<table>
<thead>
<tr>
<th>Title</th>
<th>An Exome-Chip Association Analysis in Chinese Subjects Reveals a Functional Missense Variant of GCKR That Regulates FGF21 Levels</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author(s)</td>
<td>Cheung, YY; Tang, SM; Xu, A; Lee, CHP; Au, KW; Xu, L; Fong, HY; Kwok, HM; Chow, WS; Woo, YC; Yuen, MAM; Cherny, SS; Hai, SHJJ; Cheung, BMY; Tan, KCB; Lam, TH; Tse, HF; Sham, PC; Lam, KSL</td>
</tr>
<tr>
<td>Citation</td>
<td>Diabetes, 2017, v. 66 n. 6, p. 1723-1728</td>
</tr>
<tr>
<td>Issued Date</td>
<td>2017</td>
</tr>
<tr>
<td>URL</td>
<td><a href="http://hdl.handle.net/10722/241505">http://hdl.handle.net/10722/241505</a></td>
</tr>
<tr>
<td>Rights</td>
<td>This is an author-created, uncopyedited electronic version of an article accepted for publication in Diabetes. The American Diabetes Association (ADA), publisher of Diabetes, is not responsible for any errors or omissions in this version of the manuscript or any version derived from it by third parties. The definitive publisher-authenticated version is available online at <a href="http://diabetes.diabetesjournals.org/content/66/6/1723">http://diabetes.diabetesjournals.org/content/66/6/1723</a>; This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.</td>
</tr>
</tbody>
</table>
An exome-chip association analysis in Chinese reveals a functional missense variant of \textit{GCKR} that regulates FGF21 levels
An exome-chip association analysis in Chinese reveals a functional missense variant of *GCKR* that regulates FGF21 levels

Running title: Exome-chip association analysis on FGF21 levels

Chloe YY Cheung¹, Clara S Tang², Aimin Xu¹,³,⁴,⁵, Chi-Ho Lee¹, Ka-Wing Au¹, Lin Xu⁶, Carol HY Fong¹, Kelvin HM Kwok¹, Wing-Sun Chow¹, Yu-Cho Woo¹, Michele MA Yuen¹, Stacey S Cherny⁷, JoJo Hai¹, Bernard MY Cheung¹, Kathryn CB Tan¹, Tai-Hing Lam⁶, Hung-Fat Tse¹,⁸*, Pak-Chung Sham⁷,⁹,¹⁰*, Karen SL Lam¹,³,⁴*

¹Department of Medicine, the University of Hong Kong, Hong Kong, China; ²Department of Surgery, the University of Hong Kong, Hong Kong, China; ³State Key Laboratory of Pharmaceutical Biotechnology, The University of Hong Kong, Hong Kong, China; ⁴Research Centre of Heart, Brain, Hormone and Healthy Aging, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong, China; ⁵Department of Pharmacology & Pharmacy, The University of Hong Kong, Hong Kong, China; ⁶School of Public Health, the University of Hong Kong, Hong Kong, China; ⁷Department of Psychiatry, the University of Hong Kong, Hong Kong, China; ⁸Hong Kong-Guangdong Joint Laboratory on Stem Cell and Regenerative Medicine, the University of Hong Kong, Hong Kong, China; ⁹Centre for Genomic Sciences, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong, China; ¹⁰State Key Laboratory in Brain and Cognitive Sciences, The University of Hong Kong, Hong Kong, China.

*KSL Lam, PC Sham and HF Tse contributed equally to the supervision of this work and are co-corresponding authors.
Abstract (195 words)

Fibroblast growth factor 21 (FGF21) is increasingly recognized as an important metabolic regulator of glucose homeostasis. Here, we conducted an exome-chip association analysis by genotyping 5169 Chinese individuals from a community-based cohort and two clinic-based cohorts. A custom Asian Exome-chip was used to detect genetic determinants influencing circulating FGF21 levels. Single-variant association analysis interrogating 70,444 single nucleotide polymorphisms identified a novel locus, GCKR, significantly associated with circulating FGF21 levels at genome-wide significance. In the combined analysis, the common missense variant of GCKR, rs1260326 (p.Pro446Leu), showed an association with FGF21 levels after adjustment for age and sex ($P=1.61\times10^{-12}$; $\beta[SE]:0.14[0.02]$), which remained significant on further adjustment for body mass index ($P=3.01\times10^{-14}$; $\beta[SE]:0.15[0.02]$). GCKR Leu446 may influence FGF21 expression via its ability to increase glucokinase (GCK) activity. This can lead to enhanced FGF21 expression via elevated fatty acid synthesis, consequent to the inhibition of carnitine/palmitoyl-transferase by malonyl-CoA; and via increased glucose-6-phosphate mediated activation of the carbohydrate response element binding protein, known to regulate FGF21 gene expression. Our findings shed new light on the genetic regulation of FGF21 levels. Further investigations to dissect the relationship between GCKR and FGF21, with respect to the risk of metabolic diseases, are warranted.
Introduction

Fibroblast growth factor 21 (FGF21) is a circulating metabolic hormone predominantly secreted from the liver and from other tissues, such as adipose tissue, pancreas and skeletal muscle (1; 2). Numerous animal studies have demonstrated its favorable effects on insulin sensitivity, glucose and lipid metabolism, and body weight, in obese mice and diabetic monkeys (1-5). High circulating levels of FGF21, however, have been observed in patients with obesity and a range of obesity-related conditions, including type 2 diabetes (T2DM), coronary artery disease (CAD) and the metabolic syndrome (MetS) (1). Our group previously showed that elevated levels of FGF21 independently predicted the development of T2DM (6), carotid atherosclerosis (7), and kidney disease progression in T2DM subjects (8). Observations of the paradoxical increase of FGF21 in the above diseases may represent a compensatory response to FGF21 resistance, which has been demonstrated in a previous study of mice with diet-induced obesity (9), or to combat the metabolic disturbances in these disease conditions.

The circulating levels of FGF21 have been reported to be moderately heritable in a twin study of young adults, with 40% of variation attributed to genetic determinants (10). So far, only a few genetic variants have been reported to show weak associations with circulating FGF21 levels (11; 12). No previous genome-wide or exome-wide association analyses on circulating FGF21 levels have been published to date. The heritability of circulating FGF21 levels remains largely unexplained. This study sought to identify the genetic determinants of circulating FGF21 levels. We conducted an exome-chip association analysis on circulating FGF21 levels using a specially designed Illumina HumanExome BeadChip (Asian Exome-chip) in a Chinese population (13; 14). In consideration of the possible difference in circulating FGF21 levels between individuals from the patient groups.
and those from the general population, possibly due to the intrinsic metabolic stress under
disease conditions and the treatment effects on FGF21 expression, we first conducted the
association analysis in a community-based and two clinic-based cohorts separately, followed
by a combined analysis involving a total of 5169 individuals.

Research Design and Methods

Subjects
An exome-chip association study on circulating FGF21 levels was performed in 5169
Chinese individuals who were recruited from the Hong Kong Cardiovascular Risk Factor
Prevalence Study (CRISPS; n=1172); the Hong Kong West Diabetes Registry (HKWDR;
n=2884); and the Hong Kong Chinese Coronary Artery Disease (HK-CAD) study (n=1113).
Details of the study cohorts and measurement of FGF21 levels were described in the
Supplementary notes.

Genotyping and data quality control
All subjects were genotyped using the custom Asian Exome-chip (13; 14) with an add-on
content of 58,317 single nucleotide polymorphisms (SNPs), including a custom panel of over
30,000 missense or nonsense coding variants on top of the standard content of the Infinium
HumanExome BeadChip (HumanExome-12v1_A; Illumina, CA). Genotyping of the Exome
array was performed using the Illumina iScan system platform at the Centre for Genomic
Sciences of the University of Hong Kong. The GenTrain version 2.0 in GenomeStudio
V2011.1 (Illumina) was used to perform genotyping calling. Details of genotype quality
control were described in the Supplementary notes. After all QC measures, a total of 1172
subjects and 76,389 variants remained in the association analysis of the CRISPS cohort,
whilst the analyses of the clinic-based cohorts included a total of 2884 subjects and 77598 variants from the HKWDR; and 1113 subjects and 75468 variants from the HK-CAD study. The combined analysis included only polymorphic SNPs in all three cohorts and had MAF $\geq 0.1\%$, totaling 5169 subjects and 70,444 variants. The $P$-value informed linkage disequilibrium (LD) based clumping approach with the “--clump” command implemented in PLINK was performed to address the between-SNP LD. Each index SNP represented the strongest association $P$-value from each clumped region. Each index SNP formed clumps with other variants which were within $\pm$ 500kb from the index SNP and in LD with the index SNP ($r^2 \geq 0.2$).

**Statistical analysis**

For each cohort, we carried out single-variant association tests, under the additive genetic model, for all markers that passed quality controls using PLINK version 1.9 (15). The FGF21 levels were first transformed to rank-based inverse normal residuals before analysis to ensure normality and minimize the possible effect of outliers. Age, sex and the first 2 principal components (PCs) were included as covariates in the multiple linear regression model (Model 1) (Supplementary Figure 1). To assess adiposity independent association, body mass index (BMI) was included as an additional covariate in Model 2. Meta-analysis of the association results of the three cohorts was conducted using GWAMA (http://www.geenivaramu.ee/en/tools/gwama; accessed 20 Dec 2016) (16). The inverse variance fixed-effect method was employed to meta-analyze the summary statistics of the three cohorts. The heterogeneity of effect was assessed using Cochran’s $Q$-test and $I^2$ index. All 70,444 SNPs showed no evidence of heterogeneity in the combined analysis with a Cochran’s $Q$-test $P$-value $\geq 1 \times 10^{-6}$ (a value selected to take into account the multiple testing and the stringent Bonferroni correction) (17). Exome-wide significance was defined as $P$
<7.10x10^{-7} (=0.05/70,444). As visualized in a Quantile-Quantile plot (Figure 1), the test statistics appeared well-calibrated. Associations of GCKR rs1260326 with fasting plasma glucose (FPG) and triglyceride (TG) were examined by multiple linear regression analyses in non-T2DM individuals and in subjects not taking lipid lowering medications, respectively, with adjustment for age, sex, BMI, PC1 and PC2.

**Results**

**Exome-chip association study for circulating FGF21 levels**

Of the 70,444 polymorphic SNPs (Figure 2) examined for associations with circulating FGF21 levels in 5169 Chinese individuals (Table 1), 45.06% altered protein composition and 19.96% were Asian-specific variants with MAF 0.1-5%.

The strongest association was detected at a missense variant rs1260326 (p.Pro446Leu) of GCKR (Table 2 and Supplementary Figure 2). The T (Leu446) allele of rs1260326 was significantly associated with a higher level of FGF21 levels at genome-wide significance (Model 1: $P_{\text{combined}}=1.61\times10^{-12}$; $\beta[SE]: 0.14[0.02]$; $P_{\text{heterogeneity}}=0.018$; $I^2=0.75$; Model 2: $P_{\text{combined}}=3.01\times10^{-14}$; $\beta[SE]: 0.15[0.02]$; $P_{\text{heterogeneity}}=0.041$; $I^2=0.69$; Table 2; Figure 2). GCKR rs1260326 was significantly associated with TG ($P<0.001$; $\beta[SE]:0.08[0.005]$). A modest association between rs1260326 and FPG ($P=0.013$; $\beta[SE]:-0.01[0.004]$) was also observed. The associations between rs1260326 with FGF21 levels remained significant when FPG ($P_{\text{combined}}=1.60\times10^{-14}$; $\beta[SE]:0.16[0.02]$) or TG ($P_{\text{combined}}=1.26\times10^{-7}$; $\beta[SE]:0.11[0.02]$) or both ($P_{\text{combined}}=5.42\times10^{-7}$; $\beta[SE]:0.10[0.02]$) were included in the adjustment models. The proportion of variance in FGF21 levels explained by rs1260326 was estimated to be 0.97%. 

For Peer Review Only

Page 7 of 47  
*Diabetes*
In the combined analysis, we also observed suggestive associations of an intronic SNP rs17817964 at *FTO* (*P*<sub>combined</sub>=1.44x10^-6; β[SE]:0.13[0.03]; *P*<sub>heterogeneity</sub>=0.690; *I*^2^=0) and a missense variant rs2273983 (p.Asn719Ser) at *TSR1* (*P*<sub>combined</sub>=4.61x10^-6; β[SE]:0.48[0.10]; *P*<sub>heterogeneity</sub>=0.590; *I*^2^=0) with FGF21 levels in adjustment Model 1. The associations of these two variants were slightly attenuated on further adjustment for BMI (Table 2).

Supplementary table 1 shows the 27 index SNPs which showed associations with FGF21 level at *P*<sub>combined</sub> <5x10^-4 in the combined analysis.

**Discussion**

The current study reports the first exome-chip association analysis on circulating FGF21 levels. By genotyping 5169 Chinese individuals using a custom Asian Exome-chip, we identified a genome-wide significant association between *GCKR* rs1260326, a proline to leucine substitution at amino acid 446 (p.Pro446Leu), and circulating FGF21 levels. Our study has shed light on the possible role of *GCKR* on the regulation of circulating FGF21 levels. *GCKR*, encoding the glucokinase regulator (also known as glucokinase regulatory protein [GKRP]), has been implicated in a wide range of important metabolic pathways, such as glucose homeostasis and lipid metabolism (18). Both *FGF21* and *GCKR* are highly expressed in the liver (1; 18). *GCKR* acts as a competitive inhibitor of glucokinase (GCK), the principal modulator of glucose uptake and release in the liver (18). *GCKR* forms an inhibitory complex with GCK (GCKR-GCK complex), and allosterically controls its activity in the presence of fructose. Dissociation of the GCKR-GCK complex facilitates the translocation of GCK to cytoplasm where hepatic glycolysis is stimulated (18).

The Pro446 residue of GCKR is conserved across various species, including human and rat. *GCKR* p.Pro446Leu is located in close proximity to Asp413 and Gln443 which form a
critical salt bridge with Arg186 of GCK (19). Thus, the variant could interfere directly with the binding of GCKR to GCK (19), thereby leading to reduced inhibition of GCK. This in turn contributed to an enhanced glycolytic flux and promoted uptake of glucose into the liver. Accompanying this increased rate of glycolysis is the raised levels of other liver metabolites, such as malonyl-CoA which ultimately lead to enhanced production of TG (20). This explains the reduced FPG and raised TG levels observed in individuals carrying the Leu446 (T) allele as previously reported by the French longitudinal cohort study DESIR (Data from an Epidemiological Study on the Insulin Resistance syndrome) (21) and in the current study.

GCKR Leu446 may influence FGF21 expression via its ability to increase GCK activity. This leads to elevated phosphorylation of glucose to glucose-6-phosphate (G6P) and the subsequent inhibition of carnitine/palmitoyl-transferase, the rate limiting enzyme of β-oxidation, by malonyl-CoA (22). In turn, such changes may result in enhanced FGF21 expression via increased G6P mediated activation of carbohydrate response element binding protein (ChREBP) (23); and via elevated fatty acid synthesis (1), which were known to regulate FGF21 gene expression.

GCKR has been recognized as a highly pleiotropic gene (1; 2). Common variants of GCKR, including rs1260326 or its intronic proxy SNP rs780094 ($r^2=0.92$), have been shown to be associated with multiple metabolic traits, including FPG and TG (2; 13; 21). The current study further demonstrated that rs1260326 may play a role in the regulation of FGF21 levels which is independent from the effect of BMI, FPG and TG. Future functional studies elucidating the role of GCKR in the in vivo regulation of FGF21 levels are warranted.

Recent clinical trials have demonstrated that treatments with FGF21 analogs in obese patients with T2DM resulted in obvious improved lipid profiles, lowered fasting insulin and body
weight, similar to what have been observed in animals (24; 25). In particular, FGF21 treatment has demonstrated a rapid and prominent reduction in circulating TG, which occurred as soon as two days after treatment (24). However, there was lack of significant effect on glycemic control in humans (24; 25), contradictory to the glucose lowering effect of FGF21 observed in animals (4; 5). In line with these reports, the GCKR rs1260326 was found to be strongly associated with TG but only modestly correlated with FPG in the current study (24; 25). Taken together, these findings suggest that FGF21 may play a more prominent role in regulating lipid metabolism and adiposity, rather than blood glucose. On the other hand, our finding that rs1260326 could possibly influence the circulating levels of FGF21 via enhanced GCK activity, also suggest that alterations in hepatic carbohydrate metabolism can impact on FGF21 expression and its circulating levels.

The major limitation of the current study was the lack of external validations. Further replication analysis in independent cohorts in other Asians or European populations would serve to validate our findings. The design of the exome-chip has allowed us to examine more coding variants with much lower MAF compared to conventional genome-wide association studies. However, the relatively small sample size has limited our study power to detect the association of rare or low-frequency variants. With a view to achieve a larger sample size, a meta-analysis of genome/exome-wide association studies would be useful to identify additional genetic variants influencing FGF21 levels.

In summary, we conducted an exome-chip association analysis and identified, for the first time, a highly significant association of a functional variant of GCKR, rs1260326 (p.Pro446Leu), with circulating FGF21 levels, independent of obesity and other metabolic
traits. Our findings further highlighted the pleiotropic role of GCKR and provided insights into the regulation of FGF21 levels and its relationship with metabolic diseases.

**Contribution statement**

KSLL, PCS and HFT conceived the study and undertook project leadership. CYYC wrote the first draft of the manuscript. CYYC and CST analysed and interpreted the data. PCS, CST and SSC provided useful comments to data-analysis. CYYC, AX, CHL, KWA, LX, CHYF, KHMK, WSC, YCW, MMA, JSHH, BMYC, KCBT and THL were involved in the sample collection, selection and phenotype data preparation. KSLL and HFT were involved in the database management for the study cohorts. All authors contributed to the drafting and critical revision of the manuscript. All authors approved the final version of the manuscript.

**Acknowledgements**

The authors thank all the study participants, clinical and research staffs of CRISPS, HKWDR and the Hong Kong Chinese CAD study cohorts for their contribution in this research study.

**Funding**

This work was supported by the Hong Kong Research Grant Council: Theme Based Research Scheme (T12-705/11) and Collaborative Research Fund (HKU2/CRF/12R).

**Guarantor Statement**

Professor Karen SL Lam, Professor Pak-Chung Sham and Professor Hung-Fat Tse are the guarantors of this work and, as such, had full access to the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.
Duality of interest

No potential conflicts of interest relevant to this article were reported.


15. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC: PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 2007;81:559-575


Table 1. Clinical characteristics of subjects in the community-based and the two clinic-based cohorts.

<table>
<thead>
<tr>
<th>Variables</th>
<th>CRISPS</th>
<th>HKWDR</th>
<th>P-value gauche</th>
<th>HK-CAD</th>
<th>P-value gauche</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>1172</td>
<td>2884</td>
<td>-</td>
<td>1113</td>
<td>-</td>
</tr>
<tr>
<td>Age</td>
<td>49.82 ± 10.68</td>
<td>60.87 ± 12.24</td>
<td>&lt;0.001</td>
<td>67.32 ± 10.44</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Sex (Male%)</td>
<td>43.7</td>
<td>59.1</td>
<td>&lt;0.001</td>
<td>75.1</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>24.02 ± 3.48</td>
<td>26.00 ± 4.27</td>
<td>&lt;0.001</td>
<td>25.25 ± 3.65</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>T2DM (%)</td>
<td>12.8</td>
<td>100</td>
<td>&lt;0.001</td>
<td>38.3</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>CAD (%)</td>
<td>0</td>
<td>36.7</td>
<td>&lt;0.001</td>
<td>100</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>FGF21* (pg/mL)</td>
<td>154.72(83.90-266.80)</td>
<td>161.00(82.73-287.23)</td>
<td>0.237</td>
<td>241.38(149.71-376.50)</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Data as mean ± standard deviation or median with interquartile range. N: number; BMI: body mass index; CAD: coronary artery disease; T2DM: type 2 diabetes; FGF21: fibroblast growth factor 21. CRISPS: Hong Kong Cardiovascular Risk Factor Prevalence Study; HKWDR: Hong Kong West Diabetes Registry; HK-CAD: Hong Kong Chinese Coronary Artery Disease study. *Natural-log-transformed before analysis. †HKWDR versus CRISPS. ‡HK-CAD versus CRISPS.
Table 2. Variants associated with circulating FGF21 levels at P < 1x10^{-5} in the combined analysis.

<table>
<thead>
<tr>
<th>Gene</th>
<th>SNP</th>
<th>Position</th>
<th>A1/A2 (CRISPS; n=1172)</th>
<th>Community-based</th>
<th>(SE)</th>
<th>P-value*</th>
<th>Clinic-based (HKWDR; n=2884)</th>
<th>(SE)</th>
<th>P-value*</th>
<th>(HK-CAD; n=1113)</th>
<th>Beta</th>
<th>(SE)</th>
<th>P-value*</th>
<th>Combined (CRISPS+HKWDR+HK-CAD; n=5169)</th>
<th>Beta</th>
<th>(SE)</th>
<th>P-value*</th>
<th>P-value†</th>
</tr>
</thead>
<tbody>
<tr>
<td>GCKR</td>
<td>rs1260326</td>
<td>2:27730940</td>
<td>T/C</td>
<td>0.22 (0.04)</td>
<td>3.31x10^{-5}</td>
<td>0.13 (0.03)</td>
<td>5.19x10^{-7}</td>
<td>0.06 (0.04)</td>
<td>0.184 (0.02)</td>
<td>0.455</td>
<td>0.14</td>
<td>1.61x10^{-12}</td>
<td>3.01x10^{-14}</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FTO</td>
<td>rs17817964</td>
<td>16:53828066</td>
<td>T/C</td>
<td>0.09 (0.06)</td>
<td>0.139 (0.04)</td>
<td>0.14 (0.04)</td>
<td>1.42x10^{-4}</td>
<td>0.16 (0.06)</td>
<td>7.01x10^{-3}</td>
<td>0.158</td>
<td>0.13</td>
<td>1.44x10^{-6}</td>
<td>7.59x10^{-5}</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TSRI</td>
<td>rs2273983</td>
<td>17:2227988</td>
<td>C/T</td>
<td>0.69 (0.24)</td>
<td>3.55x10^{-3}</td>
<td>0.45 (0.13)</td>
<td>8.27x10^{-4}</td>
<td>0.37 (0.24)</td>
<td>0.127 (0.10)</td>
<td>0.009</td>
<td>0.48</td>
<td>4.61x10^{-6}</td>
<td>4.90x10^{-6}</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

A1: Minor allele; A2: Major allele; MAF: Minor allele frequency; SE: Standard error. CRISPS: Hong Kong Cardiovascular Risk Factor Prevalence Study; HKWDR: Hong Kong West Diabetes Registry; HK-CAD: Hong Kong Chinese Coronary Artery Disease study. The betas are reported relative to the minor allele. *Model 1: Adjusted for age, sex, PC1 and PC2. †Model 2: Adjusted for age, sex, BMI, PC1 and PC2. Chromosomal positions are presented according to human reference genome hg19.
Figure legend

Figure 1. Quantile-Quantile plot.
Quantile-Quantile plot of association $P$-values of all tested variants for FGF21 level.

Figure 2. Manhattan plot of the combined analysis association results.
The y-axis represents the -log10 $P$ value, and the x-axis represents the genomic position. The dots represent the 70,444 SNPs analyzed, relative to their position on each chromosome (alternating black and grey). The black horizontal dashed line indicates genome-wide significance ($P = 5 \times 10^{-8}$). The diamond symbol indicates the most significantly associated SNP GCKR rs1260326 (p.Pro446Leu).
Figure 1. Quantile-Quantile plot.
Quantile-Quantile plot of association P-values of all tested variants for FGF21 level.

88x91mm (300 x 300 DPI)
Figure 2. Manhattan plot of the combined analysis association results.

The y-axis represents the -log10 P value, and the x-axis represents the genomic position. The dots represent the 70,444 SNPs analyzed, relative to their position on each chromosome (alternating black and grey). The black horizontal dashed line indicates genome-wide significance (P =5x10^{-8}). The diamond symbol indicates the most significantly associated SNP GCKR rs1260326 (p.Pro446Leu).
Supplementary notes

Subjects

All subjects of the current study were previously examined in the exome-array association studies of T2DM (1), blood lipid traits and CAD (2). Only those subjects who also had their blood samples available for measurement of circulating FGF21 levels were included in the present study. All participants gave written informed consent and the study protocol was approved by the Institutional Review Board of the University of Hong Kong/Hospital Authority Hong Kong West Cluster.

(i) Community-based cohort:

The Hong Kong Cardiovascular Risk Factor Prevalence Study (CRISPS): The community-based cohort included a total of 1172 Southern Han Chinese subjects from the Hong Kong Cardiovascular Risk Factor Prevalence Study (CRISPS) cohort (3), who were the non-CAD control subjects in our previous exome-chip association analysis on blood lipid traits and CAD (2). Details of the CRISPS cohort have been described previously (1; 2; 4). Briefly, CRISPS was first commenced as a population-based survey in 1995-1996. A total of 2895 unrelated Chinese subjects were randomly selected by their telephone numbers from the Hong Kong population. All participants were invited to undergo a comprehensive baseline assessment of cardiovascular risks at the Queen Mary Hospital, Hong Kong. The subjects were then invited for subsequent prospective follow-up visits to assess for the development of major cardiovascular risk factors, in 2000-2004 (CRISPS2), 2005-2008 (CRISPS3) and 2010-2012 (CRISPS4). The latest follow-up visit (CRISPS5) has been commenced in 2016. At each assessment, anthropometric and demographic were collected. The participants’ medical, treatment and family histories of the major cardiovascular risk factors were recorded by a
detailed questionnaire. Fasting venous blood were collected after an overnight fast for the measurement of glucose, lipids and biomarkers levels. FGF21 levels were measured in plasma samples collected at the second assessment (CRISPS2). A 75g oral glucose tolerance test (OGTT) was performed in all subjects who were not on treatment for diabetes.

(ii) Clinic-based cohorts:

The clinic-based cohort included a total of 3997 individuals recruited from the Hong Kong West Diabetes Registry (HKWDR) (5) and the Hong Kong Chinese Coronary Artery Disease (HK-CAD) study cohort (2), respectively.

The Hong Kong West Diabetes Registry cohort (HKWDR; n=2884): The HKWDR was commenced in 2008. Participants of HKWDR were type 2 diabetes (T2DM) patients who were on regular follow-up at the medical specialist clinics of the Hong Kong West Cluster. The participants were invited to undergo comprehensive clinical assessments and laboratory investigations to determine the presence of chronic diabetic complications and their control of diabetes. The anthropometric and demographic data of the study subjects were collected at each assessment. The subjects’ medical, treatment and family histories of cardiovascular risk factors were recorded with a detailed questionnaire. Venous blood samples were drawn, with informed consent, after an overnight fast, for biochemical and DNA analyses. FGF21 levels were measured from the stored serum samples collected at the first assessment.

The Hong Kong Chinese Coronary Artery Disease study cohort (HK-CAD; n=1113): The HK-CAD study cohort is an on-going prospective cohort study on the clinical outcomes and risk factors in Chinese patients with established CAD. This cohort was first initiated in the Queen Mary Hospital, Hong Kong, in 2004-2005. Briefly, consecutive patients who
underwent invasive coronary angiogram for assessment and treatment of CAD were screened. The severity of stenosis was determined by coronary angiogram. Patients who suffered from significant CAD with ≥50% stenosis in one or more of the epicardial coronary artery were invited to participate in this cohort. Detailed demographic and anthropometric data, such as major cardiovascular risk factors; and drug, family and medical histories were collected during their hospital admission or in the outpatient clinic follow-up. Blood samples were drawn from the participants after an overnight fast, with written informed consent, for biochemical and genetic analyses. FGF21 levels were measured from the plasma samples collected at recruitment.

Measurement of circulating FGF21 levels

Blood FGF21 levels of subjects were measured using a human FGF21 enzyme-linked immunosorbent assay (ELISA) kit according to the manufacturer’s protocol (Antibody and Immunoassay Services, University of Hong Kong) as previously described (6). Samples were diluted 1:1 before the assay was conducted. The 100ul diluted samples, calibrators, and quality controls were added to the 96-well microtiter plates coated with an affinity-purified polyclonal anti-human FGF21 antibody. A calibration curve was constructed by plotting the absorbance values at 450nm against FGF21 concentrations of the calibrators. The FGF21 concentrations in the study samples were then determined by using the calibration curve. The intra- and inter-assay coefficients of variation of the FGF21 ELISA were 4 to 5% and 3.5 to 10.2%, respectively.

Quality control of genotype data

Manual curation of genotype calls was conducted for over 55,000 variants which had a GenTrain score of less than 0.8; or with high missingness of greater than 1%; or were shown
to have poor genotype clustering in exome chip genotyping of more than 9000 subjects by our collaborators (7). A total of 4550 variants were excluded from further analysis due to poor genotype clustering. Individual-level quality control (QC) was conducted with regard to biological relatedness, duplication, gender mismatch, and possible sample contamination. For detection of possible existence of non-Chinese samples, a principal component (PC) analysis was conducted using a panel of >20,000 independent common SNPs (minor allele frequency [MAF] >0.05) with outliers removed from the analysis. A scree plot showing the eigenvalues, which represents the relative proportion of variance explained of the first 15 principal components, is presented in Supplementary Figure 1. For SNP-level QC, variants with more than 2% missingness; or MAF of less than 0.1%; or showed significant deviation from Hardy-Weinberg Equilibrium (HWE) with P <1x10^{-5}; or originally designed for the purpose of QC (e.g. fingerprint SNPs for sample tracking and grid SNPs for the identification of identity by descent segments) were excluded from the analysis.

References

4
Supplementary Figure 1. Scree plot

A scree plot showing the amount of variation explained by the first 15 principal components (PCs). The first two PCs have shown statistical significance with \( P \)-values of less than 0.05 in the Tracy-Widom test.
Supplementary Figure 2. Regional plot

The regional plot illustrating individual SNPs at the GCKR locus and their association to circulating FGF21 levels.
Supplementary Table 1. Index SNPs with $P_{combined} < 5 \times 10^{-4}$ in the combined analysis for association with FGF21 levels.

<table>
<thead>
<tr>
<th>Gene</th>
<th>SNP</th>
<th>Position</th>
<th>A1/A2</th>
<th>Annotation</th>
<th>MAF</th>
<th>Beta (SE)</th>
<th>P-value</th>
<th>Model 1</th>
<th>Beta (SE)</th>
<th>P-value</th>
<th>Model 2</th>
<th>Beta (SE)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>GCKR</td>
<td>rs1260326</td>
<td>2:27730940</td>
<td>T/C</td>
<td>p.Pro446Leu</td>
<td>0.455</td>
<td>0.14 (0.02)</td>
<td>1.61x10^{-12}</td>
<td></td>
<td>0.15 (0.02)</td>
<td>3.01x10^{-14}</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FTO</td>
<td>rs17817964</td>
<td>16:53828066</td>
<td>T/C</td>
<td>intronic</td>
<td>0.158</td>
<td>0.13 (0.03)</td>
<td>1.44x10^{-6}</td>
<td></td>
<td>0.11 (0.03)</td>
<td>7.59x10^{-5}</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TSR1</td>
<td>rs2273983</td>
<td>17:2227988</td>
<td>C/T</td>
<td>p.Asn719Ser</td>
<td>0.009</td>
<td>0.48 (0.10)</td>
<td>4.61x10^{-6}</td>
<td></td>
<td>0.49 (0.11)</td>
<td>4.90x10^{-6}</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EARS2</td>
<td>rs200139797</td>
<td>16:23563485</td>
<td>C/T</td>
<td>p.Met94Val</td>
<td>0.004</td>
<td>0.64 (0.15)</td>
<td>1.69x10^{-5}</td>
<td></td>
<td>0.64 (0.15)</td>
<td>1.70x10^{-5}</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GHRHR</td>
<td>rs144372265</td>
<td>7:31009485</td>
<td>T/C</td>
<td>p.Ala91Val</td>
<td>0.004</td>
<td>0.68 (0.16)</td>
<td>2.90x10^{-5}</td>
<td></td>
<td>0.62 (0.17)</td>
<td>2.41x10^{-4}</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MRPL33</td>
<td>rs3792252</td>
<td>2:27995931</td>
<td>A/G</td>
<td>intronic</td>
<td>0.417</td>
<td>0.08 (0.02)</td>
<td>3.83x10^{-5}</td>
<td></td>
<td>0.07 (0.02)</td>
<td>7.46x10^{-4}</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C7</td>
<td>rs2271708</td>
<td>5:40936541</td>
<td>C/T</td>
<td>p.Cys128Arg</td>
<td>0.033</td>
<td>0.22 (0.05)</td>
<td>7.94x10^{-5}</td>
<td></td>
<td>0.23 (0.06)</td>
<td>6.61x10^{-5}</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ENTPD4</td>
<td>rs190514217</td>
<td>8:23291891</td>
<td>T/C</td>
<td>p.Asp521Asn</td>
<td>0.004</td>
<td>-0.58 (0.15)</td>
<td>1.01x10^{-4}</td>
<td></td>
<td>-0.55 (0.15)</td>
<td>3.64x10^{-4}</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MBNL2</td>
<td>rs7320865</td>
<td>13:97976442</td>
<td>G/A</td>
<td>intronic</td>
<td>0.495</td>
<td>-0.08 (0.02)</td>
<td>1.02x10^{-4}</td>
<td></td>
<td>-0.07 (0.02)</td>
<td>7.39x10^{-4}</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CDH23</td>
<td>rs10999947</td>
<td>10:73434906</td>
<td>A/G</td>
<td>p.Ser496Asn</td>
<td>0.305</td>
<td>-0.08 (0.02)</td>
<td>1.33x10^{-4}</td>
<td></td>
<td>-0.08 (0.02)</td>
<td>1.36x10^{-4}</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CETN3-MBLAC2</td>
<td>rs6890115</td>
<td>5:89721448</td>
<td>C/T</td>
<td>intergenic</td>
<td>0.491</td>
<td>0.08 (0.02)</td>
<td>1.52x10^{-4}</td>
<td></td>
<td>0.08 (0.02)</td>
<td>1.98x10^{-4}</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RNF17</td>
<td>rs148232728</td>
<td>13:25433204</td>
<td>A/C</td>
<td>p.Pro1226Thr</td>
<td>0.047</td>
<td>-0.17 (0.05)</td>
<td>1.53x10^{-4}</td>
<td></td>
<td>-0.16 (0.05)</td>
<td>7.73x10^{-4}</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ENPP3</td>
<td>rs6901900</td>
<td>6:132012886</td>
<td>G/T</td>
<td>intronic</td>
<td>0.365</td>
<td>-0.08 (0.02)</td>
<td>1.58x10^{-4}</td>
<td></td>
<td>-0.08 (0.02)</td>
<td>2.14x10^{-4}</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DNAAF2</td>
<td>rs9989177</td>
<td>14:50092471</td>
<td>C/T</td>
<td>p.Asp720Gly</td>
<td>0.129</td>
<td>0.11 (0.03)</td>
<td>1.61x10^{-4}</td>
<td></td>
<td>0.11 (0.03)</td>
<td>1.58x10^{-4}</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AKAP9</td>
<td>rs186148498</td>
<td>7:91731945</td>
<td>A/G</td>
<td>p.Arg3712Gln</td>
<td>0.012</td>
<td>-0.35 (0.09)</td>
<td>1.65x10^{-4}</td>
<td></td>
<td>-0.34 (0.09)</td>
<td>2.65x10^{-4}</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>XPO5</td>
<td>rs186877420</td>
<td>6:43519108</td>
<td>G/A</td>
<td>p.Val552Ala</td>
<td>0.007</td>
<td>-0.45 (0.12)</td>
<td>1.73x10^{-4}</td>
<td></td>
<td>-0.47 (0.12)</td>
<td>1.45x10^{-4}</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LOC338758-LINC00615</td>
<td>rs535428191</td>
<td>12:90184174</td>
<td>A/C</td>
<td>intergenic</td>
<td>0.002</td>
<td>0.77 (0.21)</td>
<td>2.13x10^{-4}</td>
<td></td>
<td>0.73 (0.21)</td>
<td>4.43x10^{-4}</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MIA3</td>
<td>rs77464975</td>
<td>1:222810206</td>
<td>G/C</td>
<td>intronic</td>
<td>0.015</td>
<td>0.30 (0.08)</td>
<td>2.22x10^{-4}</td>
<td></td>
<td>0.28 (0.08)</td>
<td>8.70x10^{-4}</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OVOL1</td>
<td>rs77014695</td>
<td>11:65558702</td>
<td>G/T</td>
<td>intronic</td>
<td>0.041</td>
<td>0.18 (0.05)</td>
<td>2.27x10^{-4}</td>
<td></td>
<td>0.19 (0.05)</td>
<td>3.51x10^{-4}</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gene</td>
<td>rs</td>
<td>Chromosome Position</td>
<td>Allele</td>
<td>Mutation</td>
<td>MAF</td>
<td>SE</td>
<td>p Value</td>
<td>Odds Ratio</td>
<td>95% CI</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------</td>
<td>-------</td>
<td>---------------------</td>
<td>--------</td>
<td>--------------</td>
<td>------</td>
<td>---------</td>
<td>-----------</td>
<td>------------</td>
<td>-------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C6orf70</td>
<td>rs77857406</td>
<td>6:170160783</td>
<td>G/A</td>
<td>p.Lys263Arg</td>
<td>0.010</td>
<td>0.36(0.10)</td>
<td>3.13x10^-4</td>
<td>0.32(0.10)</td>
<td>2.31x10^-3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ARCN1</td>
<td>rs138250193</td>
<td>11:1118461172</td>
<td>A/G</td>
<td>p.Arg312Gln</td>
<td>0.010</td>
<td>0.36(0.10)</td>
<td>3.37x10^-4</td>
<td>0.39(0.10)</td>
<td>1.45x10^-4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DUSP10-HHIPL2</td>
<td>rs17163128</td>
<td>11:118419902</td>
<td>T/C</td>
<td>intergenic</td>
<td>0.241</td>
<td>0.08(0.02)</td>
<td>3.46x10^-4</td>
<td>0.07(0.02)</td>
<td>1.57x10^-3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TNP1-DIRC3</td>
<td>rs12623324</td>
<td>2:218016330</td>
<td>G/A</td>
<td>intergenic</td>
<td>0.221</td>
<td>-0.08(0.02)</td>
<td>4.10x10^-4</td>
<td>-0.08(0.02)</td>
<td>1.66x10^-3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CUEDC1</td>
<td>rs17762338</td>
<td>17:55551037</td>
<td>T/C</td>
<td>p.Arg169His</td>
<td>0.003</td>
<td>0.69(0.20)</td>
<td>4.47x10^-4</td>
<td>0.70(0.20)</td>
<td>4.78x10^-4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OR13G1</td>
<td>rs147056410</td>
<td>1:247835749</td>
<td>T/C</td>
<td>p.Ala199Thr</td>
<td>0.002</td>
<td>-0.82(0.24)</td>
<td>4.83x10^-4</td>
<td>-0.83(0.24)</td>
<td>6.45x10^-4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MICALCL</td>
<td>rs182486240</td>
<td>11:12348754</td>
<td>A/G</td>
<td>p.Arg557Gln</td>
<td>0.005</td>
<td>-0.50(0.14)</td>
<td>4.93x10^-4</td>
<td>-0.53(0.15)</td>
<td>3.82x10^-4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EDARADD</td>
<td>rs60808129</td>
<td>1:236557804</td>
<td>A/G</td>
<td>p.Glu20Glu</td>
<td>0.306</td>
<td>0.07(0.02)</td>
<td>4.97x10^-4</td>
<td>0.08(0.02)</td>
<td>2.86x10^-4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

A1: Minor allele; A2: Major allele; MAF: Minor allele frequency; SE: Standard error. The betas are reported relative to the minor allele. *Model 1: Adjusted for age, sex, PC1 and PC2. †Model 2: Adjusted for age, sex, BMI, PC1 and PC2. Chromosomal positions are presented according to human reference genome hg19.