Identification of causal variants for T2D and related traits

Min Jin Go

Center for Genome Science

Korea National Institute of Health

Contents

- **Introduction**
- **META results**
  - **OGTT**
  - **Metabolic traits**
  - **AGEN T2D**
  - **AGEN FPG**

Contents

- **Introduction**
- **META results**
  - **OGTT**
  - **Metabolic traits**
  - **AGEN T2D**
  - **AGEN FPG**

How to interpret a GWAS

- Use high-throughput genotyping technologies
- Assay hundreds of thousands of SNPs
- Related them to clinical conditions and measurable traits

(Pearson & Manolio, JAMA. 2008)

How to conduct GWAS and META

코호트 연구:

질병이 발생되기 이전에 질병을 일으킬 것으로 예상되는 요인들을 일정 집단에서 계속 추적해 가면서 각 요소가 질병의 발생에 관여하는 정도를 관찰

1

개체선별:

유전체 연구를 위한 개체 선별

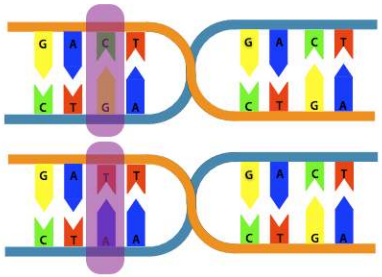
혈액샘플

단일염기다형성(SNP):

하나의 염기서열이 다른 유전적 변이의 하나

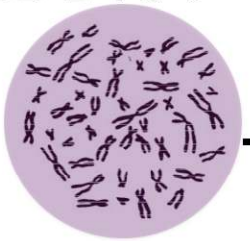
GACTGACT

GATTGACT

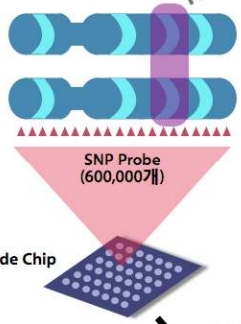


2

인간 염색체



3



SNP Genotyping:

Genome-Wide Chip을 이용하여 모든 염색체와 DNA에 대한 SNP을 알아낸다.

4

슈퍼컴퓨터를 이용한 통계 분석:

안성/안산의 코호트의 경우 10,000 명에 대한 600,000(총 6십억개의 SNP)에 대해서 Bioinformatics와 통계적인 방법을 활용하여 질병과의 연관성 분석

5

혈압 관련 유전형 발굴:

(신체 표현형을 관장하는 유전형 발굴)

인간 12번 염색체의 q21.33부분의 SNP(rs17249754)이 혈압과 연관성을 보임

GG인 경우 평균 116.62mmHg
GA인 경우 평균 115.42mmHg
AA인 경우 평균 114.10mmHg 를 보임

6

AA 유전형의 경우 GG 보다 혈압이 낮아 고혈압과 같은 질환에 대한 감수성이 낮다

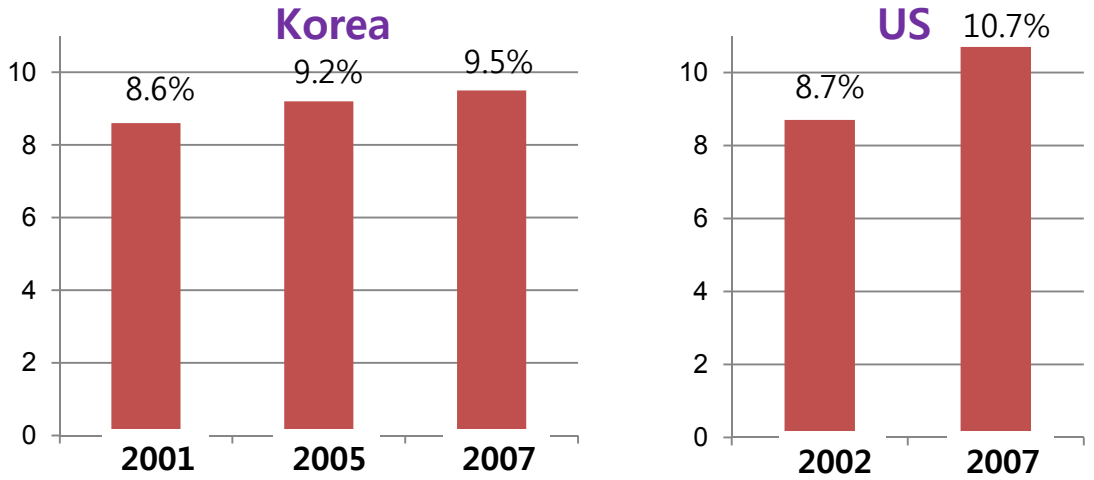
맞춤의료와 예방의학:
(질병으로부터 자유로운 세상)

대량의 인간 유전체 연구를 통하여 질환관련 인자를 확인한 경우 우리는 무엇을 할 수 있을까?

감시: 가능한 빨리 질병 발병 여부를 알아냄
위험회피: 운동과 알콜 섭취 회피
예방적 화학요법: 개인별 약물 요법

Why a genetic study of T2D ?

- **T2D huge, growing public health problem worldwide**
  - risk factors : renal failure, retinopathy, peripheral neuropathy, cardiovascular diseases and so on
  - death rate: 22.9/100,000 in Korea in 2007 (5th ranked)
  - prevalence rate growing rapidly, as below;



Age above 30 years old & FPG>126mg/dL or diabetes medication treatment

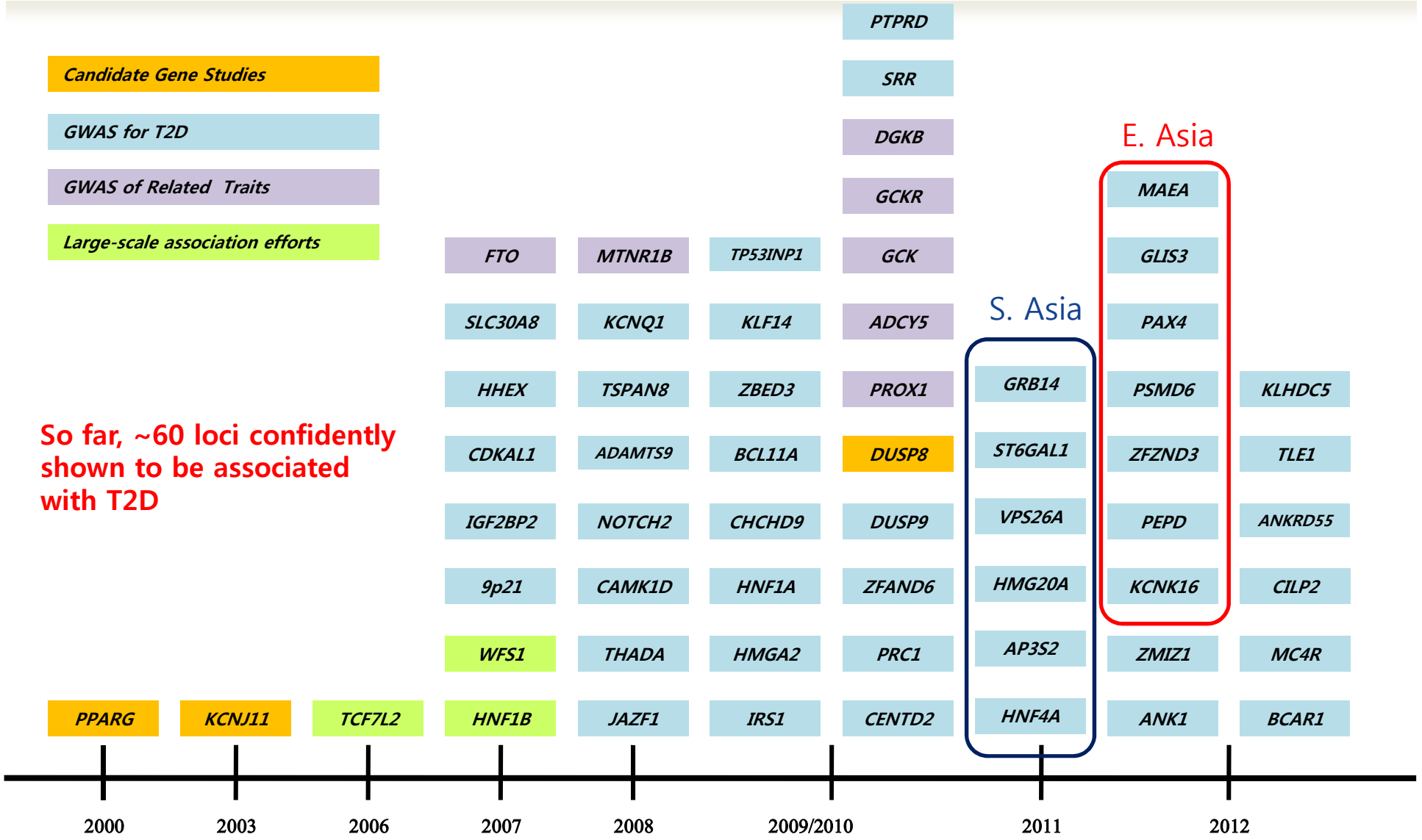
(2008 Korea National Health and Nutrition Examination Survey)

- **T2D is a chronic metabolic disease with multi-factorial pathogenesis**

- **Although the genetic contribution to T2D is well recognized , the current set of 60 established susceptibility loci, identified primarily through large-scale genome-wide association studies (GWAS), captures at best 10% of familial aggregation of the disease**

- **The characteristics (effect sizes and risk allele frequencies (RAFs)) of the variants contributing to the unexplained genetic variance remain far from clear**

GWAS and META of T2D, ~2012



Contents

- Introduction
- **META results**
  - **OGTT, *in revision***
  - Metabolic traits
  - AGEN T2D
  - AGEN FPG

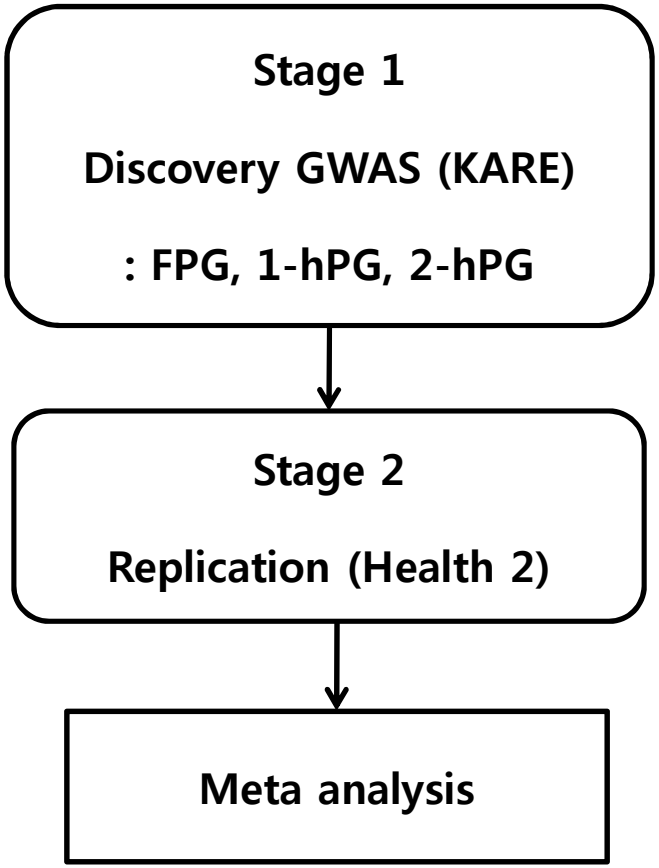
New susceptibility loci associated with one-hour plasma glucose as predisposing risk factors for type 2 diabetes

Min Jin Go[1], Joo-Yeon Hwang[1], Young Jin Kim[1], Ji Hee Oh[1], Yeon-Jung Kim[1], Soo Heon Kwak[2], Kyung Soo Park[2,3], Juyoung Lee[1], Bong-Jo Kim[1], Bok-Ghee Han[1], Myeong-Chan Cho[4], Yoon Shin Cho[1,5,*] and Jong-Young Lee[1,*]

Introduction

- Impaired glucose tolerance (IGT) and impaired fasting glucose (IFG) as an intermediate risk factor for T2D
- After adjustment for 1-hPG, the high fasting plasma glucose (FPG) level is not associated with an increase in the incidence of T2D
- 1-hPG levels are associated with elevated liver enzymes, kidney dysfunction, coronary heart disease and left ventricular hypertrophy
- Most previous studies have a limited impact on FPG and 2-hPG values for glycemic traits from Eurocentric consortium
- Identifying new genetic susceptibility predisposing to early plasma glucose condition would be more valuable as an important parameter for risk stratification in glucose abnormalities

Study scheme



Descriptive statistics

Trait	KARE (n = 7,696)	Health2 (n = 6,536)
AGE	51.64 ± 8.80	56.14 ± 7.94
SEX(M/F)	3584/4112	2584/3952
BMI	24.48 ± 3.08	24.35 ± 3.14
FPG (mmol/l)	4.62 ± 0.49	5.03 ± 0.50
1-hPG (mmol/l)	7.94 ± 2.26	8.78 ± 2.48
2-hPG (mmol/l)	6.42 ± 1.68	6.66 ± 1.79

Association results for glycemic traits

Trait	CHR	SNP	Gene	A1	KARE			Health2			Meta analysis			
					N	Beta ± sem	*P*	N	Beta ± sem	*P*	N	Beta ± sem	*P*	Hetero *P*
FPG	7	rs1799884	*GCK*	A	7685	0.065 ± 0.010	4.06E-11	6526	0.061 ± 0.011	2.14E-08	14211	0.063 ± 0.007	4.53E-18	0.8161
	11	rs10830962	*MTNR1B*	C	7552	0.036 ± 0.008	3.49E-06	6529	0.049 ± 0.009	1.60E-08	14081	0.041 ± 0.006	4.84E-13	0.2647
	12	rs2074356	*C12orf51*	T	7695	-0.054 ± 0.011	5.42E-07	6498	-0.070 ± 0.012	1.43E-08	14193	-0.061 ± 0.008	6.03E-14	0.3249
1-hPG	6	rs9348440	*CDKAL1*	A	7651	0.214 ± 0.036	3.72E-09	6522	0.290 ± 0.042	5.99E-12	14173	0.246 ± 0.028	3.13E-19	0.1707
	7	rs1799884	*GCK*	A	7641	0.232 ± 0.046	4.47E-07	6520	0.175 ± 0.054	0.0013	14161	0.208 ± 0.035	2.82E-09	0.4213
	11	rs10830962	*MTNR1B*	C	7508	0.179 ± 0.036	6.16E-07	6523	0.207 ± 0.042	1.01E-06	14031	0.191 ± 0.027	3.24E-12	0.6192
	12	**rs12229654**	***MYL2***	**G**	**7650**	**-0.262 ± 0.051**	**2.60E-07**	**6525**	**-0.299 ± 0.060**	**6.70E-07**	**14175**	**-0.277 ± 0.039**	**8.83E-13**	**0.6407**
	12	**rs2074356**	***C12orf51***	**T**	**7651**	**-0.296 ± 0.050**	**3.65E-09**	**6492**	**-0.357 ± 0.061**	**4.10E-09**	**14143**	**-0.321 ± 0.039**	**1.04E-16**	**0.4335**
	12	**rs11066453**	***OAS1***	**G**	**7595**	**-0.290 ± 0.054**	**8.52E-08**	**6523**	**-0.175 ± 0.064**	**0.0062**	**14121**	**-0.242 ± 0.041**	**4.54E-09**	**0.1717**
2-hPG	7	rs1799884	*GCK*	A	7647	0.186 ± 0.034	3.97E-08	6519	0.129 ± 0.039	0.0009	14166	0.162 ± 0.026	2.59E-10	0.2753
	12	rs2074356	*C12orf51*	T	7657	-0.182 ± 0.037	8.52E-07	6491	-0.140 ± 0.044	0.0015	14148	-0.165 ± 0.028	5.91E-09	0.4661

Summary

• Most recently, one-hour hyperglycemia has been recognized as a further risk factor for T2D

• To date, previous GWAS for glycemic traits have a limited impact on the fasting state and 2-h plasma glucose level in an oral glucose challenge

• To identify genetic susceptibility in different stages of glucose tolerance, we performed a meta-analysis for glycemic traits including 1-hPG from 14,232 non-diabetic individuals in the Korean population

• Newly implicated variants (*MYL2*, *C12orf51* and *OAS1*) were found to be significantly associated with 1-hPG. We also demonstrated associations with GDM

• Our results could provide additional insight into the genetic variation in the clinical range of glycemia

Contents

- Introduction
- **META results**
  - OGTT
  - **Metabolic traits, *published in Nat Gent***
  - AGEN T2D
  - AGEN FPG

LETTERS

nature genetics

Large-scale genome-wide association studies in east Asians identify new genetic loci influencing metabolic traits

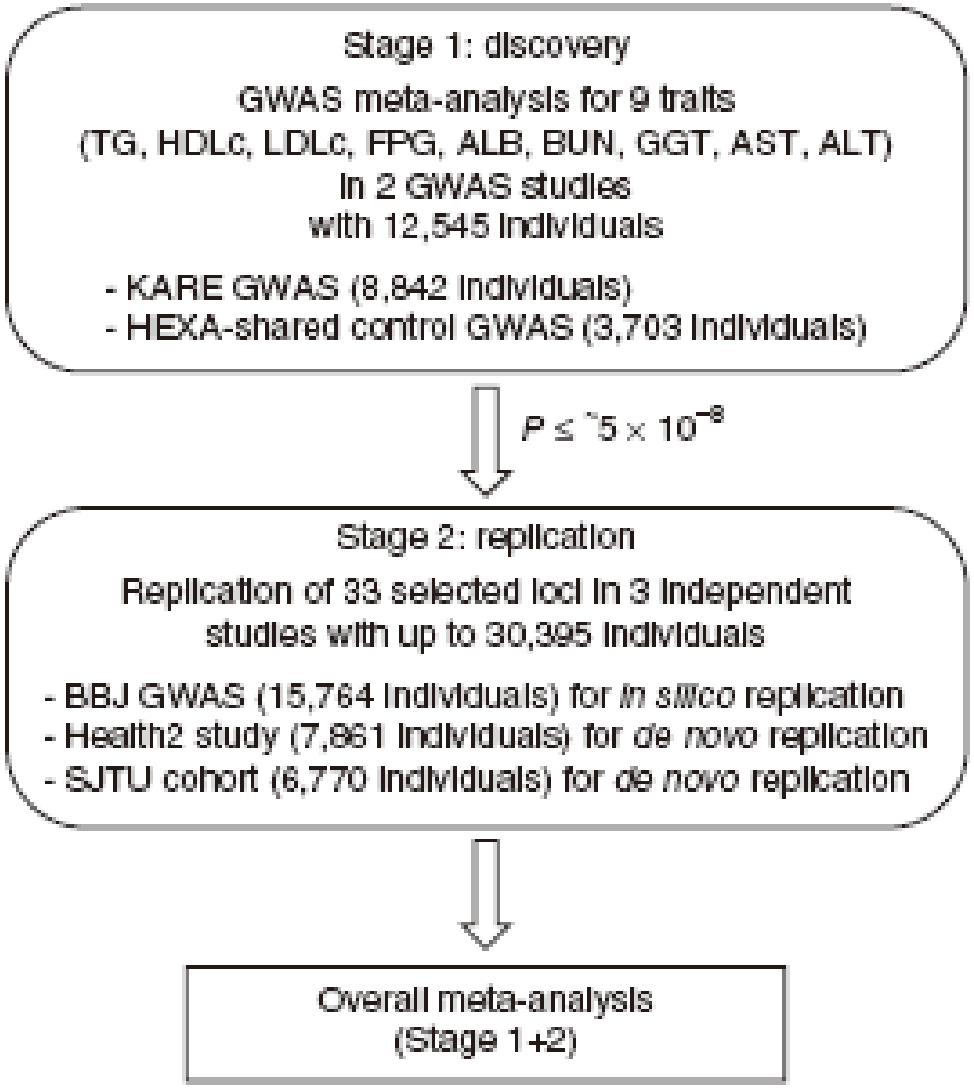
Young Jin Kim[1,14], Min Jin Go[1,14], Cheng Hu[2], Chang Bum Hong[1], Yun Kyoung Kim[1], Ji Young Lee[1], Joo-Yeon Hwang[1], Ji Hee Oh[1], Dong-Joon Kim[1], Nam Hee Kim[1], Soeui Kim[1], Eun Jung Hong[1], Ji-Hyun Kim[1], Haesook Min[1], Yeonjung Kim[1], Rong Zhang[2], Weiping Jia[2], Yukinori Okada[3,4], Atsushi Takahashi[3], Michiaki Kubo[3], Toshihiro Tanaka[3], Naoyuki Kamatani[3], Koichi Matsuda[5], MAGIC consortium[6], Taesung Park[7], Bermseok Oh[8], Kuchan Kimm[9], Daehee Kang[10], Chol Shin[11], Nam H Cho[12], Hyung-Lae Kim[1,13], Bok-Ghee Han[1], Jong-Young Lee[1] & Yoon Shin Cho[1]

To identify the genetic bases for nine metabolic traits, we conducted a meta-analysis combining Korean genome-wide

traits, we conducted a two-stage association study in individuals of east Asian ancestry (**Fig. 1** and **Supplementary Table 1**).

Kim et al. *Nature Genetics, 2011*

Study scheme

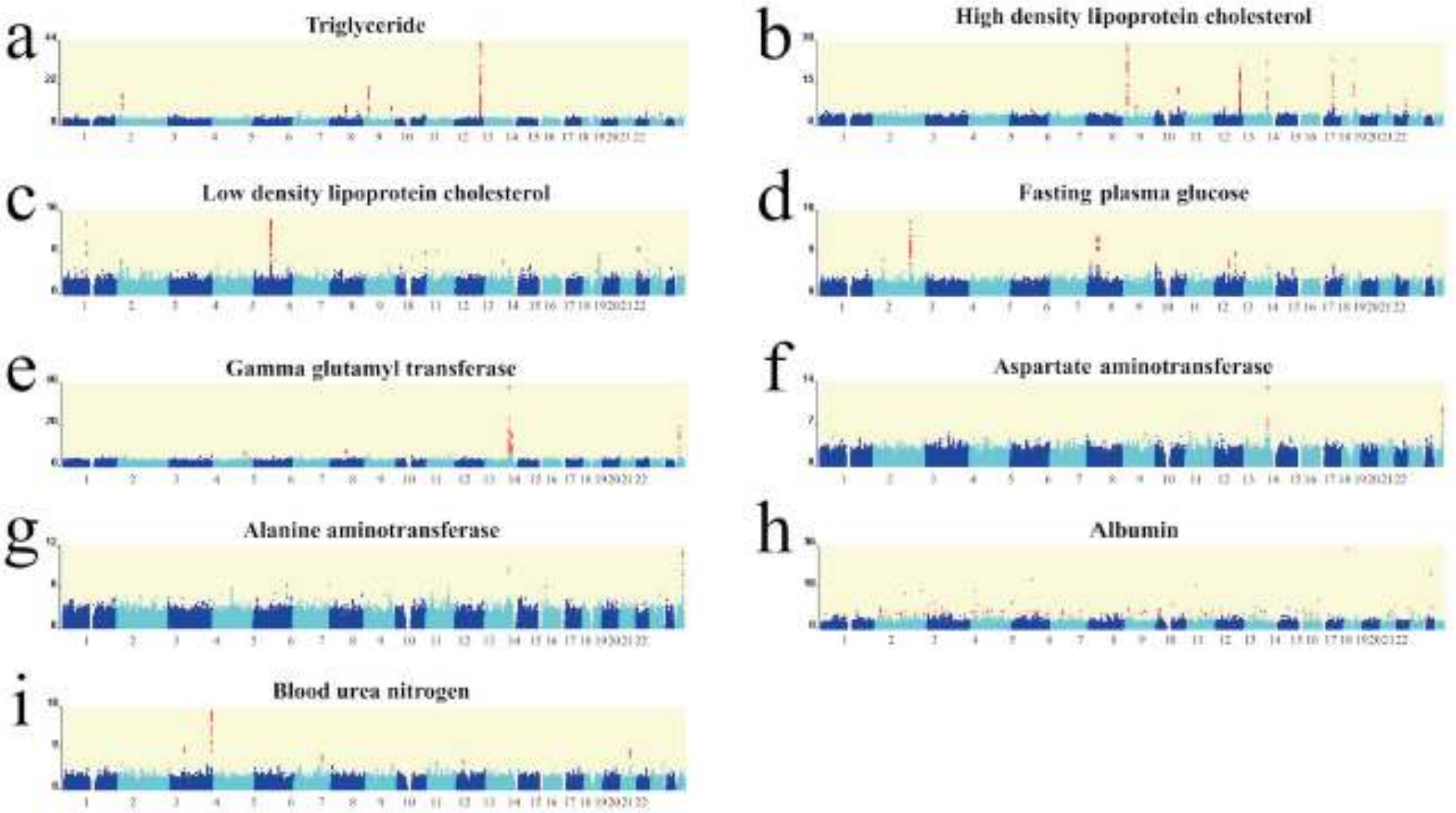


Descriptive statistics of variables

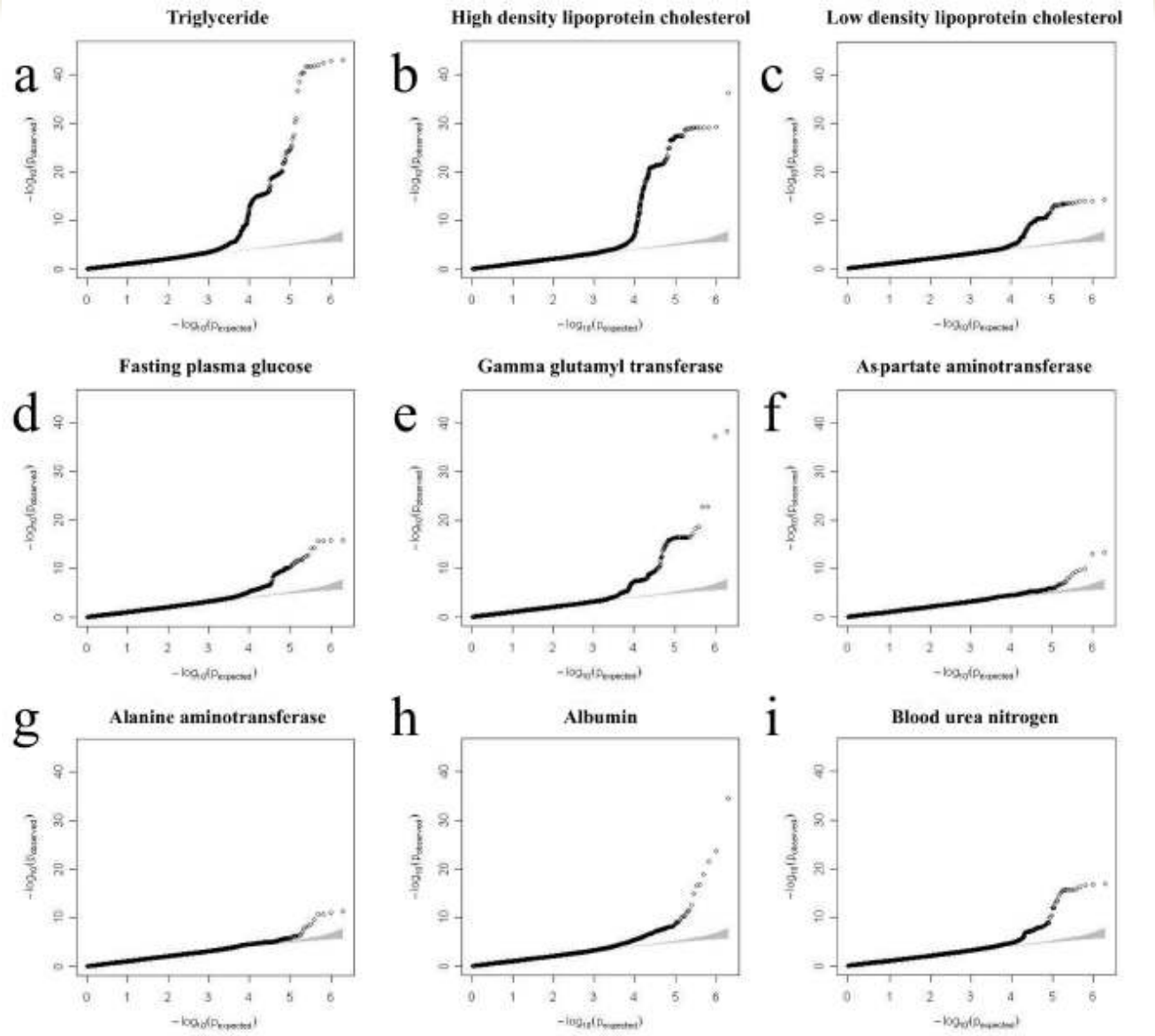
Study	KARE	HEXA-shared control	BBJ	SJTU	Health2
Stage	Discovery	Discovery	*in silico* replication	*De novo* replication	*De novo* replication
Study design	Population-based Prospective	Population-based Multicenter	Case-control Multicenter	Population-based	Population-based
Sample size	8842	3703	15764	6770	7861
Age (yr)	52.22 ± 8.92	53.21 ± 8.33	63.70 ± 10.50	55.92 ± 14.25	56.58 ± 7.85
Male/Female	4183/4689	1651/2052	10400/5364	3087/3683	3124/4647
TG (mg/dl)	162.44 ± 104.53	123.29 ± 91.40	151.20 ± 108.60	148.87 ± 188.05	157.26 ± 108.94
HDLc (mg/dl)	44.66 ± 10.10	54.61 ± 13.27	51.80 ± 14.60	47.56 ± 18.19	46.34 ± 11.24
LDLc (mg/dl)$^{\$}$	115.65 ± 32.14	119.04 ± 31.39	126.00 ± 34.00	110.71 ± 48.45	127.70 ± 34.57
FPG (mg/dl)	4.62 ± 0.49	5.00 ± 0.56	NA	NA	5.03 ± 0.50
ALB (g/dl)	4.24 ± 0.33	4.65 ± 0.29	4.26 ± 0.35	NA	4.62 ± 0.27
BUN (mg/dl)	14.37 ± 3.80	14.02 ± 3.83	15.30 ± 3.90	15.57 ± 6.84	15.77 ± 4.51
GGT (IU/L)	36.08 ± 66.17	32.03 ± 40.74	39.40 ± 26.80	29.28 ± 30.77	39.81 ± 86.43
ALT (IU/L)	28.39 ± 25.60	24.44 ± 20.04	23.70 ± 12.60	24.67 ± 23.88	26.36 ± 20.22
AST (IU/L)	29.91 ± 18.62	24.51 ± 12.94	24.10 ± 9.20	22.57 ± 17.97	29.47 ± 26.79

Nine metabolic traits are abbreviated as follows: high density lipoprotein cholesterol (HDLc), low density lipoprotein cholesterol (LDLc), triglyceride (TG), fasting plasma glucose (FPG), albumin (ALB), blood urea nitrogen (BUN), gamma glutamyl transferase (GGT), alanine aminotransferase (ALT), aspartate aminotransferase (AST). $Low density lipoprotein cholesterol is calculated by using the Friedewald's formula. KARE, Korea Association Resource study; HEXA, Health Examinee cohort; BBJ, BioBank Japan study; SJTU, Shanghai Jiao Tong University cohort; NA, data not available.

Result of Stage 1



Result of Stage 1

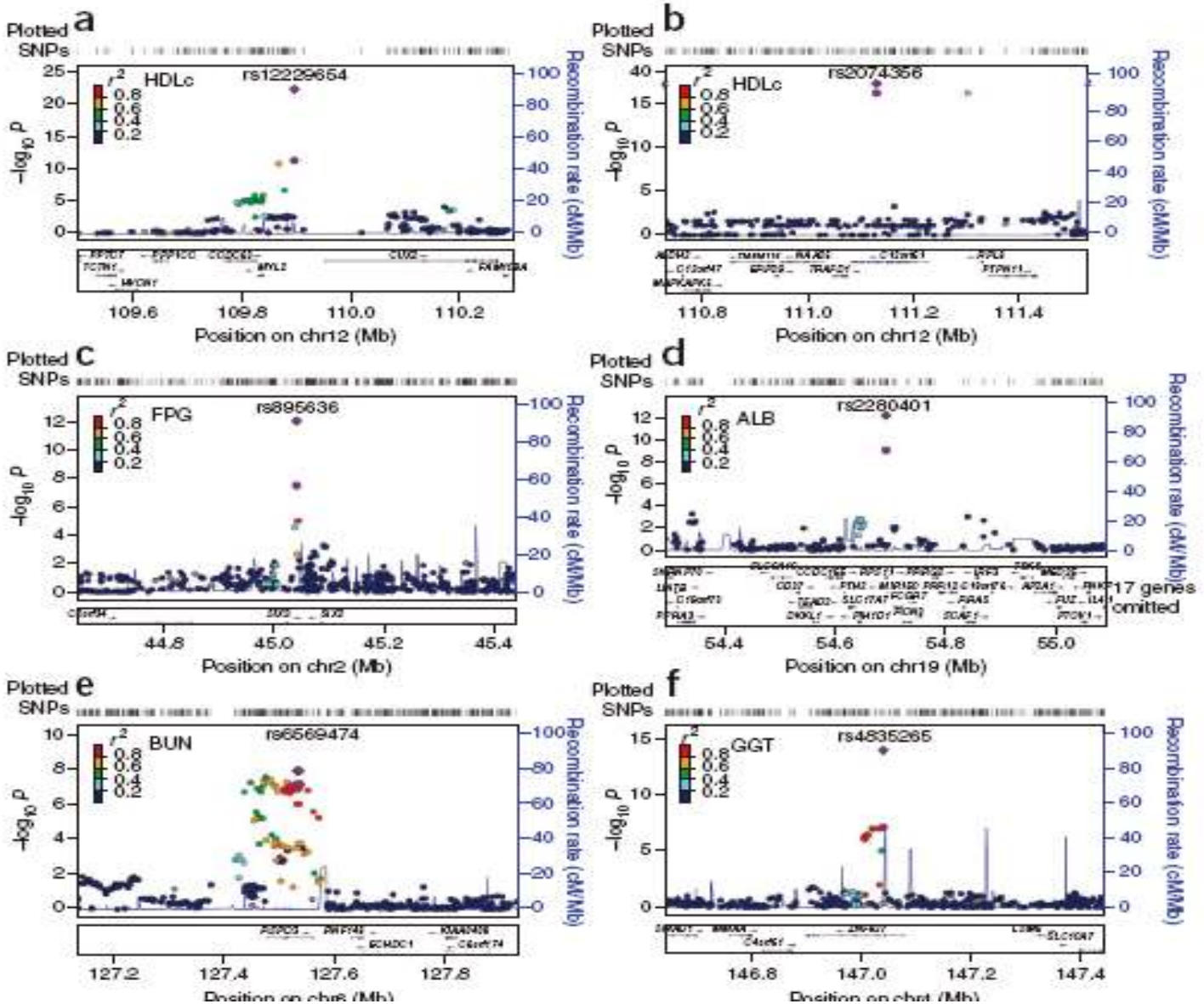


Association results for metabolic traits

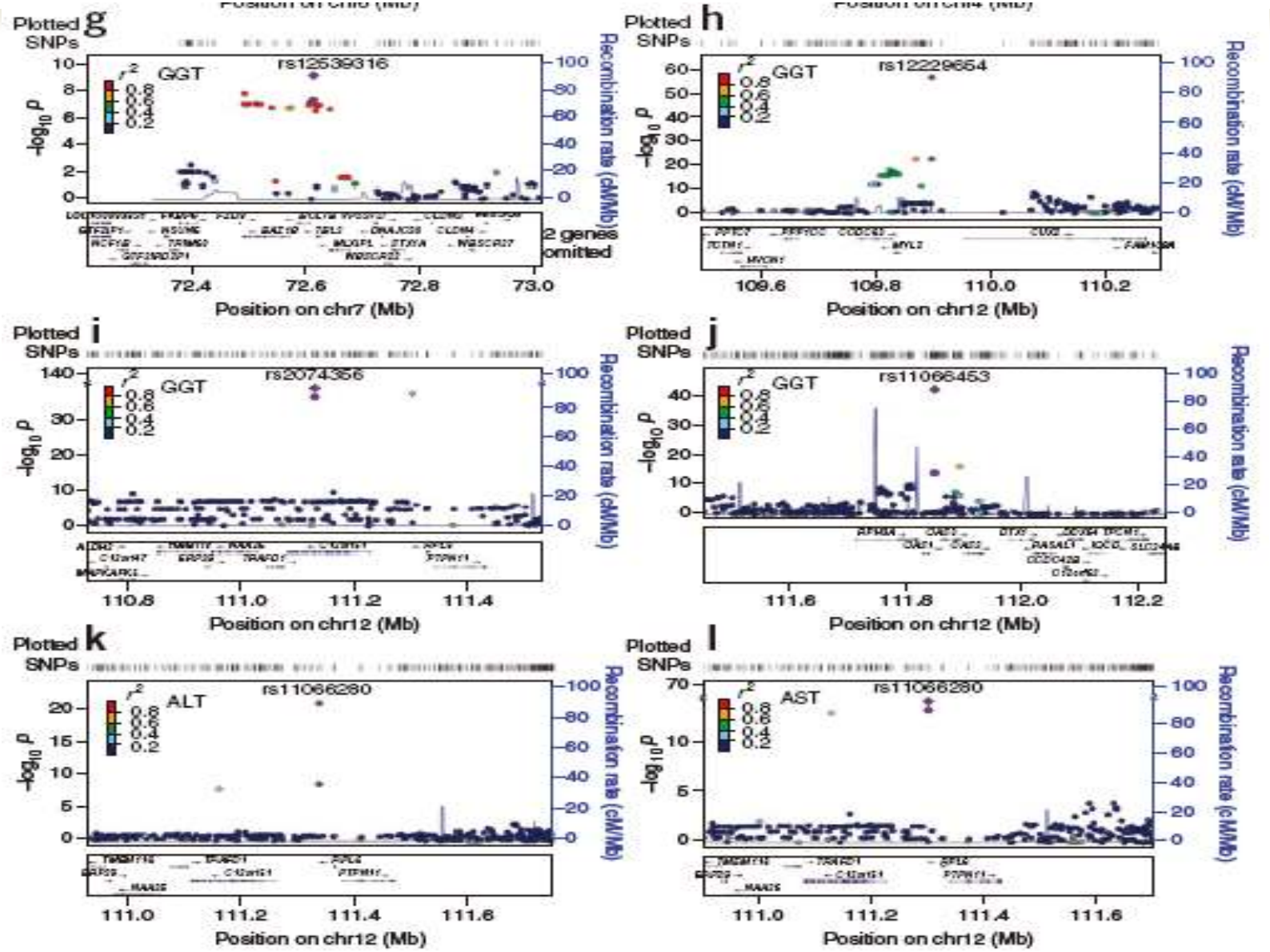
Trait	Locus	Position	Lead SNP[a]	Candidate gene	Allele[b]	MAF	Discovery meta-analysis (KARE+HEXA shared control)		
							Effect size[c]	*P*	Sample size
Plasma lipid traits									
HDLc	12q24.11	110,000,000	rs12229654	*MYL2*	G/T	0.14	−0.0284 ± 0.0041	3.20×10^{-12}	12,394
	12q24.13	110,000,000	rs2074356	*C12orf51*	T/C	0.15	−0.0338 ± 0.0040	2.31×10^{-17}	12,502
Glycemic trait									
FPG	2p21	45,041,857	rs895636	*SIX2-SIX3*	T/C	0.38	0.0392 ± 0.0070	1.87×10^{-8}	11,043
Renal-function–related traits									
ALB	19q13.33	54,691,821	rs2280401	*RPS11*	A/G	0.17	0.0312 ± 0.0051	6.43×10^{-10}	12,541
BUN	6q22.33	128,000,000	rs6569474[d]	*RSPO3*	A/T	0.47	−0.0172 ± 0.0032	4.80×10^{-8}	12,510
Liver enzymes									
GGT	4q31.22	147,000,000	rs4835265	*ZNF827*	A/C	0.42	−0.0047 ± 0.0009	6.21×10^{-8}	10,503
	7q11.23	72,615,834	rs12539316[e]	*TBL-BCL7B*	G/A	0.10	0.0081 ± 0.0014	2.41×10^{-8}	10,492
	12q24.11	110,000,000	rs12229654	*MYL2*	G/T	0.14	0.0132 ± 0.0012	2.23×10^{-26}	10,450
	12q24.13	111,000,000	rs2074356	*C12orf51*	T/C	0.15	0.0165 ± 0.0012	2.05×10^{-41}	10,505
	12q24.13	112,000,000	rs11066453	*OAS1*	G/A	0.13	0.0105 ± 0.0013	1.55×10^{-15}	10,434
ALT	12q24.13	111,000,000	rs11066280	*C12orf51*	A/T	0.17	0.0044 ± 0.0008	6.19×10^{-9}	12,241
AST	12q24.13	111,000,000	rs11066280	*C12orf51*	A/T	0.17	0.0013 ± 0.0002	9.74×10^{-14}	12,241

		Replication meta-analysis (BBJ+Health2+SJTU)			Overall meta-analysis (Discovery+Replication)			
Trait	Locus	Effect size[c]	*P*	Sample size	Effect size[c]	*P*	Sample size	P_{het}
Plasma lipid traits								
HDLc	12q24.11	−0.0277 ± 0.0039	1.17×10^{-12}	13,784	−0.0280 ± 0.0028	3.41×10^{-23}	26,178	9.02×10^{-1}
	12q24.13	−0.0360 ± 0.0038	3.59×10^{-21}	13,784	−0.0350 ± 0.0028	6.95×10^{-37}	26,286	6.90×10^{-1}
Glycemic trait								
FPG	2p21	0.0395 ± 0.0090	1.13×10^{-5}	6,574	0.0393 ± 0.0055	9.99×10^{-13}	17,617	6.69×10^{-1}
Renal-function–rela								
ALB	19q13.33	0.0267 ± 0.0067	6.82×10^{-5}	9,469	0.0293 ± 0.0041	8.73×10^{-13}	22,010	5.95×10^{-1}
BUN	6q22.33	−0.0044 ± 0.0017	8.93×10^{-3}	22,076	−0.0090 ± 0.0016	1.26×10^{-8}	34,586	4.12×10^{-4}
Liver enzymes								
GGT	4q31.22	−0.004 ± 0.0007	1.03×10^{-9}	13,874	−0.0043 ± 0.0006	1.01×10^{-14}	24,377	5.39×10^{-1}
	7q11.23	0.0029 ± 0.0010	5.44×10^{-3}	13,902	0.0051 ± 0.0008	5.81×10^{-10}	24,394	2.51×10^{-3}
	12q24.11	0.0109 ± 0.0009	3.29×10^{-34}	13,840	0.0119 ± 0.0007	8.76×10^{-58}	24,290	1.29×10^{-1}
	12q24.13	0.0158 ± 0.0008	6.25×10^{-85}	17,807	0.0161 ± 0.0007	2.88E-126	28,312	6.27×10^{-1}
	12q24.13	0.0092 ± 0.0008	2.41×10^{-28}	17,852	0.0097 ± 0.0007	6.27×10^{-44}	28,286	3.94×10^{-1}
ALT	12q24.13	0.0046 ± 0.0006	2.54×10^{-14}	21,456	0.0045 ± 0.0005	7.62×10^{-22}	33,697	8.36×10^{-1}
AST	12q24.13	0.0018 ± 0.0001	1.85×10^{-38}	21,646	0.0016 ± 0.0001	2.77×10^{-63}	33,887	2.53×10^{-2}

Summary of GWAS for metabolic traits



Summary of GWAS for metabolic traits



Multiple diverse genetic effect

Table 2 Pleiotropic loci detected from GWAS for various traits

	This study								
Locus	SNP ID	Position (bp)	Candidate gene	Affected Trait	Variable transformation	Effect allele (frequency)	Effect size[a]	P^b	Affected trait(s) identified from previous studies
2p23.3	rs780092	27,596,658	*GCKR*	ALB	No	C (0.33)	−0.0235 ± 0.0040	$4.75 \times 10^{-9,b}$	HTG[19], TGC[20], eGFRcrea[21], SUA[22], CRP[23], FI[24], HOMA-IR[24], FPG[24], SU[25]
				TG	Ln	C (0.33)	−0.0500 ± 0.0046	4.58×10^{-27}	
7q11.23	rs12539316	72,615,834	*TBL2-BCL7B*	GGT	1/sqrt	G (0.10)	0.0051 ± 0.0008	5.81×10^{-10}	TG[13], SLE[26], eGFRcrea[21]
	rs2286276	72,625,290		TG	Ln	A (0.10)	−0.0652 ± 0.0082	1.44×10^{-15}	
8p21.3	rs10503669	19,891,970	*LPL*	TG	Ln	A (0.12)	−0.0857 ± 0.0065	6.84×10^{-39}	HTG[19], MCV[1]
				HDLc	Ln	A (0.12)	0.0426 ± 0.0031	8.04×10^{-43}	
9q34.2	rs651007	135,143,696	*ABO*	LDLc	No	A (0.26)	2.2026 ± 0.3841	9.78×10^{-9}	MCHC[1], PC[27], VTE[28], E-selectin[29], P-selectin[30], ICAM-1[30], RBC[1], HB[1], HT[1], ALP[1], ACE[31]
11q23.3	rs11216126	116,122,450	*ZNF259-APOA1/C3/A4/A5-BUD13*	HDLc	Ln	C (0.20)	0.0322 ± 0.0026	2.6×10^{-34}	ATC[32], HDLc[13], HTG[19], Glioma[33], SLE[26]
	rs603446	116,159,645		TG	Ln	T (0.23)	−0.0875 ± 0.0051	2.03×10^{-65}	
12q24.11	rs12229654	109,898,844	*MYL2*	HDLc	Ln	G (0.14)	−0.0280 ± 0.0028	3.41×10^{-23}	HDLc[13]
				GGT	1/sqrt	G (0.14)	0.0119 ± 0.0007	8.76×10^{-58}	
12q24.13	rs2074356	111,129,784	*C12orf51*	HDLc	Ln	T (0.15)	−0.0350 ± 0.0028	6.95×10^{-37}	T1D[34], WHR[4]
				GGT	1/sqrt	T (0.15)	0.0161 ± 0.0007	2.88×10^{-126}	
	rs11066280	111,302,166	*C12orf51*	ALT	1/sqrt	A (0.17)	0.0045 ± 0.0005	7.62×10^{-22}	
				AST	Reciprocal	A (0.17)	0.0016 ± 0.0001	2.77×10^{-63}	

Previous studies were retrieved based on $P < 5 \times 10^{-7}$. Information for SNP ID and chromosomal position is based on NCBI genome build 36 and dbSNP build 129. ALB, albumin; TG, triglyceride; GGT, gamma glutamyl transferase; HDLc, high density lipoprotein cholesterol; LDLc, low density lipoprotein cholesterol; AST, aspartate aminotransferase; HTG, hypertriglyceridemia; TGC, two hour glucose; eGFRcrea, estimated glomerular filtration rate by serum creatinine; SUA, serum uric acid; CRP, c-reactive protein; FI, fasting insulin; HOMA-IR, insulin resistance; FPG, fasting plasma glucose; SU, serum urate; SLE, serum lupus erythematosus; MCV, mean corpuscular volume; MCHC, mean corpuscular hemoglobin concentration; PC, pancreatic cancer; VTE, venous thromboembolism; ICAM-1, intercellular adhesion molecule-1; RBC, erythrocyte count; HB, hemoglobin concentration; HT, hematocrit; ALP, alkaline phosphatase; ACE, angiotensin-converting enzyme activity; ATC, alpha-tocopherol; T1D, type 1 diabetes; WHR, waist hip ratio; Ln, natural log; 1/sqrt, inverse square root.

[a]All *P* values are the overall meta-analysis (discovery+replication) *P* value except the *P* value given for [b]ALB, which is the discovery *P* value. [a]Effect sizes are listed as β ± s.e.m.

Summary

To identify the genetic bases for 9 metabolic traits, we conducted a meta-analysis combining Korean genome-wide association results from the KARE project (n = 8,842) and the HEXA shared control study (n = 3,703)
Replication study; *in silico* from the BioBank Japan GWAS (n = 15,764) and *de novo* from the Health2 (n = 7,861) and Shanghai Jiao Tong University Diabetes cohorts (n = 6,770)

We identified ten genome-wide significant signals newly associated with traits from an overall meta- analysis

Novel associations involved;
12q24.11 (near *MYL2*) and 12q24.13 (in *C12orf51*) for HDLc,
2p21 (near *SIX2*-*SIX3*) for FPG,
19q13.33 (in *RPS11*) and 6q22.33 (in *RSPO3*) for renal traits,
12q24.11 (near *MYL2*), 12q24.13 (in *C12orf51,* near *OAS1*), 4q31.22 (in *ZNF827*) and 7q11.23 (near *TBL2*-*BCL7B*) for hepatic traits

These findings highlight previously unknown biological pathways for metabolic traits investigated in this study.

Contents

- Introduction
- **META results**
  - OGTT
  - Metabolic traits
  - **AGEN T2D, *published in Nat Gent***
  - AGEN FPG

nature genetics

Meta-analysis of genome-wide association studies identifies 8 new loci for type 2 diabetes in East Asians

Yoon Shin Cho[1,§,¶], Chien-Hsiun Chen[2,3,§], Cheng Hu[4,§], Jirong Long[5,§], Rick Twee Hee Ong[6,§], Xueling Sim[7,§], Fumihiko Takeuchi[8,§], Ying Wu[9,§], Min Jin Go[1,§], Toshimasa Yamauchi[10,§], Yi-Cheng Chang[11,§], Soo Heon Kwak[12,§], Ronald C.W. Ma[13,§], Ken Yamamoto[14,§], Linda S. Adair[15], Tin Aung[16,17], Qiuyin Cai[5], Li-Ching Chang[2], Yuan-Tsong Chen[2], Yutang Gao[18], Frank B. Hu[19], Hyung-Lae Kim[1,20], Sangsoo Kim[21], Young Jin Kim[1], Jeannette Jen-Mai Lee[22], Nanette R. Lee[23], Yun Li[9,43], Jian Jun Liu[24], Wei Lu[25], Jiro Nakamura[26], Eitaro Nakashima[26,27], Daniel Peng-Keat Ng[22], Wan Ting Tay[16], Fuu-Jen Tsai[3], Tien Yin Wong[16,17,28], Mitsuhiro Yokota[29], Wei Zheng[5], Rong Zhang[4], Congrong Wang[4], Wing Yee So[13], Keizo Ohnaka[30], Hiroshi Ikegami[31], Kazuo Hara[10], Young Min Cho[12], Nam H Cho[32], Tien-Jyun Chang[11], Yuqian Bao[4], Åsa K Hedman[44], Andrew P Morris[44], Mark I McCarthy[44,45], DIAGRAM consortium[46], MuTHER consortium[46], Ryoichi Takayanagi[30,*], Kyong Soo Park[12,33,*], Weiping Jia[4,*], Lee-Ming Chuang[11,34,*], Juliana C.N. Chan[13,*], Shiro Maeda[35,*], Takashi Kadowaki[10,*], Jong-Young Lee[1,*], Jer-Yuarn Wu[2,3,*], Yik Ying Teo[22,24,36,37,38,*], E Shyong Tai[39,40,41,*,¶], Xiao Ou Shu[5,*], Karen L Mohlke[9,*], Norihiro Kato[8,*], Bok-Ghee Han[1,*], Mark Seielstad[24,42,47,*,¶]

Cho et al. *Nature Genetics, 2012, 44; 67-72*

International collaboration network

Stage	Representative	Study	Ethnic group	Sample size		
				case	control	total
Stage 1	KNIH	KARE	Korean	1042	2943	3985
	NUS/SERI	SDCS/SP2(1)	Chinese	1082	1006	2088
		SDCS/SP2(2)	Chinese	928	939	1867
		SiMES	Malay	794	1240	2034
	NCGM	CAGE	Japanese	931	1404	2335
	VU/SCI	SDGS	Chinese	1019	1710	2729
	SINICA	TDS	Chinese	997	999	1996
	UNC	CLHNS	Filipino	159	1624	1783
	total			**6952**	**11865**	**18817**
Stage 2	RIKEN/UT	BBJ	Japanese	4470	3071	7541
	KNIH	H2T2DS	Korean	1183	1305	2488
	SJTU	SJTUDS	Chinese	190	198	388
	total			**5843**	**4574**	**10417**
Stage 3	NCGM	CAGE	Japanese	5253	5903	11156
	SJTU	SDIID/SDS	Chinese	3410	3412	6822
	CUHK	CUHKDS	Chinese	1477	1584	3061
	NTUH	NTUHDS	Chinese	1512	1512	3024
	SNUH	SNUHDS	Korean	632	761	1393
	total			**12284**	**13172**	**25456**
Overall	**AGEN**	**AGEN-T2D**	**East Asian**	**25079**	**29611**	**54690**

Stage1 : Discovery

✓**GWA meta-analysis combining 8 T2D GWA studies**
(6,952 cases vs. 11,865 controls)

$P < 5 \times 10^{-4}$

Stage2 : *in silico* replication

✓**Validation of 3,756 SNPs selected from Stage1**
(297 lead SNPs + their proxy SNPs)
in 3 T2D GWA studies (5,843 cases vs. 4,574 controls)
✓**Combined meta-analysis (Stages 1+2)**

$P < 10^{-5}$

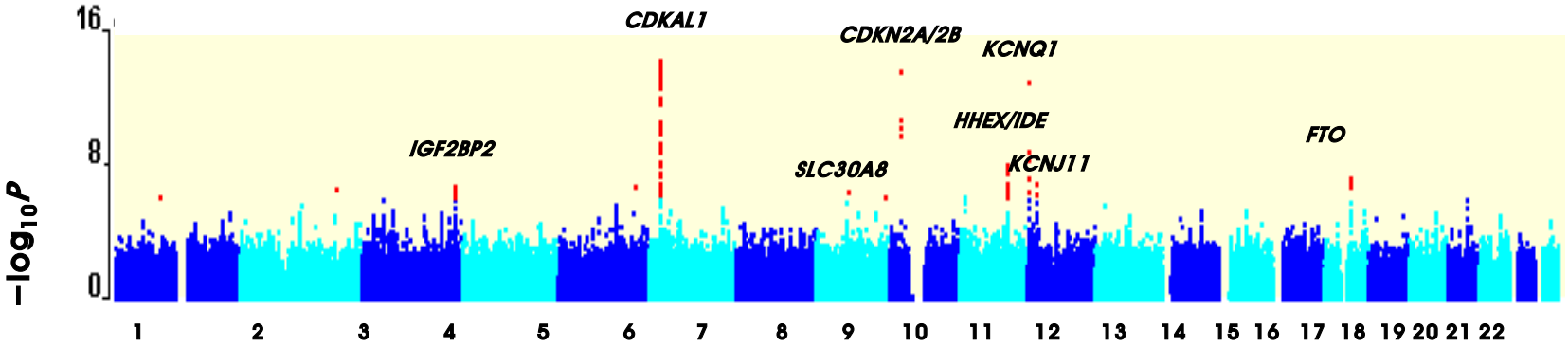
Stage3 : *de novo* replication

✓**Validation of 19 new SNPs selected from Stage2**
in 5 T2D studies (12,284 cases vs. 13,172 controls)
✓**Combined meta-analysis (Stages 1+2+3)**

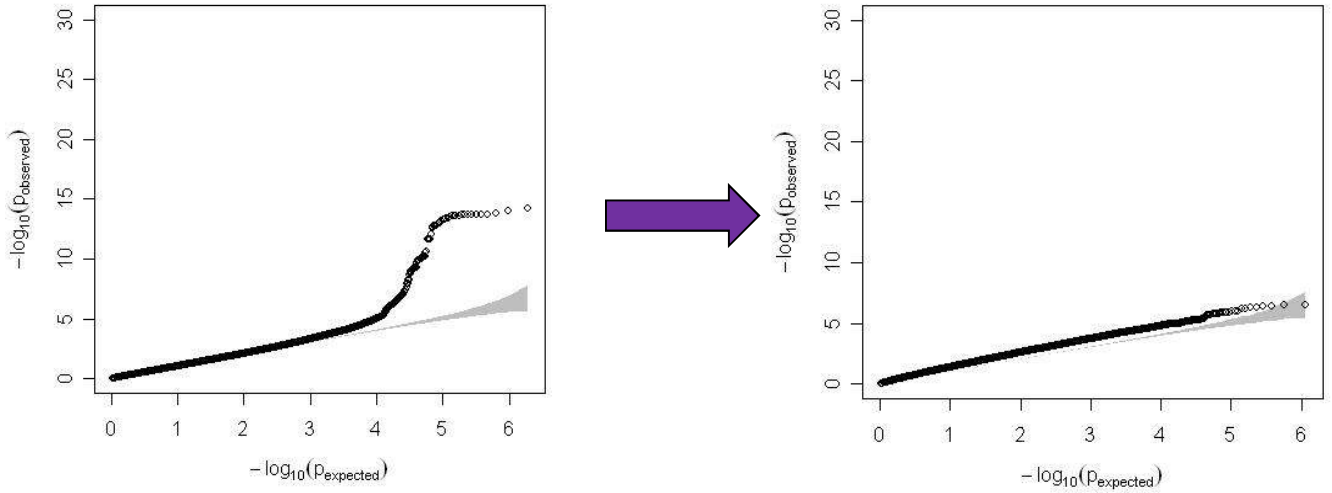
$P < 5 \times 10^{-8}$

8 novel T2D SNPs

Manhattan plot and QQ plots for the EA T2D stage 1 meta-analysis



Remove 20 known T2D loci



Association results for T2D

SNP	Chr	Position (bp)	Nearby gene	Risk allele	Other allele	Stage 1 (discovery)		Stage 2 (*in silico* replication)		Stage3 (*de novo* replication)		Combined (stage 1+2+3)	
						OR (CI)	*P*-value	OR (CI)	*P*-value	OR (CI)	*P*-value	OR (CI)	*P*-value
						up to 6,952 cases and 11,865 controls		up to 5,843 cases and 4,574 controls		up to 12,284 cases and 13,172 controls		up to 25,079 cases and 29,611 controls	
Loci showing strong evidence of association with T2D													
rs6815464	4	1299901	*MAEA*	c	g	1.09 (1.04-1.14)	8.21E-04	1.13 (1.07-1.20)	3.67E-05	1.16 (1.11-1.20)	4.15E-15	1.13 (1.10-1.16)	1.57E-20
rs7041847	9	4277466	*GLIS3*	a	g	1.09 (1.04-1.14)	1.29E-04	1.09 (1.03-1.15)	2.20E-03	1.11 (1.07-1.15)	2.89E-09	1.10 (1.07-1.13)	1.99E-14
rs6017317	20	42380380	*FITM2-R3HDML-HNF4A*	g	t	1.10 (1.05-1.15)	2.43E-05	1.07 (0.99-1.15)	8.42E-02	1.10 (1.06-1.14)	3.96E-07	1.09 (1.07-1.12)	1.12E-11
rs6467136	7	126952194	*GCC1*	g	a	1.12 (1.06-1.18)	6.47E-05	1.11 (1.04-1.18)	2.09E-03	1.10 (1.05-1.15)	2.31E-05	1.11 (1.07-1.14)	4.96E-11
rs831571	3	64023337	*PSMD6*	c	t	1.11 (1.06-1.17)	4.85E-06	1.06 (1.00-1.13)	4.46E-02	1.08 (1.05-1.12)	1.41E-05	1.09 (1.06-1.12)	8.41E-11
rs9470794	6	38214822	*ZFAND3*	c	t	1.11 (1.05-1.17)	1.45E-04	1.09 (1.02-1.17)	1.48E-02	1.16 (1.09-1.23)	3.20E-06	1.12 (1.08-1.16)	2.06E-10
rs3786897	19	38584848	*PEPD*	a	g	1.14 (1.08-1.20)	3.74E-06	1.05 (0.99-1.12)	1.28E-01	1.11 (1.04-1.17)	5.46E-04	1.10 (1.07-1.14)	1.30E-08
rs1535500	6	39392028	*KCNK16*	t	g	1.11 (1.06-1.16)	5.34E-06	1.07 (1.01-1.15)	3.33E-02	1.06 (1.02-1.10)	3.50E-03	1.08 (1.05-1.11)	2.30E-08
Loci showing moderate evidence of association with T2D													
rs16955379*	16	80046874	*CMIP*	c	t	1.13 (1.07-1.20)	2.20E-05	1.10 (1.03-1.17)	6.59E-03	1.05 (1.01-1.10)	2.19E-02	1.08 (1.05-1.12)	2.84E-07
rs17797882	16	77964419	*WWOX*	t	c	1.12 (1.05-1.18)	1.76E-04	1.09 (1.02-1.16)	1.21E-02	1.06 (1.01-1.11)	1.61E-02	1.08 (1.05-1.12)	9.49E-07

*Proxy SNP, rs9930117 (r^2=1), was genotyped in the stage3 CAGE study.

Summary

To identify susceptibility loci for T2D in east Asian populations, we followed our stage 1 meta-analysis of 8 T2D GWASs (6,952 cases and 11,865 controls) with a stage 2 *in silico* replication analysis (5,843 cases and 4,574 controls) and a stage 3 *de novo* replication analysis (12,284 cases and 13,172 controls)

8 new T2D loci from combined analysis;
GLIS2, PEPD, FITM2-R3HDML-HNF4A, KCNK16, MAEA, GCC1-PAX4, PSMD6 and ZFAND3

These findings, derived from an east Asian population, provide new perspectives on the etiology of T2D

Contents

- Introduction
- **META results**
  - OGTT
  - Metabolic traits
  - AGEN T2D
  - **AGEN FPG, *under process***

Identification of novel variants implicated in

Fasting plasma glucose in East Asians

Introduction

- Type 2 diabetes is relatively characterized by high circulating levels of fasting plasma glucose (FPG), one of the major risk factor.

- To identify novel T2D and FPG loci in a large sample using staged genome-wide association studies (GWAS) with meta-analysis.

- We performed three-stage design involving a discovery set (stage1) of 23,500 populations and *in silico* replication (stage 2) and *de novo* replication (stage 3) from independent East Asians populations.

Study Scheme

STAGE 1 GWAS meta-analysis of FPG	**STAGE 2** *In silico* replication	**STAGE 3** *De novo* replication , Follow-up analyses
Approximately 2.4 M directly or imputed genotyped SNPs were tested using the inverse variance meta-analysis in **23,500** individuals from **12** studies with East Asian ancestry	**3,500** individuals from **1** studies with East Asian ancestry	**21,300** individuals from **11** studies with East Asian ancestry Comparison published data and Magic lookup • Comparison AGEN-T2D • Pathway and protein-protein interaction analysis • eQTL

SNPs selected from Stage1 with $P < 10^{-4}$ forward stage2 → **Joint meta analyses of 22 SNPs (Stage1 + Stage2)** → Strongly associated SNP from **4** new regions showing the most compelling evidence for association ($P < 5 \times 10^{-8}$)

Previous FPG loci :
GCKR, SIX2-SIX3, G6PC2/ABCB11, CDKAL1, TMEM195, GCK, SLC30A8, GLIS3, CDKN2A/B, FADS1, MTNR1B, VPS13C

KNIH

Cohort characteristics

Stage	Study	Ethnic group	N	Study design	% Females	Mean age years	FPG	BMI
Stage 1 : discovery								
KNIH	KARE	Korean	7696	Population-based, prospective	4112	51.64 ± 8.80	4.62 ± 0.49	24.48 ± 3.08
	HEXA	Korean	3385	Population-based	1932	52.70 ± 8.17	5.00 ± 0.56	23.88 ± 2.89
NUS	SP2	Chinese	2282	Population-based	339	46.79 ± 10.21	4.73 ± 0.46	
					815	47.79 ± 11.11	4.70 ± 0.52	
					72	49.02 ± 12.51	4.77 ± 0.54	
VU/SCI	SWHS	Chinese	5284	Population-based	2034	49.37 ± 8.47	5.14 ± 1.52	
					3250			
GenSalt	GenSalt	Han Chinese	1832	Family-based				
NCGM	CAGE	Japanese	756	Case-control multicenter; population-base	345	65.14 ± 7.39	5.32 ± 0.51	
UNC	CLHNS	Filipino	1624	Population-based	1624	48.33 ± 6.10	5.03 ± 0.62	
CRC	CRC	Chinese	733	population-based	559	43.7 ± 7.8	4.87 ± 0.44	
Yonsei Uni		Korean	956	population-based				
			24548					
Stage 2 : in silico replication								
KNIH	Rural	Korean	3187	Population-based				
			3187					
Stage 3 : follow-up replication								
KNIH	Health2	Korean	5277	Population-based				
BBJ		Japanese	1835					
	Cheng	Han Chinese	3412					
TABARA	Nomura	Japanese	1946					
	Yokohama	Japanese	1868					
	Takashima	Japanese	1516					
	Shigaraki	Japanese	1964					
	Ohasama	Japanese	771					
	Matsuyama	Japanese	732					
	Toon	Japanese	1502					
HK		Han Chinese	474					
			21297					
Overall	**AGEN**	**East Asian**	**49032**					

Asian Genetic Epidemiology Network FPG (AGEN-FPG) : 23,500

- Korea Association Resource (KARE) Project : 7,696
- Singapore Prospective Study Program (SP2) : 2,282
- Shanghai Women's Health Study (SWHS) : 4,360
- Genetic Epidemiology Network of Salt-Sensitivity (GenSalt) : 1,832
- National Center for Global Health and Medicine, Japan (NCGM) : 756
- Cebu Longitudinal Health and Nutritional Survey (CLHNS) : 1,624
- Cardiometabolic Risk in Chinese (CRC) study : 733
- Health Examinees Study (HEXA) : 3,385
- Yonsei : 800

Association analysis

- Non-diabetes individuals were only tested for fasting plasma glucose (exclude known diabetes, individuals treated with anti-diabetic medicine and individuals with fasting glucose ≥ 7 mmol/l)

- Within each study, the rank-based inverse normal transformation (INT) was applied to FPG and performed linear regression analysis in an additive genetic model with adjustment for sex, BMI, ethnicity (if applicable) and principal component factors (if applicable)

- Because of the family design of GenSalt, family relationship was adjusted using the linear mixed model in which family ID was used as a random effect

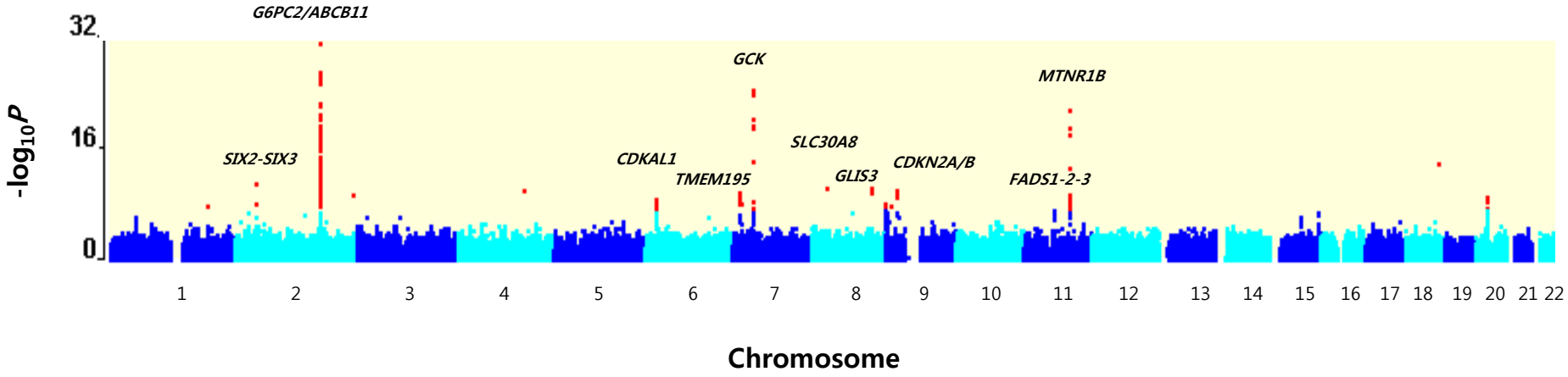
Stage 1 meta analysis (DISCOVERY):

- After quality control (exclude study-specific imputation quality), approximately 2.4 million directly genotyped or imputed markers were tested using the inverse variance meta-analysis assuming fixed effects. We applied genomic control correction at the each study and meta-analysis to avoid population stratification.

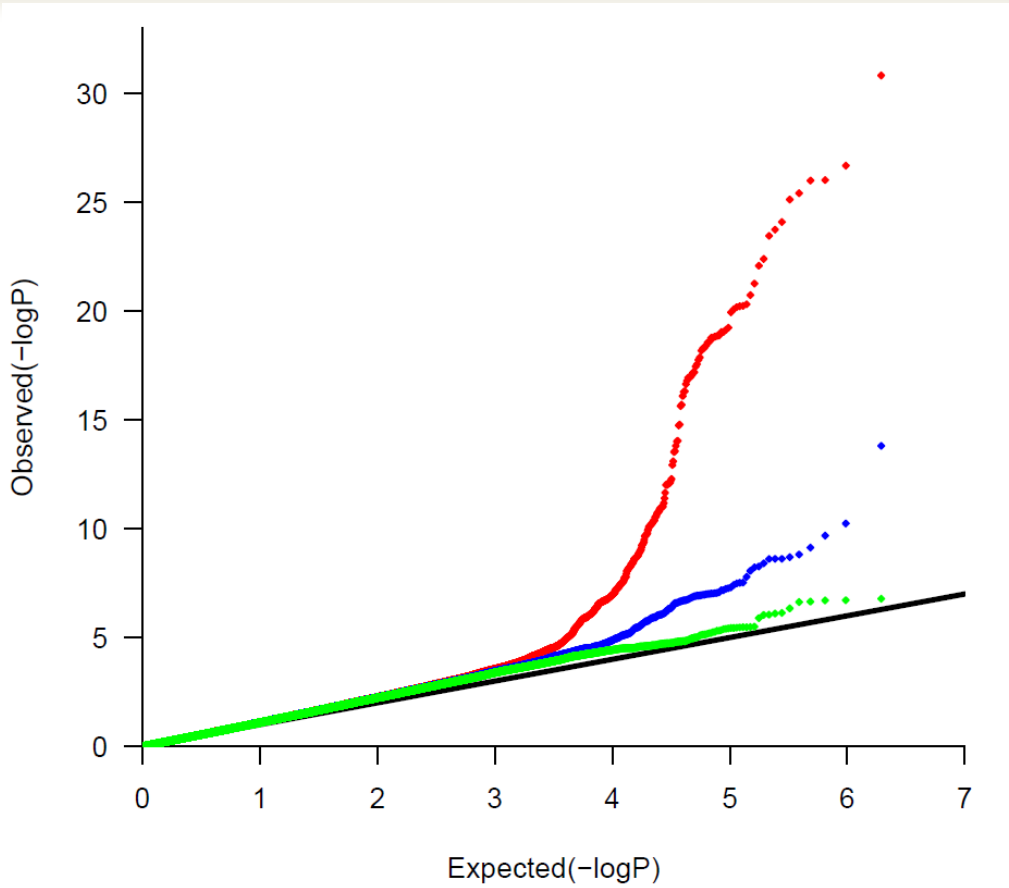
SNP selection for Stage1

- *P*-value < 1 X 10^{-5}
- Presented SNPs that are common (MAF >= 5%)
- subjects from more than eight study sites (of 12) GWAS were available for then meta analysis
- Heterogeneity *P*-value < 1 X 10^{-3}
- which is in LD with the reported SNP (r^2 > 0.2)

Manhattan plots for the AGEN stage 1 meta-analysis



Qunatile-quantile plot for the AGEN stage 1 meta-analysis



The expected null distribution is plotted along the **black** diagonal, the entire distribution of observed P values is plotted in **red** and a distribution that excludes the 12 reported previously loci is plotted is **blue** and excludes the newly discovered loci is plotted is **green**.

Ways to explain missing heritability

1. GWA meta-analysis & ethnic specific GWAS
- more common variants
- ethnic specific variants

2. Fine mapping of candidate T2D loci or Exome sequencing
- low frequency variants
- rare variants
- causal variants

3. Structural variants
- CNVs
- indels

4. Others
- GXG interaction, GXE interaction
- epigenetic modifications

Acknowledgements

국립보건연구원

형질연구과

유전체역학과

생물자원은행과

한복기, 조명찬

한국인 유전체역학 코호트

조남한, 신철, 강대희,

최보율, 정혜원, 성주헌

서울의대: 박경수

MERK: 김규찬

이화여자대학교: 김형래

경희대학교: 오범석

한림대학교 : 조윤신

All AGEN members