SUPPLEMENTARY DATA

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 eignenvalues, 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

SUPPLEMENTARY DATA

References

1. Cheung CY, Tang CS, Xu A, Lee CH, Au KW, Xu L, Fong CH, Kwok KH, Chow WS, Woo YC, Yuen MM, Hai JS, Jin YL, Cheung BM, Tan KC, Cherny SS, Zhu F, Zhu T, Thomas GN, Cheng KK, Jiang CQ, Lam TH, Tse HF, Sham PC, Lam KS: Exome-chip association analysis reveals an Asian-specific missense variant in PAX4 associated with type 2 diabetes in Chinese individuals. Diabetologia 2017;60:107-115

2. Tang CS, Zhang H, Cheung CY, Xu M, Ho JC, Zhou W, Cherny SS, Zhang Y, Holmen O, Au KW, Yu H, Xu L, Jia J, Porsch RM, Sun L, Xu W, Zheng H, Wong LY, Mu Y, Dou J, Fong CH, Wang S, Hong X, Dong L, Liao Y, Wang J, Lam LS, Su X, Yan H, Yang ML, Chen J, Siu CW, Xie G, Woo YC, Wu Y, Tan KC, Hveem K, Cheung BM, Zollner S, Xu A, Eugene Chen Y, Jiang CQ, Lam TH, Ganesh SK, Huo Y, Sham PC, Lam KS, Willer CJ, Tse HF, Gao W: Exome-wide association analysis reveals novel coding sequence variants associated with lipid traits in Chinese. Nat Commun 2015;6:10206

3. Janus ED: Epidemiology of cardiovascular risk factors in Hong Kong. Clin Exp Pharmacol Physiol 1997;24:987-988

4. Chen C, Cheung BM, Tso AW, Wang Y, Law LS, Ong KL, Wat NM, Xu A, Lam KS: High plasma level of fibroblast growth factor 21 is an Independent predictor of type 2 diabetes: a 5.4-year population-based prospective study in Chinese subjects. Diabetes Care 2011;34:2113-2115

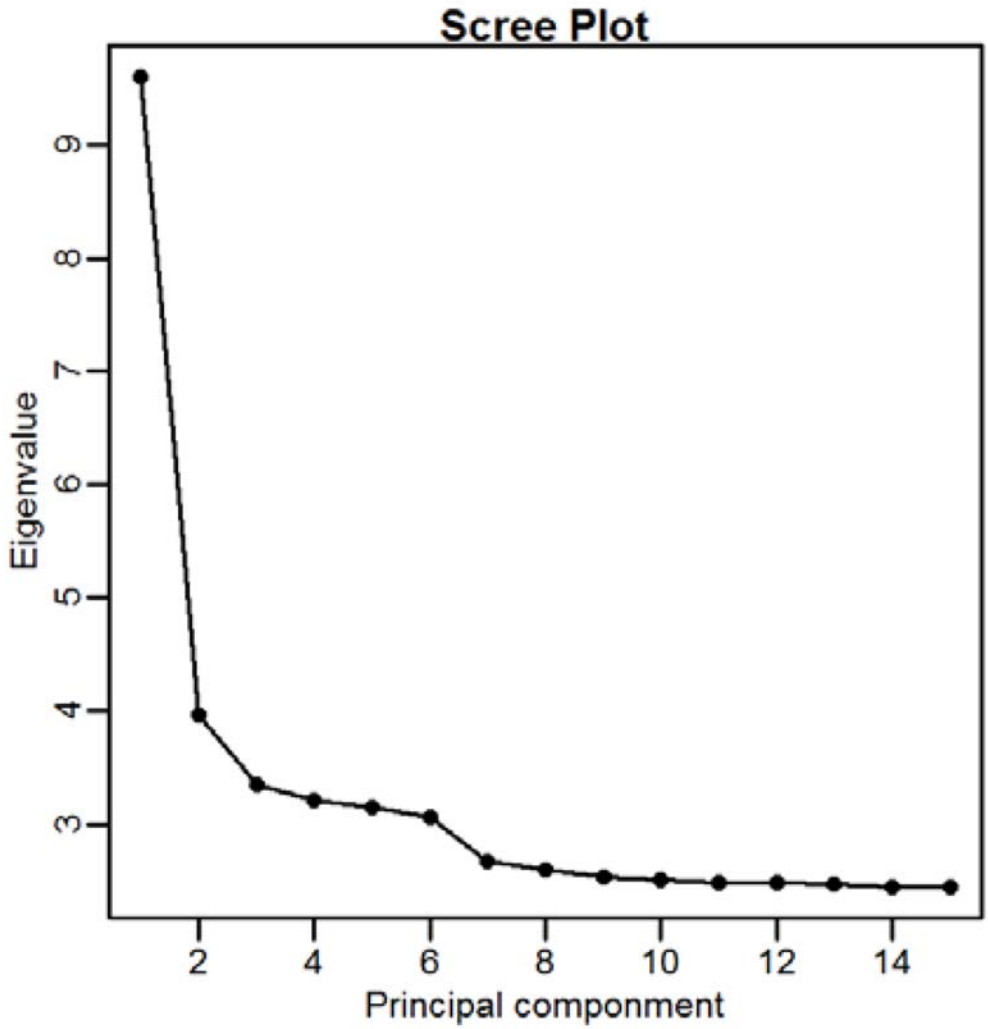
5. Hui E, Yeung CY, Lee PC, Woo YC, Fong CH, Chow WS, Xu A, Lam KS: Elevated circulating pigment epithelium-derived factor predicts the progression of diabetic nephropathy in patients with type 2 diabetes. J Clin Endocrinol Metab 2014;99:E2169-2177

6. Lee CH, Hui EY, Woo YC, Yeung CY, Chow WS, Yuen MM, Fong CH, Xu A, Lam KS: Circulating fibroblast growth factor 21 levels predict progressive kidney disease in subjects with type 2 diabetes and normoalbuminuria. J Clin Endocrinol Metab 2015;100:1368-1375

7. Zhang Y, Long J, Lu W, Shu XO, Cai Q, Zheng Y, Li C, Li B, Gao YT, Zheng W: Rare coding variants and breast cancer risk: evaluation of susceptibility Loci identified in genome-wide association studies. Cancer Epidemiol Biomarkers Prev 2014;23:622-628

SUPPLEMENTARY DATA

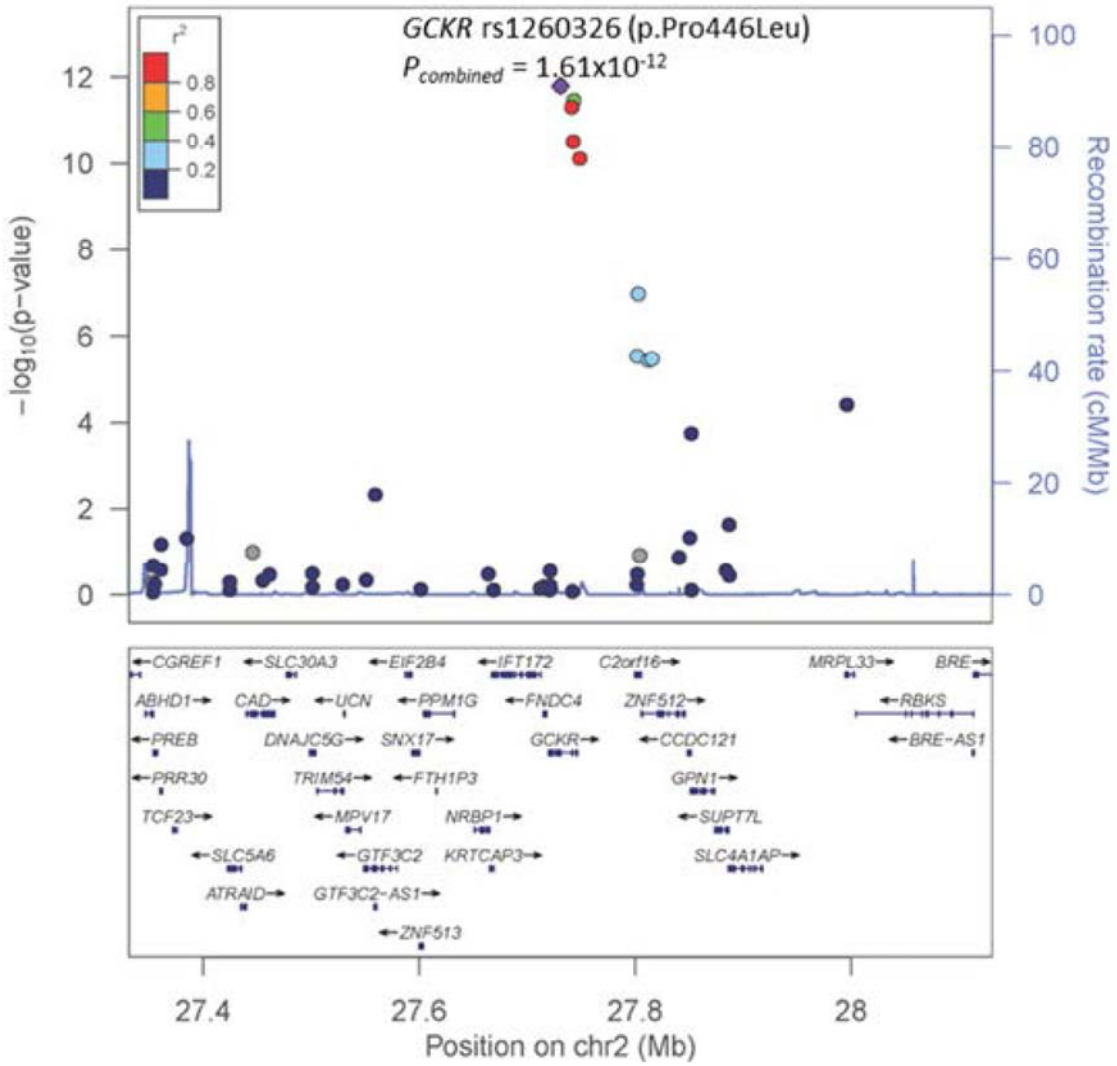
Supplementary Figure 1. Scree plot



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 DATA

Supplementary Figure 2. Regional plot



The regional plot illustrating individual SNPs at the *GCKR* locus and their association to circulating FGF21 levels.

SUPPLEMENTARY DATA

Supplementary Table 1. Index SNPs with *P*combined <$5x10^{-4}$ in the combined analysis for association with FGF21 levels.

						Model 1		Model 2	
Gene	**SNP**	**Position**	**A1/A2**	**Annotation**	**MAF**	**Beta(SE)**	***P-value***	**Beta(SE)**	***P-value***
GCKR	rs1260326	2:27730940	T/C	p.Pro446Leu	0.455	0.14(0.02)	$1.61x10^{-12}$	0.15(0.02)	$3.01x10^{-14}$
FTO	rs17817964	16:53828066	T/C	intronic	0.158	0.13(0.03)	$1.44x10^{-6}$	0.11(0.03)	$7.59x10^{-5}$
TSR1	rs2273983	17:2227988	C/T	p.Asn719Ser	0.009	0.48(0.10)	$4.61x10^{-6}$	0.49(0.11)	$4.90x10^{-6}$
EARS2	rs200139797	16:23563485	C/T	p.Met94Val	0.004	0.64(0.15)	$1.69x10^{-5}$	0.64(0.15)	$1.70x10^{-5}$
GHRHR	rs144372265	7:31009485	T/C	p.Ala91Val	0.004	0.68(0.16)	$2.90x10^{-5}$	0.62(0.17)	$2.41x10^{-4}$
MRPL33	rs3792252	2:27995931	A/G	intronic	0.417	0.08(0.02)	$3.83x10^{-5}$	0.07(0.02)	$7.46x10^{-4}$
C7	rs2271708	5:40936541	C/T	p.Cys128Arg	0.033	0.22(0.05)	$7.94x10^{-5}$	0.23(0.06)	$6.61x10^{-5}$
ENTPD4	rs190514217	8:23291891	T/C	p.Asp521Asn	0.004	-0.58(0.15)	$1.01x10^{-4}$	-0.55(0.15)	$3.64x10^{-4}$
MBNL2	rs7320865	13:97976442	G/A	intronic	0.495	-0.08(0.02)	$1.02x10^{-4}$	-0.07(0.02)	$7.39x10^{-4}$
CDH23	rs10999947	10:73434906	A/G	p.Ser496Asn	0.305	-0.08(0.02)	$1.33x10^{-4}$	-0.08(0.02)	$1.36x10^{-4}$
CETN3-MBLAC2	rs6890115	5:89721448	C/T	intergenic	0.491	0.08(0.02)	$1.52x10^{-4}$	0.08(0.02)	$1.98x10^{-4}$
RNF17	rs148232728	13:25433204	A/C	p.Pro1226Thr	0.047	-0.17(0.05)	$1.53x10^{-4}$	-0.16(0.05)	$7.73x10^{-4}$
ENPP3	rs6901900	6:132012886	G/T	intronic	0.365	-0.08(0.02)	$1.58x10^{-4}$	-0.08(0.02)	$2.14x10^{-4}$
DNAAF2	rs9989177	14:50092471	C/T	p.Asp720Gly	0.129	0.11(0.03)	$1.61x10^{-4}$	0.11(0.03)	$1.58x10^{-4}$
AKAP9	rs186148498	7:91731945	A/G	p.Arg3712Gln	0.012	-0.35(0.09)	$1.65x10^{-4}$	-0.34(0.09)	$2.65x10^{-4}$
XPO5	rs186877420	6:43519108	G/A	p.Val552Ala	0.007	-0.45(0.12)	$1.73x10^{-4}$	-0.47(0.12)	$1.45x10^{-4}$
LOC338758-LINC00615	rs535428191	12:90184174	A/C	intergenic	0.002	0.77(0.21)	$2.13x10^{-4}$	0.73(0.21)	$4.43x10^{-4}$
MIA3	rs77464975	1:222810206	G/C	intronic	0.015	0.30(0.08)	$2.22x10^{-4}$	0.28(0.08)	$8.70x10^{-4}$
OVOL1	rs77014695	11:65558702	G/T	intronic	0.041	0.18(0.05)	$2.27x10^{-4}$	0.19(0.05)	$3.51x10^{-4}$

SUPPLEMENTARY DATA

C6orf70	rs77857406	6:170160783	G/A	p.Lys263Arg	0.010	0.36(0.10)	$3.13x10^{-4}$	0.32(0.10)	$2.31x10^{-3}$
ARCN1	rs138250193	11:118461172	A/G	p.Arg312Gln	0.010	0.36(0.10)	$3.37x10^{-4}$	0.39(0.10)	$1.45x10^{-4}$
DUSP10-HHIPL2	rs17163128	1:222619902	T/C	intergenic	0.241	0.08(0.02)	$3.46x10^{-4}$	0.07(0.02)	$1.57x10^{-3}$
TNP1-DIRC3	rs12623324	2:218016330	G/A	intergenic	0.221	-0.08(0.02)	$4.10x10^{-4}$	-0.08(0.02)	$1.66x10^{-3}$
CUEDC1	rs17762338	17:55951037	T/C	p.Arg169His	0.003	0.69(0.20)	$4.47x10^{-4}$	0.70(0.20)	$4.78x10^{-4}$
OR13G1	rs147056410	1:247835749	T/C	p.Ala199Thr	0.002	-0.82(0.24)	$4.83x10^{-4}$	-0.83(0.24)	$6.45x10^{-4}$
MICALCL	rs182486240	11:12348754	A/G	p.Arg557Gln	0.005	-0.50(0.14)	$4.93x10^{-4}$	-0.53(0.15)	$3.82x10^{-4}$
EDARADD	rs60808129	1:236557804	A/G	p.Glu20Glu	0.306	0.07(0.02)	$4.97x10^{-4}$	0.08(0.02)	$2.86x10^{-4}$

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. 9