fradin D, Heath S, Lathrop M, Bougnères P. *Quantitative trait loci for fasting glucose in young Europeans replicate previous findings for type 2 diabetes in 2q23-24 and other locations. Diabetes* 2007; 56: 1742-5.
 5. Huang QY, Cheng MR, Ji SL. *Linkage and association studies of the susceptibility genes for type 2 diabetes. Yi Chuan Xue Bao* 2006; 33: 573-89.
 6. Frayling TM. *Genome-wide association studies provide new insights into type 2 diabetes aetiology. Nat Rev Genet* 2007; 8: 657-62.
 7. Martin N, Boomsma D, Machin G. *A twin-pronged attack on complex traits. Nat Genet* 1997; 17: 387-92.
 8. Abecasis GR, Cherny SS, Cookson WO, Cardon LR. *Merlin: rapid analysis of dense genetic maps using sparse gene flow trees. Nat Genet* 2002; 30: 97-101.
 9. Sung J, Lee K, Song YM. *Heritabilities of the metabolic syndrome phenotypes and related factors in Korean twins. J Clin Endocrinol Metab* 2009; 94: 4946-52.
 10. Sham PC, Purcell S, Cherny SS, Abecasis GR. *Powerful regression-based quantitative-trait linkage analysis of general pedigrees. Am J Hum Genet* 2002; 71: 238-53.
 11. Elston RC. *On fisher's method of combining p-values. Biom J* 1991; 33: 339-45.
 12. Meigs JB, Manning AK, Fox CS, Florez JC, Liu C, Cupples LA, Dupuis J. *Genome-wide association with diabetes-related traits in the Framingham Heart Study. BMC Med Genet* 2007; 8: S16.
 13. Busfield F, Duffy DL, Kesting JB, Walker SM, Lovelock PK, Good D, Tate H, Watego D, Marczak M, Hayman N, et al. *A genomewide search for type 2 diabetes-susceptibility genes in indigenous Australians. Am J Hum Genet* 2002; 70: 349-57.
 14. Hsueh WC, St Jean PL, Mitchell BD, Pollin TI, Knowler WC, Ehm MG, Bell CJ, Sakul H, Wagner MJ, Burns DK, et al. *Genome-wide and fine-mapping linkage studies of type 2 diabetes and glucose traits in the Old Order Amish: evidence for a new diabetes locus on chromosome 14q11 and confirmation of a locus on chromosome 1q21-q24. Diabetes* 2003; 52: 550-7.
 15. Lindgren CM, Mahtani MM, Widén E, McCarthy MI, Daly MJ, Kirby A, Reeve MP, Kruglyak L, Parker A, Meyer J, et al. *Genomewide search for type 2 diabetes mellitus susceptibility loci in Finnish families: the Botnia Study. Am J Hum Genet* 2002; 70: 509-16.
 16. Vionnet N, Hani EH, Dupont S, Gallina S, Francke S, Dotte S, De Matos F, Durand E, Leprêtre F, Lecoœur C, et al. *Genomewide search for type 2 diabetes-susceptibility genes in French whites: evidence for a novel susceptibility locus for early-onset diabetes on chromosome 3q27-qter and independent replication of a type 2-diabetes locus on chromosome 1q21-q24. Am J Hum Genet* 2000; 67: 1470-80.
 17. Ghosh S, Watanabe RM, Hauser ER, Valle T, Magnuson VL, Erdos MR, Langefeld CD, Balow J Jr, Ally DS, Kohtamaki K, et al. *Type 2 diabetes: evidence for linkage on chromosome 20 in 716 Finnish affected sib pairs. Proc Natl Acad Sci U S A* 1999; 96: 2198-203.
 18. Ghosh S, Watanabe RM, Valle TT, Hauser ER, Magnuson VL, Langefeld CD, Ally DS, Mohlke KL, Silander K, Kohtamäki K, et al. *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: 1174-85.
 19. Tsai FJ, Yang CF, Chen CC, Chuang LM, Lu CH, Chang CT, Wang TY, Chen RH, Shiu CF, Liu YM, et al. *A genome-wide association study identifies susceptibility variants for type 2 diabetes in Han Chinese. PLoS Genet* 2010; 6: e1000847.
 20. Hanson RL, Bogardus C, Duggan D, Kobes S, Knowlton M, Infante AM, Marovich L, Benitez D, Baier LJ, Knowler WC. *A search for variants associated with young-onset type 2 diabetes in American Indians in a 100K genotyping array. Diabetes* 2007; 56: 3045-52.
 21. Kapp K, Metzinger E, Kellerer M, Häring HU, Lammers R. *The protein tyrosine phosphatase alpha modifies insulin secretion in INS-1E cells. Biochem Biophys Res Commun* 2003; 311: 361-4.
 22. Ahmad F, Azevedo JL, Cortright R, Dohm GL, Goldstein BJ. *Alterations in skeletal muscle protein-tyrosine phosphatase activity and expression in insulin-resistant human obesity and diabetes. J Clin Invest* 1997; 100: 449-58.
 23. Hunter T. *Protein kinases and phosphatases: the yin and yang of protein phosphorylation and signaling. Cell* 1995; 80: 225-36.
 24. Ostman A, Hellberg C, Böhmer FD. *Protein-tyrosine phosphatases and cancer. Nat Rev Cancer* 2006; 6: 307-20.
 25. Zheng X, Resnick RJ, Shalloway D. *Apoptosis of estrogen-receptor negative breast cancer and colon cancer cell lines by PTP alpha and src RNAi. Int J Cancer* 2008; 122: 1999-2007.
 26. Wu CW, Kao HL, Li AF, Chi CW, Lin WC. *Protein tyrosine-phosphatase expression profiling in gastric cancer tissues. Cancer Lett* 2006; 242: 95-103.
 27. Ishii H, Jirousek MR, Koya D, Takagi C, Xia P, Clermont A, Bursell SE, Kern TS, Ballas LM, Heath WF, et al. *Amelioration of vascular dysfunctions in diabetic rats by an oral PKC beta inhibitor. Science* 1996; 272: 728-31.
 28. Bossenmaier B, Mosthaf L, Mischak H, Ullrich A, Häring HU. *Protein kinase C isoforms beta 1 and beta 2 inhibit the tyrosine kinase activity of the insulin receptor. Diabetologia* 1997; 40: 863-6.
 29. Standaert ML, Bandyopadhyay G, Galloway L, Soto J, Ono Y, Kikkawa U, Farese RV, Leitges M. *Effects of knockout of the protein kinase C beta gene on glucose transport and glucose homeostasis. Endocrinology* 1999; 140: 4470-7.
 30. Kaneto H, Suzuma K, Sharma A, Bonner-Weir S, King GL, Weir GC. *Involvement of protein kinase C beta 2 in c-myc induction by high glucose in pancreatic beta-cells. J Biol Chem* 2002; 277: 3680-5.
 31. Kawano Y, Rincon J, Soler A, Ryder JW, Nolte LA, Zierath JR, Wallberg-Henriksson H. *Changes in glucose transport and protein kinase Cbeta(2) in rat skeletal muscle induced by hyperglycaemia. Diabetologia* 1999; 42: 1071-9.
 32. Osterhoff MA, Heuer S, Pfeiffer M, Tasic J, Kaiser S, Isken F, Spranger J, Weickert MO, Möhlig M, Pfeiffer AF. *Identification of a functional protein kinase Cbeta promoter polymorphism in humans related to insulin resistance. Mol Genet Metab* 2008; 93: 210-5.
 33. Wang Q, Lin JL, Reinking BE, Feng HZ, Chan FC, Lin CI, Jin JP, Gustafson-Wagner EA, Scholz TD, Yang B, et al. *Essential roles of an intercalated disc protein, mXinbeta, in postnatal heart growth and survival. Circ Res*

- 2010; 106: 1468-78.
34. Takeuchi F, Katsuya T, Chakrewarthy S, Yamamoto K, Fujioka A, Serizawa M, Fujisawa T, Nakashima E, Ohnaka K, Ikegami H, et al. *Common variants at the GCK, GCKR, G6PC2-ABCB11 and MTNR1B loci are associated with fasting glucose in two Asian populations.* *Diabetologia* 2010; 53: 299-308.
35. Rasmussen-Torvik LJ, Alonso A, Li M, Kao W, Köttgen A, Yan Y, Couper D, Boerwinkle E, Bielinski SJ, Pankow JS. *Impact of repeated measures and sample selection on genome-wide association studies of fasting glucose.* *Genet Epidemiol* 2010; 34: 665-73.
36. Ramos E, Chen G, Shriner D, Doumatey A, Gerry NP, Herbert A, Huang H, Zhou J, Christman MF, Adeyemo A, et al. *Replication of genome-wide association studies (GWAS) loci for fasting plasma glucose in African-Americans.* *Diabetologia* 2011; 54: 783-8.
37. Chen WM, Erdos MR, Jackson AU, Saxena R, Sanna S, Silver KD, Timpson NJ, Hansen T, Orrù M, Grazia Piras M, et al. *Variations in the G6PC2/ABCB11 genomic region are associated with fasting glucose levels.* *J Clin Invest* 2008; 118: 2620-8.
38. Shu XO, Long J, Cai Q, Qi L, Xiang YB, Cho YS, Tai ES, Li X, Lin X, Chow WH, et al. *Identification of new genetic risk variants for type 2 diabetes.* *PLoS Genet* 2010; 6. pii: e1001127.
39. Ma RC, Tam CH, Wang Y, Luk AO, Hu C, Yang X, Lam V, Chan AW, Ho JS, Chow CC, et al. *Genetic variants of the protein kinase C-beta 1 gene and development of end-stage renal disease in patients with type 2 diabetes.* *JAMA* 2010; 304: 881-9.
40. Ikeda Y, Suehiro T, Osaki F, Tsuzura S, Kumon Y, Hashimoto K. *Polymorphisms in the 5'-upstream region of the PKCbeta gene in Japanese patients with Type 2 diabetes.* *Diabet Med* 2004; 21: 1113-20.