



OPEN

Genome-wide association study for circulating fibroblast growth factor 21 and 23

Gwo-Tsann Chuang^{1,2,14}, Pi-Hua Liu^{3,4,14}, Tsui-Wei Chyan⁵, Chen-Hao Huang⁵, Yu-Yao Huang^{4,6}, Chia-Hung Lin^{4,7}, Jou-Wei Lin⁸, Chih-Neng Hsu⁸, Ru-Yi Tsai⁸, Meng-Lun Hsieh⁹, Hsiao-Lin Lee⁹, Wei-shun Yang^{2,9}, Cassianne Robinson-Cohen¹⁰, Chia-Ni Hsiung¹¹, Chen-Yang Shen^{12,13} & Yi-Cheng Chang^{2,9,12}✉

Fibroblast growth factors (FGFs) 21 and 23 are recently identified hormones regulating metabolism of glucose, lipid, phosphate and vitamin D. Here we conducted a genome-wide association study (GWAS) for circulating FGF21 and FGF23 concentrations to identify their genetic determinants. We enrolled 5,000 participants from Taiwan Biobank for this GWAS. After excluding participants with diabetes mellitus and quality control, association of single nucleotide polymorphisms (SNPs) with log-transformed FGF21 and FGF23 serum concentrations adjusted for age, sex and principal components of ancestry were analyzed. A second model additionally adjusted for body mass index (BMI) and a third model additionally adjusted for BMI and estimated glomerular filtration rate (eGFR) were used. A total of 4,201 participants underwent GWAS analysis. rs67327215, located within *RGS6* (a gene involved in fatty acid synthesis), and two other SNPs (rs12565114 and rs9520257, located between *PHC2-ZSCAN20* and *ARGLU1-FAM155A* respectively) showed suggestive associations with serum FGF21 level ($P = 6.66 \times 10^{-7}$, 6.00×10^{-7} and 6.11×10^{-7} respectively). The SNPs rs17111495 and rs17843626 were significantly associated with FGF23 level, with the former near *PCSK9* gene and the latter near *HLA-DQA1* gene ($P = 1.04 \times 10^{-10}$ and 1.80×10^{-8} respectively). SNP rs2798631, located within the *TGFB2* gene, was suggestively associated with serum FGF23 level ($P = 4.97 \times 10^{-7}$). Additional adjustment for BMI yielded similar results. For FGF23, further adjustment for eGFR had similar results. We conducted the first GWAS of circulating FGF21 levels to date. Novel candidate genetic loci associated with circulating FGF21 or FGF23 levels were found. Further replication and functional studies are needed to support our findings.

Fibroblast growth factor (FGF) superfamily includes 22 family members. Except for the endocrine FGF19, FGF21, and FGF23, other family members act as autocrine or paracrine factors¹. Recent evidence showed that FGF21 exerts metabolic effects by acting both centrally and peripherally, mostly in the liver, adipose tissue and pancreas. In prolonged starvation, FGF21 is secreted by the liver to promote gluconeogenesis, increases hepatic

¹Department of Pediatrics, National Taiwan University Hospital, College of Medicine, National Taiwan University, Taipei, Taiwan, ROC. ²Graduate Institute of Medical Genomics and Proteomics, National Taiwan University, 5F, No.2, Xuzhou Rd., Zhongzheng Dist., Taipei 100, Taiwan, ROC. ³Clinical Informatics and Medical Statistics Research Center, College of Medicine, Chang Gung University, Taoyuan, Taiwan, ROC. ⁴Division of Endocrinology and Metabolism, Department of Internal Medicine, Chang Gung Memorial Hospital At Linkou, Taoyuan, Taiwan, ROC. ⁵Institute of Molecular Medicine, College of Medicine, National Taiwan University, Taipei, Taiwan, ROC. ⁶Department of Medical Nutrition Therapy, Chang Gung Memorial Hospital At Linkou, Taoyuan, Taiwan, ROC. ⁷Department of Chinese Medicine, College of Medicine, Chang Gung University, Taoyuan, Taiwan, ROC. ⁸Cardiovascular Center, National Taiwan University Hospital Yun-Lin Branch, Taipei, Taiwan, ROC. ⁹Department of Internal Medicine, College of Medicine, National Taiwan University, Taipei, Taiwan, ROC. ¹⁰Division of Nephrology, Department of Medicine, Vanderbilt University Medical Center, Nashville, TN, USA. ¹¹Data Science Statistical Cooperation Center, Institute of Statistical Science, Academia Sinica, Taipei, Taiwan, ROC. ¹²Institute of Biomedical Sciences, Academia Sinica, Taipei, Taiwan, ROC. ¹³College of Public Health, China Medical University, Taichung, Taiwan, ROC. ¹⁴These authors contributed equally: Gwo-Tsann Chuang and Pi-Hua Liu. ✉email: b83401040@gmail.com

Characteristic	
Total participants (N)	4,201
Age, year	49.54 ± 10.82
Males (%)	1,446 (34.42)
BMI, kg/m ²	24.32 ± 6.67
Serum glucose, mg/dl	93.01 ± 8.26
eGFR (MDRD), ml/min/1.73 m ²	106.80 ± 24.98
FGF21, ng/ml	0.16 (0.10–0.27)
FGF23, ng/ml	0.36 (0.11–1.74)

Table 1. Study characteristics of participants. Data are mean ± SD, median (IQR) or number (%), as appropriate. Age is at specimen collection. *BMI* body mass index; *MDRD* modification of diet in renal disease equation.

SNP	Chr	Position	Nearest gene	FGF21 Increasing allele	Other allele	FGF21 increasing allele frequency	Model 1		Model 2	
							P value	Beta (SEM)	P value	Beta (SEM)
rs12565114	1	33,921,202	<i>PHC2-ZSCAN20</i>	G	A	0.69	6.00 × 10 ⁻⁷	0.048 (0.010)	4.69 × 10 ⁻⁷	0.049 (0.010)
rs9520257	13	107,631,394	<i>ARGLU1-FAM155A</i>	C	T	0.85	6.11 × 10 ⁻⁷	0.062 (0.012)	3.96 × 10 ⁻⁷	0.063 (0.012)
rs67327215	14	72,694,874	<i>RGS6</i>	G	A	0.62	6.66 × 10 ⁻⁷	0.046 (0.009)	6.93 × 10 ⁻⁷	0.046 (0.009)

Table 2. Top genetic polymorphisms associated with log-transformed FGF21 level. Model 1 was adjusted for age, sex and the first ten principal components of ancestry. Model 2 was additionally adjusted for BMI. *SNP* single nucleotide polymorphism; *Chr* chromosome. rs12565114 and rs9520257 are genotyped. rs67327215 is imputed with high info score (0.984).

fatty acid oxidation and ketogenesis², enhances secretion of glucocorticoid, and reduces behaviors that waste energy such as ovulation³. In contrast, in fed status, FGF21 promotes glucose uptake by fat cells and increase heat production of brown fat⁴. In the pancreatic islet, FGF21 preserves the survival of beta-cell and restores insulin synthesis under the stress of nutrient excess⁵. Of note, FGF21 acts on the brain to induce sympathetic nerve activity, thereby increasing energy expenditure and promoting weight loss⁶. FGF21 is regulated by both peroxisome proliferator-activated receptor (PPAR) γ and PPAR α and is an important mediator of the downstream effects of PPAR γ and PPAR α ^{6,7}. FGF21 is currently a drug target for treating lipid disorders and fatty liver⁸.

FGF23 is the physiological regulator of phosphate and vitamin D serum levels. FGF23 lowers apical membrane expression of sodium-dependent phosphate co-transporter 2A and sodium-dependent phosphate co-transporter 2C in the kidney, both of which primarily mediate renal tubular phosphate reabsorption. Furthermore, FGF23 lowers serum 1,25-dihydroxyvitamin D₃ (1,25(OH)₂D₃) levels, thereby preventing hyperphosphatemia and hypervitaminosis^{1,9}. FGF23 has structure and biological features similar to those of FGF19 and FGF21, which were known to regulate glucose and lipid metabolism¹⁰. FGF23 level was also found to be associated with obesity and dyslipidemia in studies of elderly cohorts¹¹. FGF23 is also an important drug target for treating phosphate disorder and bone disease¹². Robinson-Cohen et al. performed a genome-wide association study (GWAS) of circulating FGF23 concentrations among people of European ancestry, and found five genome-wide significant loci¹³.

To understand better the regulation of circulating FGF21 and FGF23 levels, we performed a GWAS to identify genetic determinants of circulating FGF21 and FGF23 levels in Taiwanese population.

Results

Of the 5,000 subjects enrolled from Taiwan Biobank, one withdrew from the study, 559 were excluded due to their having diabetes mellitus, and 239 were excluded by quality control procedures. A total of 617,073 genotyped autosomal SNPs remained. After imputation and post-imputation quality control, 7,897,704 SNPs remained. We performed GWAS analysis in the remaining 4,201 subjects, where log-transformed FGF21 or FGF23 level was the quantitative phenotype. Study characteristics are shown in Table 1.

The results of association testing for FGF21 and FGF23 are listed, showing SNP with strongest association in each region (Tables 2 and 3, respectively). Manhattan plots and quantile–quantile plots of log-transformed FGF21 and FGF23 are shown in Figs. 1 and 2, respectively. Regional GWAS plots of each SNP are shown in Figs 3 and 4. For FGF21, no significantly associated SNP is found. One suggestive locus is located at the intergenic region between *PHC2* and *ZSCAN20* genes on chromosome 1 (rs12565114; $P = 6.00 \times 10^{-7}$, Fig. 3a), another is between *ARGLU1* and *FAM155A* genes on chromosome 13 (rs9520257; $P = 6.11 \times 10^{-7}$, Fig. 3b), and the other is within the *RGS6* gene on chromosome 14 (rs67327215; $P = 6.66 \times 10^{-7}$, Fig. 3c). For FGF23, two significant loci are found. One locus is near gene *PCSK9* on chromosome 1 (rs17111495; $P = 1.04 \times 10^{-10}$, Fig. 4a), and the other is near the *HLA-DQA1* gene on chromosome 6 (rs17843626; $P = 1.80 \times 10^{-8}$, Fig. 4b). There is also a suggestive locus within the *TGFB2* gene on chromosome 1 (rs2798631; $P = 4.97 \times 10^{-7}$, Fig. 4c). The array-based heritability estimates for FGF21 and FGF23 were 0.1171 and 0.1104 respectively. Proportion of variance in serum FGF21 level explained

SNP	Chr	Position	Nearest gene	FGF23 increasing allele	Other allele	FGF23 increasing allele frequency	Model 1		Model 2		Model 3	
							P value	Beta (SEM)	P value	Beta (SEM)	P value	Beta (SEM)
rs17111495	1	55,500,706	PCSK9	G	C	0.98	1.04×10^{-10}	0.463 (0.071)	1.57×10^{-10}	0.458 (0.071)	1.55×10^{-10}	0.458 (0.071)
rs17843626	6	32,621,013	HLA-DQA1	G	A	0.49	1.80×10^{-8}	0.117 (0.021)	1.70×10^{-8}	0.117 (0.021)	1.75×10^{-8}	0.117 (0.021)
rs2798631	1	218,611,878	TGFB2	G	A	0.75	4.97×10^{-7}	0.115 (0.023)	3.73×10^{-7}	0.116 (0.023)	3.71×10^{-7}	0.116 (0.023)

Table 3. Top genetic polymorphisms associated with log-transformed FGF23 level. Model 1 was adjusted for age, sex and the first ten principal components of ancestry. Model 2 was additionally adjusted for BMI. Model 3 was additionally adjusted for BMI and eGFR. SNP single nucleotide polymorphism; Chr chromosome. rs17111495, rs2798631 and rs17843626 are imputed with high info scores (0.809, 0.997 and 0.990 respectively).

by rs12565114, rs9520257 and rs67327215 were 0.5%, 0.6% and 0.5% respectively; for serum FGF23 level, proportion of variance explained by rs17111495, rs2798631 and rs17843626 were 1.0%, 0.6% and 0.7% respectively.

Further excluding two more subjects with severe renal impairment (eGFR < 30 ml/min/1.73m²) did not alter the results of top SNPs associated with FGF23 level (Supplementary Table S1).

We looked for the association of the five SNPs identified to be associated with serum FGF23 levels by Robinson-Cohen et al.¹³ in the European population but failed to find significant association with the FGF23 level in our cohort (Supplementary Table S2). In view of low minor allele frequency of rs17216707, we performed gene-based association analysis using MAGMA v1.07b¹⁴ and still found no significant association (CYP24A1, $P=0.18$). The regional plots regarding association with serum FGF23 levels in Taiwan Biobank at the five loci identified previously by Robinson-Cohen and colleagues are shown in Supplementary Fig. S1a–e. We also performed meta-analysis pooling our results and those reported by Robinson-Cohen et al.¹³ using METAL¹⁵ but the results were not significant as shown in Supplementary Table S3 (P values were 0.00070, 0.0010 and 0.00029 for rs17111495, rs2798631 and rs17843626, respectively).

Discussion

In this GWAS for evaluating both circulating FGF21 and FGF23 levels, we identified several loci for each trait that showed associations. SNP rs67327215, associated with FGF21 level, is located within *RGS6* (Regulator of G protein signaling 6). *RGS6* belongs to the R7 subfamily of RGS proteins¹⁶ and is expressed in a variety of organs and tissues, including the central nervous system and liver¹⁷. RGS proteins regulate G protein-coupled receptor-initiated signaling by acting as GTPase-activating proteins for G α subunits, which hydrolyze GTP and restore the inactive G α_{GDP} $\beta\gamma$ heterotrimer¹⁸. A previous study has shown *Rgs16* knockout mice expressing increased *Fgf21* expression in the liver while overexpression of *Rgs16* decreased *Fgf21* expression. It was postulated that RGS16 inhibits G α_i /G α_q -mediated fatty acid oxidation, thereby decreasing the induced *Fgf21* expression by peroxisome proliferator-activated receptor alpha (PPAR α)¹⁹. RGS6 knockout mice revealed lower mRNA levels of PPAR α and PPAR γ ²⁰. These data support that RGS6 may also act as a regulator of circulating FGF21 levels through fatty acid synthesis and oxidation.

The SNPs rs12565114 and rs9520257, also suggestively associated with FGF21 level, are intergenic SNPs on chromosome 1 and 13 respectively. rs12565114 is located between *PHC2* (Polyhomeotic Homolog 2) and *ZSCAN20* (Zinc Finger And SCAN Domain Containing 20) genes. *PHC2* is associated with conventional angiosarcoma²¹. *ZSCAN20* is associated with diabetic neuropathic pain²². rs9520257 is located between gene *ARGLU1* (Arginine And Glutamate Rich 1) and *FAM155A* (Family With Sequence Similarity 155 Member A) genes. *ARGLU1* acts in cooperation with MED1 (Mediator Complex Subunit 1) and is required for estrogen-dependent gene transcription and breast cancer cell growth²³. *FAM155A* is associated with diverticulitis²⁴. None of the genes above were related to FGF21 according to our current understanding. There might be microRNAs or long non-coding RNAs encoded in these intergenic regions affecting FGF21 levels.

SNP rs17111495, strongly associated with FGF23 levels, is located upstream of *PCSK9* (Proprotein convertase subtilisin/kexin type 9). *PCSK9* is a member of proprotein convertase family PCSK (subtilisin-like proprotein convertases previously) and is synthesized as a soluble zymogen that undergoes autocatalytic intramolecular processing in the endoplasmic reticulum²⁵. FGF23 is inactivated when being cleaved intracellularly by PCSK at the minimum consensus sequence RHTR¹⁷⁹ between Arg¹⁷⁹ and Ser^{180,26,27}. Autosomal dominant hypophosphatemic rickets is caused by gain of function mutations in FGF23 that renders it resistant to PCSK cleaving at site RHTR¹⁷⁹, thus resulting in elevated circulating FGF23 level and the consequent renal phosphate wasting, rickets, and osteomalacia²⁶. Hyperphosphatemic familial tumoral calcinosis due to GALNT3 mutation involves deficient O-glycosylation of the threonine residue in the RTHR¹⁷⁹ proprotein convertase-processing site, thus favoring intracellular degradation of intact FGF23 by PCSK and the resulting reduced phosphate renal excretion²⁸. Collectively, previous studies and our findings suggest that PCSK is a critical regulator of FGF23 secretion.

SNP rs17843626, significantly associated with circulating FGF23 levels, is located between *HLA-DQA1* and *HLA-DQB1* (Human Leukocyte Antigen, Class II, DQ Alpha 1 and Beta 1). In general, these HLA class II loci are known to be associated with autoimmune diseases such as type 1 diabetes, multiple sclerosis and rheumatoid arthritis²⁹. No literature has reported association between these loci and FGF23 level or its regulators, parathyroid hormone and vitamin D. This is also a new finding.

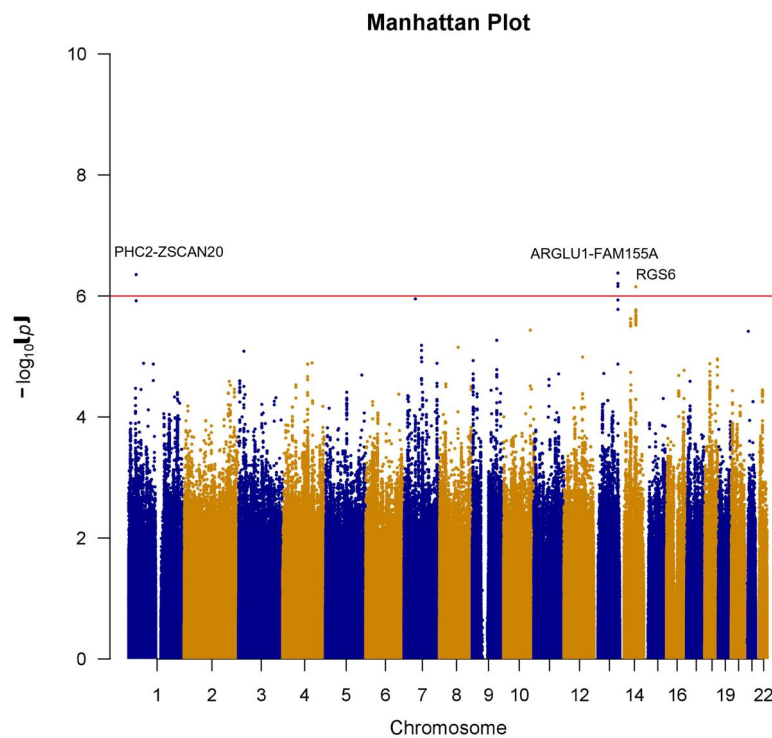
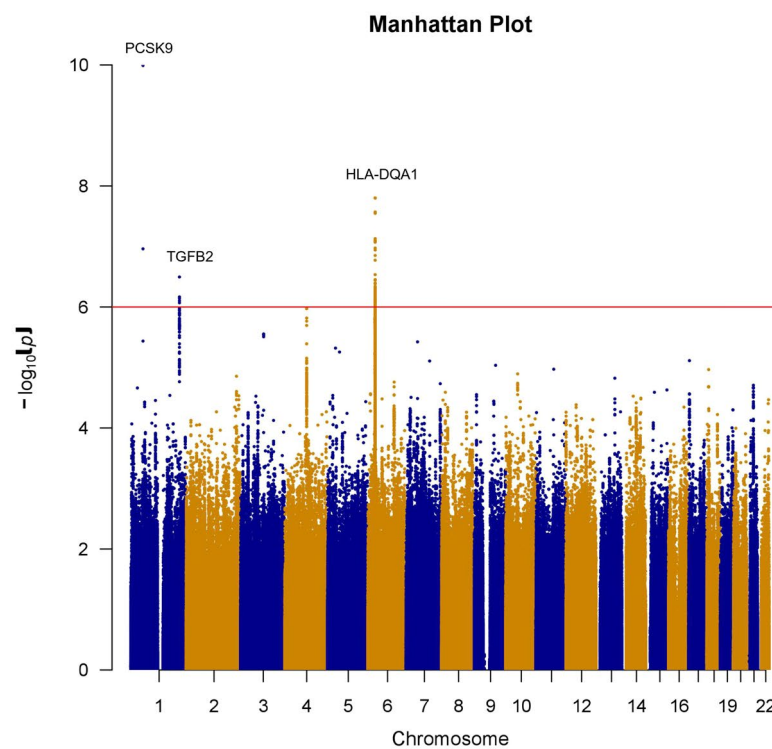
(a) Circulating levels of FGF21**(b) Circulating levels of FGF23**

Figure 1. Manhattan plot of the GWAS results for FGF21. SNPs are plotted on the x axis according to their chromosome position against association with (a) log FGF21 or (b) log FGF23 on the y axis. The red horizontal line represents the suggestive threshold of $P = 1.0 \times 10^{-6}$.

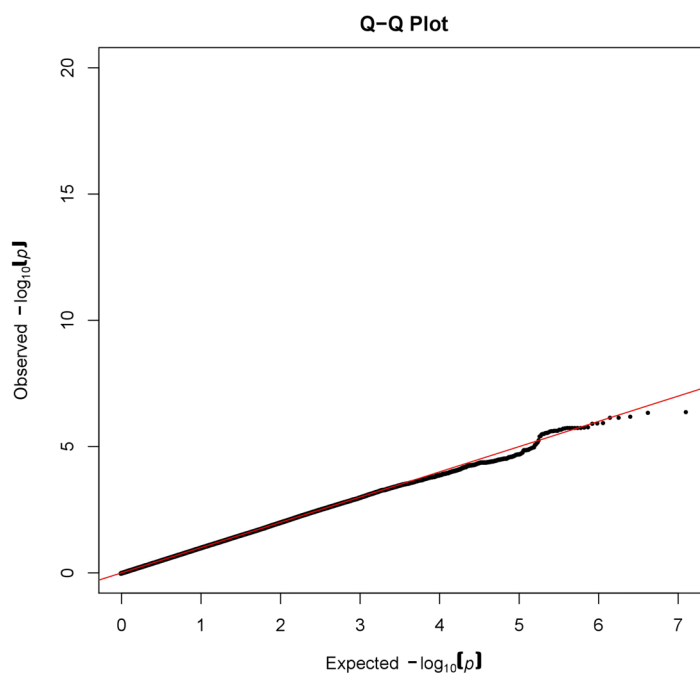
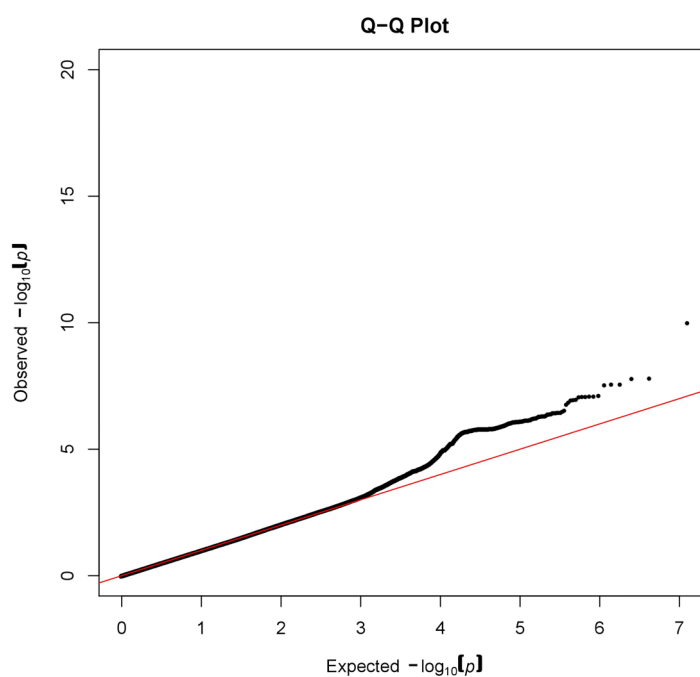
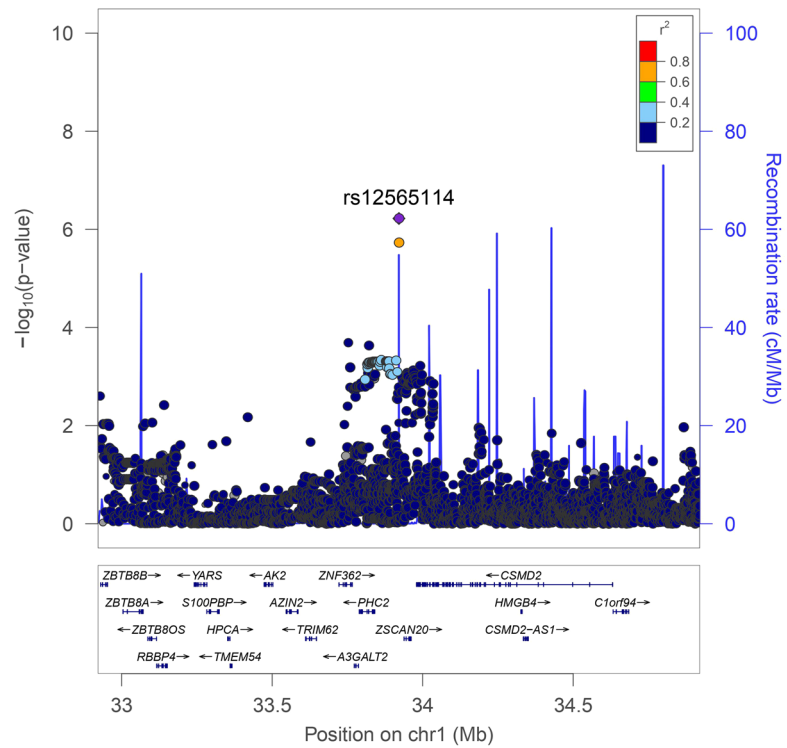
(a) Circulating FGF21 levels**(b) Circulating FGF23 levels**

Figure 2. Quantile–quantile plots of **(a)** log FGF21 and **(b)** log FGF23.

SNP rs2798631, suggestively associated with circulating FGF23 levels, is located within *TGFB2* (Transforming growth factor beta-2), which is involved in cell proliferation, differentiation, inflammation and apoptosis. A recent study has demonstrated that TGF- β 2 directly enhanced FGF23 production in rat osteoblast-like cells through increasing calcium entry³⁰. 1,25(OH)₂D₃, the active vitamin D metabolite which induces FGF23 production, also inhibits TGF- β downstream signaling³¹. Downstream TGF- β and 1,25(OH)₂D₃ signaling are characterized by intense crosstalk³². The expression of FGF23 by both modulators also seemed to be subjective to this crosstalk³³.

For more comprehensive bioinformatic annotations of these top SNPs, we used GRASP v2.0 to search for documented associations between these SNPs and phenotypes³⁴. One study showed association between rs12565114

(a) rs12565114



(b) rs9520257

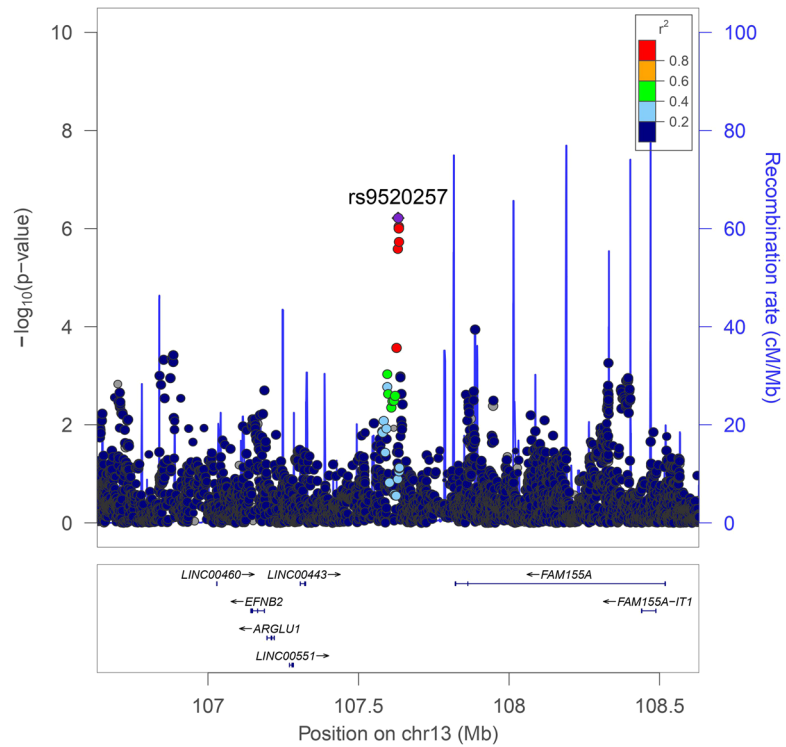
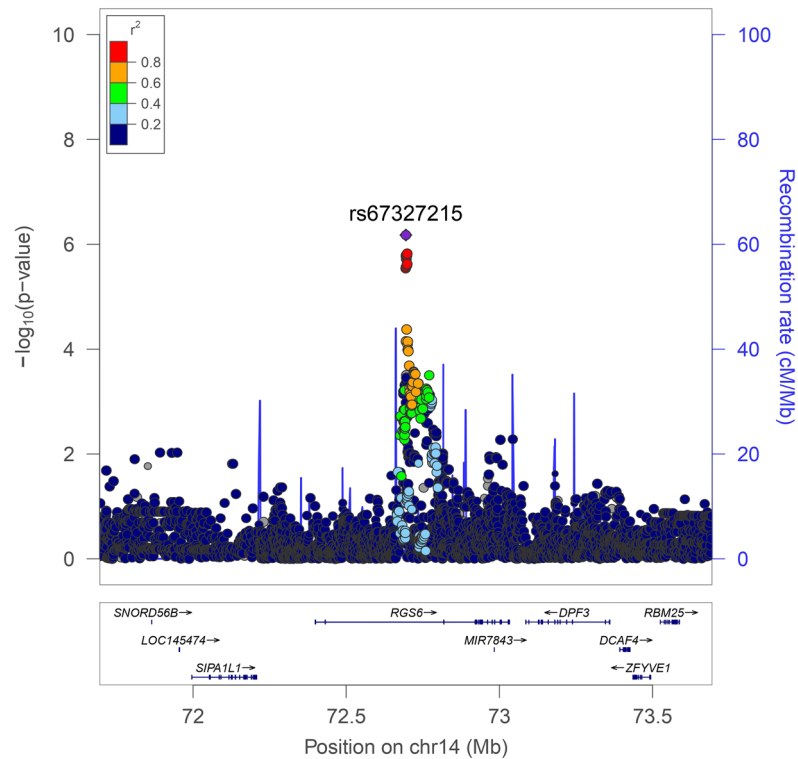


Figure 3. Regional association plots of log FGF21.

(e) rs67327215

**Figure 3.** (continued)

and schizophrenia³⁵. Several studies demonstrated association between rs17111495 and serum lipid levels^{36,37}. rs2798631 was associated with height³⁸, refractive error³⁹ and Parkinson's disease⁴⁰. There were no results for rs9520257, rs67327215 and rs17843626.

The absolute circulating FGF23 levels measured in this study were different from that obtained by Robinson-Cohen et al. This study used Duoset ELISA Development kit (R&D system, Inc.) to measure intact FGF23 levels; and the median intact FGF23 level in the studied population was 360 (110–1,740) pg/ml (expressed as median and interquartile range). For published researches using this ELISA kit, the median of intact FGF23 level was 269.9 (109.2–1,014.0) pg/ml in Chinese renal transplant donors⁴¹, and 150 (46–583) pg/ml in Mexicans with normal kidney function⁴². In contrary, Robinson-Cohen et al. used a different ELISA kit (Kainos Laboratories, Inc., Tokyo, Japan), and the population for their study was of European ancestry. Their mean intact FGF23 level was around 40–50 pg/ml¹³. The variations in FGF23 level obtained may be attributed to ethnicity of the studied populations and different ELISA kits used.

Our study has some unique strength and some limitations. It is the first known study identifying suggestive genetic determinants of circulating FGF21 level. We also evaluated SNPs associated with circulating FGF23 level from a population of ancestry different from those studied in previous GWAS, and novel candidate loci were found.

As mentioned above, our findings differed from those of Robinson-Cohen et al.¹³, possibly due to differences in ethnicity, differences in gene-environment and gene-gene interactions, and our relatively small sample size. As rs17216707 minor allele frequency was low in our cohort, gene-based association analysis was conducted as an alternative way in effort to replicate the findings but the result was not significant. A meta-analysis conducted on both this study and that of Robinson-Cohen et al. found no statistically significant results either.

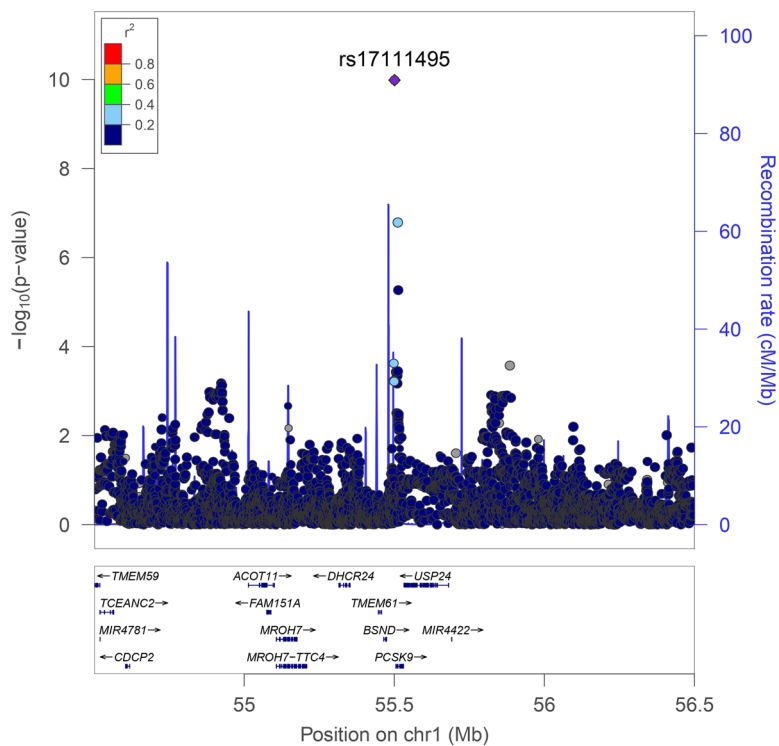
Regarding the lack of replication in an external cohort of the study findings, we searched for GWAS summary statistics of circulating FGF21 or FGF23 phenotypes in UK Biobank and Biobank Japan but found them unavailable. Further replication may be needed in the future.

In conclusion, this study is the first GWAS on circulating FGF21 level to date. Novel candidate genetic loci possibly related to circulating FGF21 and FGF23 levels were found. Further replication and functional studies are needed to support the present findings.

Materials and methods

Study population. Five thousand subjects were enrolled from Taiwan Biobank by random sampling as study population. Subjects are aged between 30 and 70 years, without cancer history and volunteered to join the Taiwan Biobank study population. Eating and drinking were refrained before blood samples were drawn. Delinked databases including biological specimens, personal data and clinical information were used in this investigation. The estimated glomerular filtration rate (eGFR) was calculated using the 4-variable equation from

(a) rs17111495



(b) rs17843626

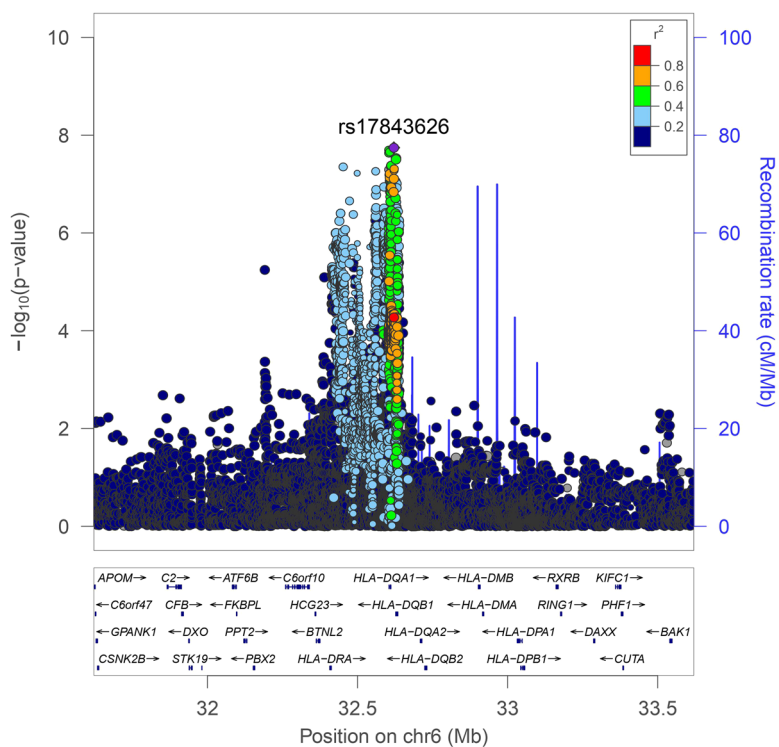
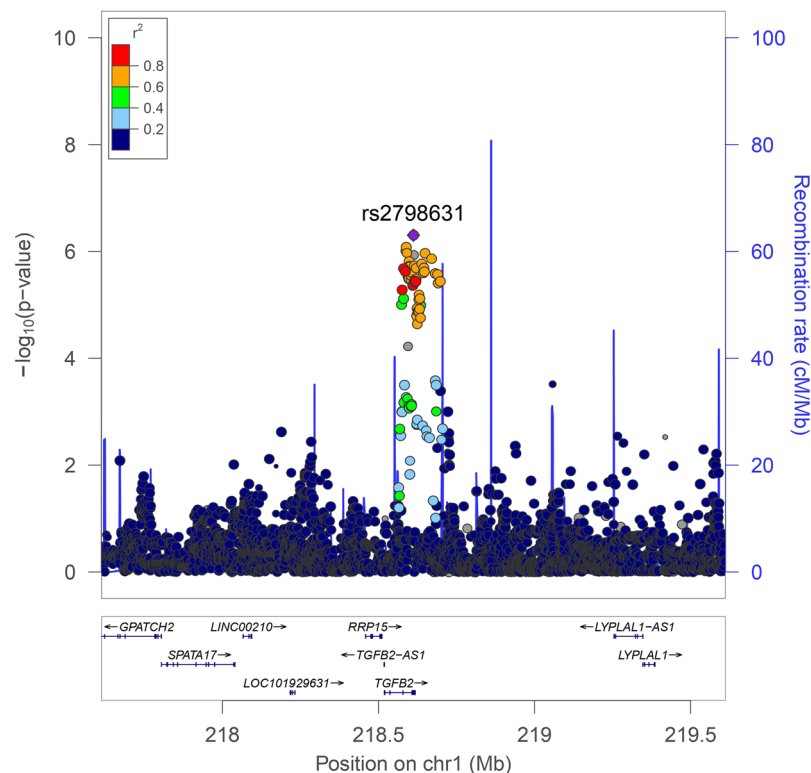


Figure 4. Regional association plots of log FGF23.

(e) rs2798631

**Figure 4.** (continued)

the Modification of Diet in Renal Disease (MDRD) Study⁴³. There were only two individuals with severe renal impairment (eGFR < 30 ml/min/1.73 m²), and their serum FGF23 levels were 0.259 and 0.709 ng/ml.

All subjects provided written informed consent. Individuals diagnosed with diabetes mellitus (either as stated in the questionnaire or HbA1c ≥ 6.5%) were excluded.

FGF21 and FGF23 measurements. Intact circulating FGF21 and FGF23 levels in serum of Taiwan Biobank subjects were measured using Enzyme-Linked Immunosorbent Assay (Duoset ELISA Development kit, R&D system, Inc.). The assays have high sensitivity and exhibit no interference or cross-reactivity with recombinant human FGF R1a/Fc Chimera, FGF R2a/Fc Chimera, FGF R3/Fc Chimera, FGF R4/Fc Chimera and Klotho. All standards by serial dilution were assayed in duplicates.

Genotyping, quality control and imputation. Genome-wide genotyping of all subjects was carried out at the National Center for Genome Medicine of Academia Sinica using the Axiom-Taiwan Biobank Array Plate (TWB chip; Affymetrix Inc, Santa Clara, California)⁴⁴. TWB chip consists of 653,291 SNPs, and was specifically customized for the Taiwanese population, who are mainly of Han-Chinese lineage, by including SNPs with detected polymorphisms in Taiwanese-based genotyping results from the Axiom Genome-Wide CHB 1 Array plate (Affymetrix Inc). SNPs from ancestry information panels, GWAS and cancer studies, as well as pharmacogenetics arrays were also incorporated into the TWB chip. PLINK (version 1.07), an open-source whole-genome association analysis toolset was used for quality control⁴⁵. Genotypes for SNPs with batch effect were set as missing. Individuals with high missing genotype rate (> 5%), with extreme heterozygosity rate (above or below 5 standard deviations of mean heterozygosity rate) or being closely related as assessed by Identity-By-Descent (IBD) estimation (IBD ≥ 0.1875) were excluded from analyses. SNPs with high missing genotype rate (> 5%), low frequency (minor allele frequency < 1%), or deviation from Hardy-Weinberg Equilibrium (HWE) (P value < 10⁻⁵) were excluded, with 617,073 SNPs remaining. Population structure was assessed using principal component analysis. Using PLINK, we computed the principal components on an LD-pruned ($r^2 < 0.2$) set of autosomal variants obtained by removing high-LD regions. Genome-wide genotype imputation was performed by SHAPEIT⁴⁶ and IMPUTE2⁴⁷ with 1,000 Genomes Project (1000GP) Phase 3 East Asian (EAS) population as reference panel. Quality control after imputation was as follows: We filtered the SNPs with imputation quality score of IMPUTE2 (info score⁴⁸) greater than 0.3 for further analysis. Indels were removed with VCFtools⁴⁹. SNPs with low frequency (minor allele frequency < 1%) were excluded.

Statistical analyses. Age, body mass index (BMI), glucose level and eGFR were expressed as mean and standard deviation. FGF21 and FGF23 levels were expressed as median and interquartile range. GWAS analyses were performed using an additive genetic model in PLINK v1.07. Skewed variables including serum FGF21

and FGF23 were logarithmically transformed to approximate normal distribution. We performed regression diagnostics plots to check the assumptions of linearity, normality, and homoscedasticity, as well as to identify the influential observations for each outcome of interest (logFGF21 and logFGF23) and SNPs, adjusting for covariates. We did not find any severe violations of assumptions. Further, variance inflation factor (VIF) was used to check for multicollinearity. No correlations between independent variables were found. Linear regression was employed to analyze association between SNPs and log-transformed FGF21 and FGF23 levels. Population stratification was not obvious ($\lambda = 1.002$ for FGF21, $\lambda = 0.999$ for FGF23). Age, sex and the first ten principal components of ancestry were adjusted in Model 1, and BMI was additionally adjusted in Model 2. For log FGF23, BMI and eGFR were additionally adjusted in Model 3. The threshold for genome-wide significance was set at $P = 5 \times 10^{-850}$. In view of the relatively small sample size and to avoid false negative results caused by too stringent threshold⁵¹, the P value threshold for suggestive results were set at 1.0×10^{-652} . Manhattan plots and quantile–quantile plots were drawn using the qqman R package^{53,54}. Regional association plots were generated by LocusZoom⁵⁵. 1,000 Genomes Project Phase 3 East Asian Ancestry was used for the reference population and Genome Reference Consortium Human Build 37 was used for gene annotation. The array-based heritability was calculated by linkage disequilibrium (LD) score regression⁵⁶. The proportion of phenotypic variance explained by single SNP was estimated by the following formula:

$$\frac{2\beta^2 * \text{MAF} * (1 - \text{MAF})}{2\beta^2 * \text{MAF} * (1 - \text{MAF}) + \text{SE}^2 * 2N * \text{MAF} * (1 - \text{MAF})}$$

where β , SE, N, and MAF are the effect size estimate of each minor allele on the relative concentration of FGF21/FGF23, standard error of the effect size, sample size, and minor allele frequency for the SNP, respectively⁵⁷. Gene-based association analysis was performed using MAGMA v1.07b¹⁴ with LD information retrieved from 1,000 Genomes Project Phase 3 East Asian Panel. Meta-analysis pooling our study and those reported by Robinson-Cohen et al. was performed using by METAL using a random effects model¹⁵.

Bioethics statement. All methods were carried out in accordance with relevant guidelines and regulations. The experimental protocols were approved by the Institutional Review Board of National Taiwan University Hospital.

Data availability

Individual researchers may request to use the data for specific projects on a collaborative basis.

Received: 24 November 2019; Accepted: 6 August 2020

Published online: 03 September 2020

References

1. Degirolamo, C., Sabba, C. & Moschetta, A. Therapeutic potential of the endocrine fibroblast growth factors FGF19, FGF21 and FGF23. *Nat. Rev. Drug Discov.* **15**, 51–69. <https://doi.org/10.1038/nrd.2015.9> (2016).
2. Badman, M. K. *et al.* Hepatic fibroblast growth factor 21 is regulated by PPARalpha and is a key mediator of hepatic lipid metabolism in ketotic states. *Cell Metab.* **5**, 426–437. <https://doi.org/10.1016/j.cmet.2007.05.002> (2007).
3. Bookout, A. L. *et al.* FGF21 regulates metabolism and circadian behavior by acting on the nervous system. *Nat. Med.* **19**, 1147–1152. <https://doi.org/10.1038/nm.3249> (2013).
4. Chartoumpakis, D. V. *et al.* Brown adipose tissue responds to cold and adrenergic stimulation by induction of FGF21. *Mol. Med.* **17**, 736–740. <https://doi.org/10.2119/molmed.2011.00075> (2011).
5. Wenthe, W. *et al.* Fibroblast growth factor-21 improves pancreatic beta-cell function and survival by activation of extracellular signal-regulated kinase 1/2 and Akt signaling pathways. *Diabetes* **55**, 2470–2478. <https://doi.org/10.2337/db05-1435> (2006).
6. Owen, B. M. *et al.* FGF21 acts centrally to induce sympathetic nerve activity, energy expenditure, and weight loss. *Cell Metab.* **20**, 670–677. <https://doi.org/10.1016/j.cmet.2014.07.012> (2014).
7. Dutchak, P. A. *et al.* Fibroblast growth factor-21 regulates PPARgamma activity and the antidiabetic actions of thiazolidinediones. *Cell* **148**, 556–567. <https://doi.org/10.1016/j.cell.2011.11.062> (2012).
8. Kharitonov, A. & Adams, A. C. Inventing new medicines: The FGF21 story. *Mol. Metab.* **3**, 221–229. <https://doi.org/10.1016/j.molmet.2013.12.003> (2014).
9. Kanbay, M. *et al.* Novel faces of fibroblast growth factor 23 (FGF23): iron deficiency, inflammation, insulin resistance, left ventricular hypertrophy, proteinuria and acute kidney injury. *Calcif. Tissue Int.* **100**, 217–228. <https://doi.org/10.1007/s00223-016-0206-7> (2017).
10. Angelin, B., Larsson, T. E. & Rudling, M. Circulating fibroblast growth factors as metabolic regulators—a critical appraisal. *Cell Metab.* **16**, 693–705. <https://doi.org/10.1016/j.cmet.2012.11.001> (2012).
11. Mirza, M. A. *et al.* Circulating fibroblast growth factor-23 is associated with fat mass and dyslipidemia in two independent cohorts of elderly individuals. *Arterioscler. Thromb. Vasc. Biol.* **31**, 219–227. <https://doi.org/10.1161/ATVBAHA.110.214619> (2011).
12. Kinoshita, Y. & Fukumoto, S. X-linked hypophosphatemia and FGF23-related hypophosphatemic diseases: prospect for new treatment. *Endocr. Rev.* **39**, 274–291. <https://doi.org/10.1210/er.2017-00220> (2018).
13. Robinson-Cohen, C. *et al.* Genetic variants associated with circulating fibroblast growth factor 23. *J. Am. Soc. Nephrol.* **29**, 2583–2592. <https://doi.org/10.1681/ASN.2018020192> (2018).
14. de Leeuw, C. A., Mooij, J. M., Heskes, T. & Posthuma, D. MAGMA: generalized gene-set analysis of GWAS data. *PLoS Comput. Biol.* **11**, e1004219. <https://doi.org/10.1371/journal.pcbi.1004219> (2015).
15. Willer, C. J., Li, Y. & Abecasis, G. R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190–2191. <https://doi.org/10.1093/bioinformatics/btq340> (2010).
16. Chatterjee, T. K., Liu, Z. & Fisher, R. A. Human RGS6 gene structure, complex alternative splicing, and role of N terminus and G protein gamma-subunit-like (GGL) domain in subcellular localization of RGS6 splice variants. *J. Biol. Chem.* **278**, 30261–30271. <https://doi.org/10.1074/jbc.M212687200> (2003).
17. Uhlen, M. *et al.* Proteomics. Tissue-based map of the human proteome. *Science* **347**, 1260419. <https://doi.org/10.1126/science.1260419> (2015).

18. Dohlman, H. G. & Thorner, J. RGS proteins and signaling by heterotrimeric G proteins. *J. Biol. Chem.* **272**, 3871–3874. <https://doi.org/10.1074/jbc.272.7.3871> (1997).
19. Pashkov, V. et al. Regulator of G protein signaling (RGS16) inhibits hepatic fatty acid oxidation in a carbohydrate response element-binding protein (ChREBP)-dependent manner. *J. Biol. Chem.* **286**, 15116–15125. <https://doi.org/10.1074/jbc.M110.216234> (2011).
20. Stewart, A. et al. Regulator of G protein signaling 6 is a critical mediator of both reward-related behavioral and pathological responses to alcohol. *Proc. Natl. Acad. Sci. USA* **112**, E786–795. <https://doi.org/10.1073/pnas.1418795112> (2015).
21. Perry, K. D. et al. Composite hemangioendothelioma with neuroendocrine marker expression: an aggressive variant. *Mod. Pathol.* **30**, 1589–1602. <https://doi.org/10.1038/modpathol.2017.83> (2017).
22. Meng, W., Deshmukh, H. A. & Donnelly, L. A. A genome-wide association study provides evidence of sex-specific involvement of Chr1p35.1 (ZSCAN20-TLR12P) and Chr8p23.1 (HMGB1P46) With Diabetic Neuropathic Pain. *EbioMedicine* **2**, 1386–1393. <https://doi.org/10.1016/j.ebiom.2015.08.001> (2015).
23. Zhang, D., Jiang, P., Xu, Q. & Zhang, X. Arginine and glutamate-rich 1 (ARGLU1) interacts with mediator subunit 1 (MED1) and is required for estrogen receptor-mediated gene transcription and breast cancer cell growth. *J. Biol. Chem.* **286**, 17746–17754. <https://doi.org/10.1074/jbc.M110.206029> (2011).
24. Sigurdsson, S. et al. Sequence variants in ARHGAP15, COLQ and FAM155A associate with diverticular disease and diverticulitis. *Nat. Commun.* **8**, 15789. <https://doi.org/10.1038/ncomms15789> (2017).
25. Lagace, T. A. PCSK9 and LDLR degradation: regulatory mechanisms in circulation and in cells. *Curr. Opin. Lipidol.* **25**, 387–393. <https://doi.org/10.1097/MOL.000000000000114> (2014).
26. White, K. E. et al. Autosomal-dominant hypophosphatemic rickets (ADHR) mutations stabilize FGF-23. *Kidney Int.* **60**, 2079–2086. <https://doi.org/10.1046/j.1523-1755.2001.00064.x> (2001).
27. Benet-Pages, A. et al. FGF23 is processed by proprotein convertases but not by PHEX. *Bone* **35**, 455–462. <https://doi.org/10.1016/j.bone.2004.04.002> (2004).
28. Kato, K. et al. Polypeptide GalNAc-transferase T3 and familial tumoral calcinosis. Secretion of fibroblast growth factor 23 requires O-glycosylation. *J. Biol. Chem.* **281**, 18370–18377. <https://doi.org/10.1074/jbc.M602469200> (2006).
29. Jones, E. Y., Fugger, L., Strominger, J. L. & Siebold, C. MHC class II proteins and disease: a structural perspective. *Nat. Rev. Immunol.* **6**, 271–282. <https://doi.org/10.1038/nri1805> (2006).
30. Feger, M. et al. The production of fibroblast growth factor 23 is controlled by TGF-beta2. *Sci. Rep.* **7**, 4982. <https://doi.org/10.1038/s41598-017-05226-y> (2017).
31. Qiu, T. et al. TGF-beta type II receptor phosphorylates PTH receptor to integrate bone remodelling signalling. *Nat. Cell. Biol.* **12**, 224–234. <https://doi.org/10.1038/ncb2022> (2010).
32. Zerr, P. et al. Vitamin D receptor regulates TGF-beta signalling in systemic sclerosis. *Ann. Rheum. Dis.* **74**, e20. <https://doi.org/10.1136/annrheumdis-2013-204378> (2015).
33. Subramaniam, N. et al. Cross-talk between 1,25-dihydroxyvitamin D3 and transforming growth factor-beta signaling requires binding of VDR and Smad3 proteins to their cognate DNA recognition elements. *J. Biol. Chem.* **276**, 15741–15746. <https://doi.org/10.1074/jbc.M011033200> (2001).
34. Leslie, R., O'Donnell, C. J. & Johnson, A. D. GRASP: analysis of genotype-phenotype results from 1390 genome-wide association studies and corresponding open access database. *Bioinformatics* **30**, i185–i194. <https://doi.org/10.1093/bioinformatics/btu273> (2014).
35. Need, A. C. et al. A genome-wide investigation of SNPs and CNVs in schizophrenia. *PLoS Genet.* **5**, e1000373. <https://doi.org/10.1371/journal.pgen.1000373> (2009).
36. Kathiresan, S. et al. Common variants at 30 loci contribute to polygenic dyslipidemia. *Nat. Genet.* **41**, 56–65. <https://doi.org/10.1038/ng.291> (2009).
37. Teslovich, T. M. et al. Biological, clinical and population relevance of 95 loci for blood lipids. *Nature* **466**, 707–713. <https://doi.org/10.1038/nature09270> (2010).
38. Lango Allen, H. et al. Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature* **467**, 832–838. <https://doi.org/10.1038/nature09410> (2010).
39. Stambolian, D. et al. Meta-analysis of genome-wide association studies in five cohorts reveals common variants in RBFOX1, a regulator of tissue-specific splicing, associated with refractive error. *Hum. Mol. Genet.* **22**, 2754–2764. <https://doi.org/10.1093/hmg/ddt116> (2013).
40. Pankratz, N. et al. Meta-analysis of Parkinson's disease: identification of a novel locus, RIT2. *Ann. Neurol.* **71**, 370–384. <https://doi.org/10.1002/ana.22687> (2012).
41. Deng, G. et al. The value of older donors' klotho level in predicting recipients' short-term renal function. *Med. Sci. Monit.* **24**, 7936–7943. <https://doi.org/10.12659/MSM.913274> (2018).
42. Farias-Basulto, A. et al. Circulating levels of soluble klotho and fibroblast growth factor 23 in diabetic patients and its association with early nephropathy. *Arch. Med. Res.* **49**, 451–455. <https://doi.org/10.1016/j.arcmed.2019.01.008> (2018).
43. Levey, A. S. et al. Using standardized serum creatinine values in the modification of diet in renal disease study equation for estimating glomerular filtration rate. *Ann. Intern. Med.* **145**, 247–254. <https://doi.org/10.7326/0003-4819-145-4-200608150-00004> (2006).
44. Chen, C. H. et al. Population structure of Han Chinese in the modern Taiwanese population based on 10,000 participants in the Taiwan Biobank project. *Hum. Mol. Genet.* **25**, 5321–5331. <https://doi.org/10.1093/hmg/ddw346> (2016).
45. Purcell, S. et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575. <https://doi.org/10.1086/519795> (2007).
46. Delaneau, O., Zagury, J. F. & Marchini, J. Improved whole-chromosome phasing for disease and population genetic studies. *Nat. Methods* **10**, 5–6. <https://doi.org/10.1038/nmeth.2307> (2013).
47. Howie, B., Fuchsberger, C., Stephens, M., Marchini, J. & Abecasis, G. R. Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. *Nat. Genet.* **44**, 955–959. <https://doi.org/10.1038/ng.2354> (2012).
48. Marchini, J. & Howie, B. Genotype imputation for genome-wide association studies. *Nat. Rev. Genet.* **11**, 499–511. <https://doi.org/10.1038/nrg2796> (2010).
49. Danecek, P. et al. The variant call format and VCFtools. *Bioinformatics* **27**, 2156–2158. <https://doi.org/10.1093/bioinformatics/btr330> (2011).
50. Péter, I., Yelensky, R., Altshuler, D. & Daly, M. J. Estimation of the multiple testing burden for genomewide association studies of nearly all common variants. *Genet. Epidemiol.* **32**, 381–385. <https://doi.org/10.1002/gepi.20303> (2008).
51. Fallin, M. D. & Kao, W. H. Is, “X”-WAS the future for all of epidemiology?. *Epidemiology* **22**, 457–459. <https://doi.org/10.1097/EDE.0b013e31821d3a9f> (2011).
52. Duggal, P., Gillanders, E. M., Holmes, T. N. & Bailey-Wilson, J. E. Establishing an adjusted p-value threshold to control the family-wide type 1 error in genome wide association studies. *BMC Genom.* **9**, 516. <https://doi.org/10.1186/1471-2164-9-516> (2008).
53. Turner, S. *qqman: An R Package for Visualizing GWAS Results Using Q-Q and Manhattan Plots.* (2014).
54. Team, R. D. C. R: *A Language and Environment for Statistical Computing.* (2008).
55. Pruim, R. J. et al. LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics* **26**, 2336–2337. <https://doi.org/10.1093/bioinformatics/btq419> (2010).

56. Bulik-Sullivan, B. K. *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.* **47**, 291–295. <https://doi.org/10.1038/ng.3211> (2015).
57. Shim, H. *et al.* A multivariate genome-wide association analysis of 10 LDL subfractions, and their response to statin treatment, in 1868 Caucasians. *PLoS ONE* **10**, e0120758. <https://doi.org/10.1371/journal.pone.0120758> (2015).

Acknowledgements

This work is supported by grants from the Ministry of Science and Technology in Taiwan (MOST 105-2314-B-182-062, MOST 106-2314-B-182-043), the Translational Medical Research Program of Academia Sinica (AS-TM-108-01-04), National Taiwan University Hospital, Yunlin Branch (NTUHYL106.X003) and Chang Gung University, Taoyuan, Taiwan (NMRPD1F154, NMRPD1G0711, and BMRPD08). The authors would also like to thank the Data Science Statistical Cooperation Center of Academia Sinica (AS-CFII-108-117) for statistical support. The manuscript has been edited by the English editing service of K. T. Li Foundation for Development of Science and Technique, Taiwan (no. 2804_20200708121153).

Author contributions

G.T.C., P.H.L. and Y.C.C. conceived and designed the study. Analysis was done by G.T.C., P.H.L., C.R.C., C.N.H. and C.Y.S. Sample procurement and data generation was done by G.T.C., P.H.L., T.W.C., C.H.H., Y.Y.H., C.H.L., J.W.L., C.N.H., R.Y.T., M.L.H., H.L.L., W.S.Y. and Y.C.C. Manuscript writing was done by G.T.C., P.H.L. and Y.C.C.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-020-71569-8>.

Correspondence and requests for materials should be addressed to Y.-C.C.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020