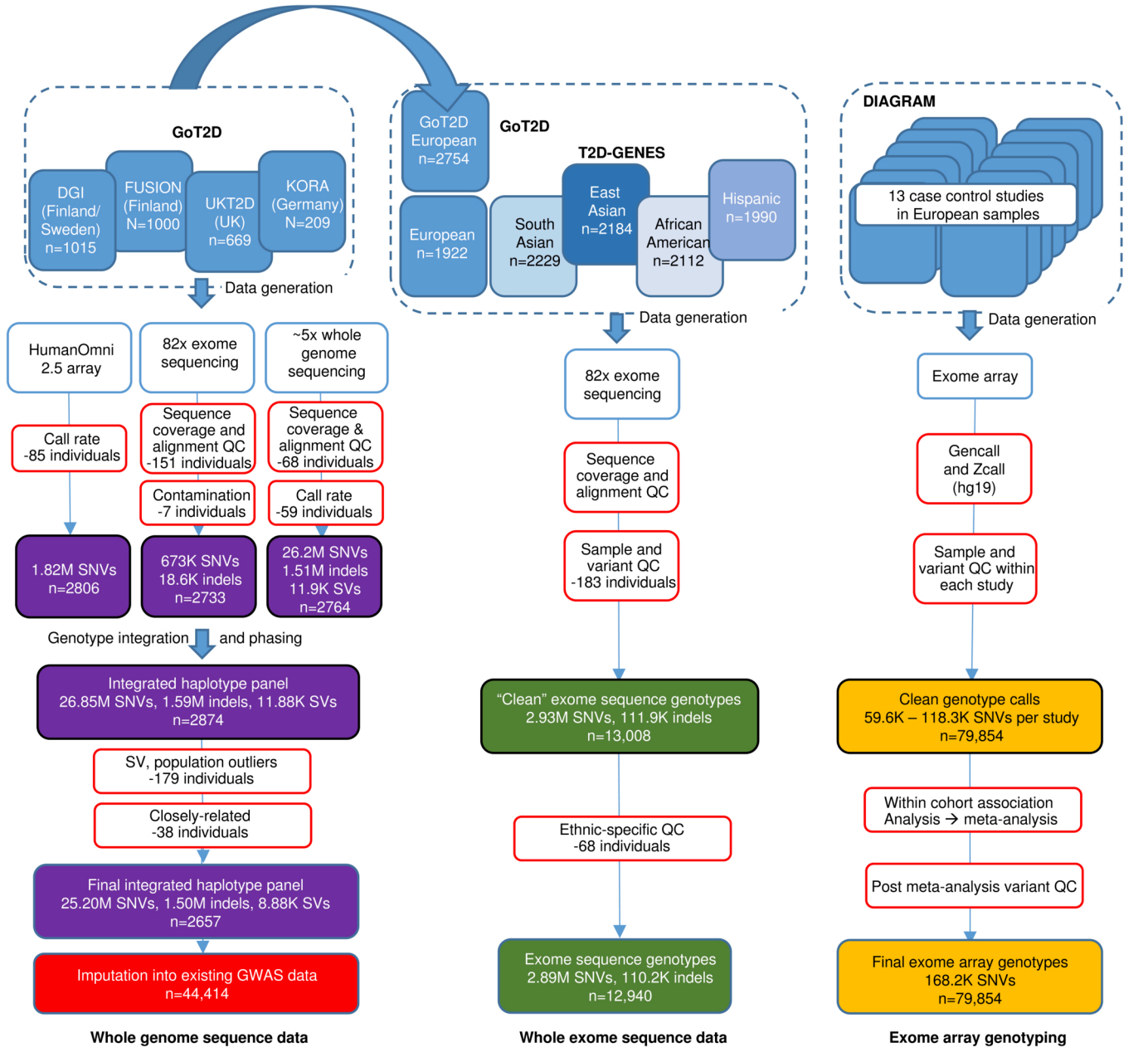

GoT2D
DGI (Finland/ Sweden) n=1015
FUSION (Finland) N=1000
UKT2D (UK) n=669
KORA (Germany) N=209
Data generation
HumanOmni 2.5 array
82x exome sequencing
~5x whole genome sequencing
Call rate -85 individuals
Sequence coverage and alignment QC -151 individuals
Sequence coverage & alignment QC -68 individuals
Contamination -7 individuals
Call rate -59 individuals
1.82M SNVs n=2806
673K SNVs 18.6K indels n=2733
26.2M SNVs 1.51M indels 11.9K SVs n=2764
Genotype integration and phasing
Integrated haplotype panel 26.85M SNVs, 1.59M indels, 11.88K SVs n=2874
SV, population outliers -179 individuals
Closely-related -38 individuals
Final integrated haplotype panel 25.20M SNVs, 1.50M indels, 8.88K SVs n=2657
Imputation into existing GWAS data n=44,414
Whole genome sequence data
GoT2D European n=2754
GoT2D
T2D-GENES
European n=1922
South Asian n=2229
East Asian n=2184
African American n=2112
Hispanic n=1990
Data generation
82x exome sequencing
Sequence coverage and alignment QC
Sample and variant QC -183 individuals
"Clean" exome sequence genotypes 2.93M SNVs, 111.9K indels n=13,008
Ethnic-specific QC -68 individuals
Exome sequence genotypes 2.89M SNVs, 110.2K indels n=12,940
Whole exome sequence data
DIAGRAM
13 case control studies in European samples
Data generation
Exome array
Gencall and Zcall (hg19)
Sample and variant QC within each study
Clean genotype calls 59.6K – 118.3K SNVs per study n=79,854
Within cohort association Analysis → meta-analysis
Post meta-analysis variant QC
Final exome array genotypes 168.2K SNVs n=79,854
Exome array genotyping

Ancestry	Study	Countries of Origin	Num. of Cases (% female)	Num. of Controls (% female)	Effective Sample Size
Whole Genome Sequencing Studies					
European	Finland-United States Investigation of NIDDM Genetics (FUSION) Study	Finland	493 (41.5)	486 (45.2)	979
European	Kooperative Gesundheitsforschung in der Region Augsburg (KORA)	Germany	101 (44.5)	104 (66.3)	205
European	Malmo-Botnia Study	Finland, Sweden	410 (51.5)	419 (44.1)	829
European	UK Type 2 Diabetes Genetics Consortium (UKT2D)	UK	322 (46.2)	322 (82.2)	644
	Total Whole Genome Sequence		**1,326**	**1,331**	**2657**
Genome-Wide Array Studies					
European	INTERACT	France, Germany, Italy, Netherlands, Spain, Sweden, UK	4624 (51.8)	4668 (64.2)	9292
European	Wellcome Trust Case Control Consortium (WTCCC)	UK	1586 (40.9)	2938 (50.8)	4120
European	Kooperative Gesundheitsforschung in der Region Augsburg (KORA)	Germany	993 (45.1)	2985 (52.2)	2980
European	Framingham Heart Study (FHS)	US	673 (42.6)	7660 (55.1)	2475
European	Finland-United States Investigation of NIDDM Genetics (FUSION) Study	Finland	1060 (43.1)	1090 (51.3)	2150
European	Diabetes Genetics Initiative (DGI)	Finland, Sweden	899 (46.6)	1057 (49.6)	1943
European	Estonian Genome Center, University of Tartu (EGCUT-OMNI)	Estonia	389 (58.6)	6013 (54.2)	1461
European	Diabetes Gene Discovery Group (DGDG)	France, Canada	677 (39.3)	697 (59.7)	1374
European	Mt Sinai BioMe Biobank Platform (BioMe (Illumina))	US	255 (29.0)	1647 (51.4)	883
European	Uppsala Longitudinal Study of Adult Men (ULSAM)	Sweden	166 (0)	953 (0)	565
European	Mt Sinai BioMe Biobank Platform (BioMe)	US	132 (26.5)	455 (34.7)	409
European	Prospective Investigation of the Vasculature in Uppsala Seniors (PIVUS)	Sweden	111 (41.4)	838 (51.2)	392
European	Estonian Genome Center, University of Tartu (EGCUT-370)	Estonia	80 (48.8)	1768 (51)	306
	Total Genome-Wide Array		**11,645**	**32,769**	**44,414**
	Total Whole Genome Sequence + Genome-Wide Array		**12,971**	**34,100**	**47,071**
Whole Exome Sequencing Studies					
African American	Jackson Heart Study	US	500 (66.6)	526 (63.3)	1,026
African American	Wake Forest School of Medicine Study	US	518 (59.5)	530 (56.0)	1,048
East Asian	Korea Association Research Project	Korea	526 (45.6)	561 (58.5)	1,086
East Asian	Singapore Diabetes Cohort Study; Singapore Prospective Study Program	Singapore (Chinese)	486 (52.1)	592 (61.3)	1,068
European	Ashkenazi	US, Israel	506 (47.0)	355 (56.9)	834
European	Metabolic Syndrome in Men Study (METSIM)	Finland	484 (0)	498 (0)	982
European	Finland-United States Investigation of NIDDM Genetics (FUSION) Study	Finland	472 (42.6)	476 (45.0)	948
European	Kooperative Gesundheitsforschung in der Region Augsburg (KORA)	Germany	97 (44.3)	90 (63.3)	186
European	UK Type 2 Diabetes Genetics Consortium (UKT2D)	UK	322 (45.7)	320 (82.8)	642
European	Malmo-Botnia Study	Finland, Sweden	478 (54.8)	443 (43.8)	920
Hispanic	San Antonio Family Heart Study, San Antonio Family Diabetes/ Gallbladder Study, Veterans Administration Genetic Epidemiology Study, and the Investigation of Nephropathy and Diabetes Study Family Component	US	272 (58.8)	218 (58.7)	484
Hispanic	Starr County, Texas	US	749 (59.7)	704 (71.9)	1,452
South Asian	London Life Sciences Population Study (LOLIPOP)	UK (Indian Asian)	531 (14.1)	538 (15.8)	1,068
South Asian	Singapore Indian Eye Study	Singapore (Indian Asian)	563 (44.4)	585 (49.2)	1,148
	Total Whole Exome Sequence		**6,504**	**6,436**	**12,892**
Exome Array Studies					
European	ADDITION; Steno Diabetes Centre (SDC); Health06; Health08; Vejle Biobank; Inter99	Denmark	5813 (40.0)	7987 (54.4)	13,458
European	Wellcome Trust Case Control Consortium (UK Type 2 Diabetes Consortium); Young Diabetics Study (YDX); Genetics of Diabetes and Audit Research Tayside Study (GoDARTS); Oxford Biobank; TwinsUK; 1958 Birth Cohort (BC58)	UK	3576 (51.7)	12675 (41.2)	11,156
European	Finland-United States Investigation of NIDDM Genetics (FUSION) Study; Finrisk2007; Metabolic Syndrome in Men Study (METSIM); Dose-Responses to Exercise Training (DR'sEXTRA); D2D2007	Finland	3593 (33.4)	8222 (26.0)	10,001
European	Malmo Diabetes Cohort (MDC); All New Diabetics in Skane (ANDIS)	Sweden	4633(41.0)	5404 (59.5)	9,978
European	Prevalence, Prediction and Prevention of Diabetes (PPP); Diabetes Register in Vaasa (DIREVA)	Finland	2910 (43.7)	4596 (53.7)	7,127
European	Nurses' Health Study (NHS)	US	1413 (100.0)	1695 (100.0)	3,082
European	Health Professionals Follow-up Study (HPFS)	US	1184 (0.0)	1287 (0.0)	2,467
European	The Exeter Family Study of Child Health (EFSOCH)	UK	1446 (39.0)	1567 (52.0)	3,008
European	Kooperative Gesundheitsforschung in der Region Augsburg (KORA)	Germany	933 (45.3)	2705 (51.7)	2,775
European	Estonian Genome Center at the University of Tartu (EGCUT)	Estonia	882 (43.7)	1506 (44.2)	2,225
European	Gene-Lifestyle Interactions and Complex Traits Involved in Elevated Disease Risk (GLACIER)	Sweden	960 (47.6)	957 (54.5)	1,917
European	Fenland cohort of the European Prospective Investigation of Cancer (Fen-EPIC)	UK	691(47.0)	1157 (54.5)	1,730
European	The Prospective Investigation of the Vasculature in Uppsala Seniors (PIVUS); Uppsala Longitudinal Study of Adult Men (ULSAM)	Sweden	271(16.9)	1791 (23.9)	942
	Total Exome Array		**28,305**	**51,549**	**69,886**
	Total Whole Exome Sequence + Exome Array		**34,809**	**57,985**	**82,758**

a

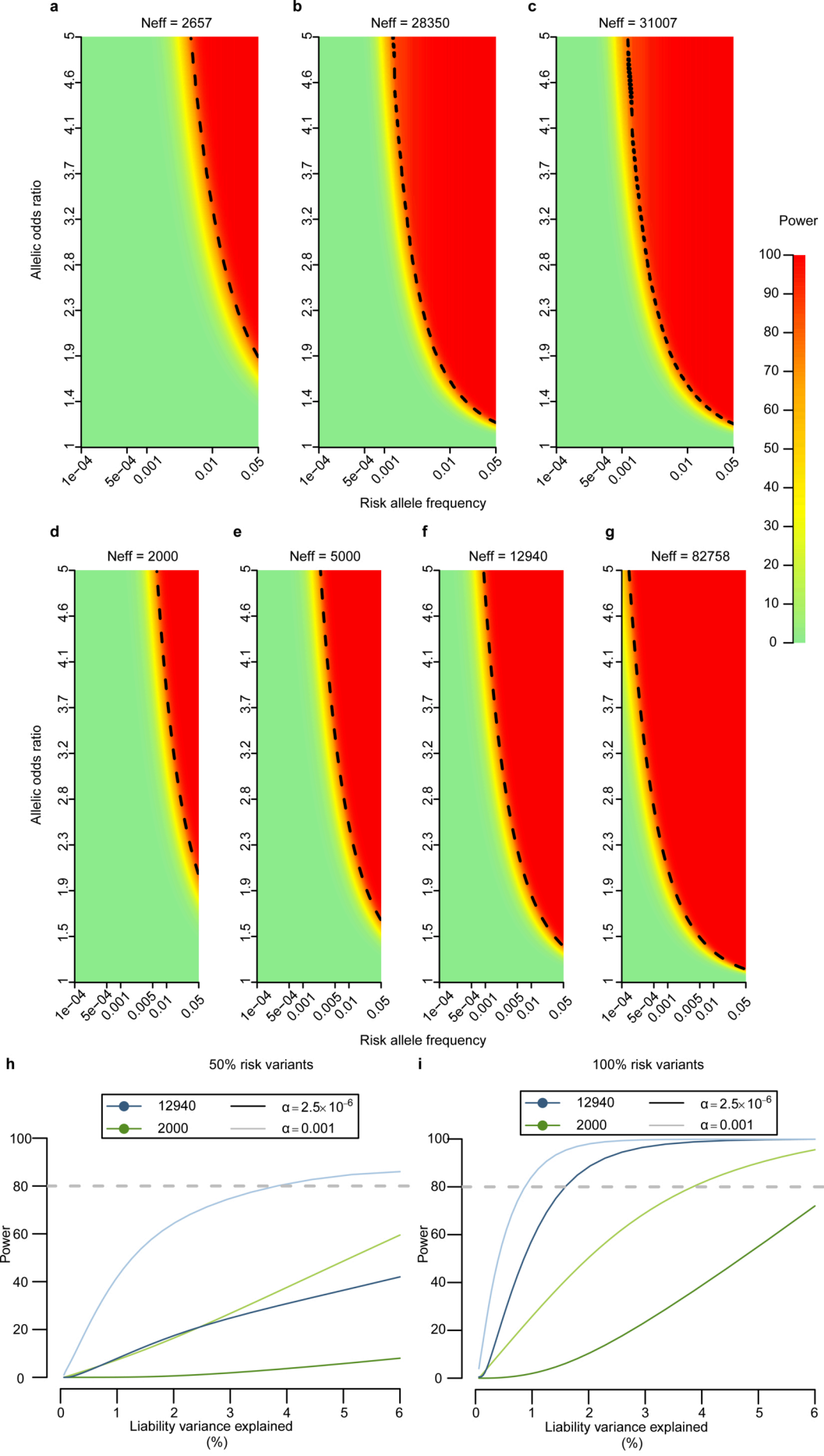
Genomes integrated panel

	SNV	Indel	SV
Variant Type N (%total)	25.2M (94%)	1.50M (5.6%)	8,876 (0.03%)
	Coding		Non-coding
Function N (%total)	888K (3.3%)		25.8M (97%)
	Rare (MAF<0.5%)	Low frequency (0.5<MAF<5%)	Common (MAF>5%)
Frequency spectrum N (%total)	6.26M (23%)	4.16M (16%)	16.3M (61%)
	b137		Novel
dbSNV N (%total)	14.6M (55%)		12.1M (45%)

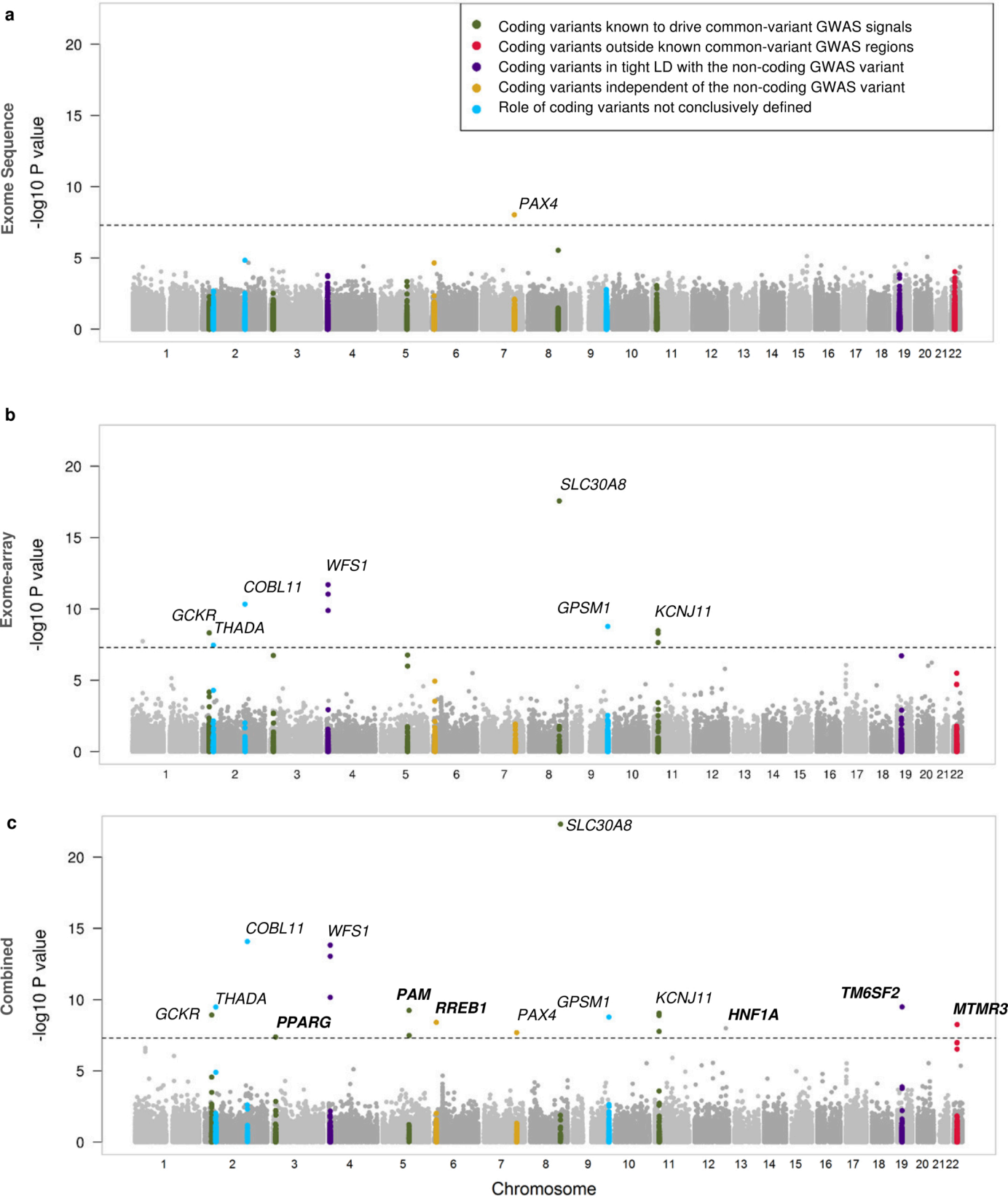
b

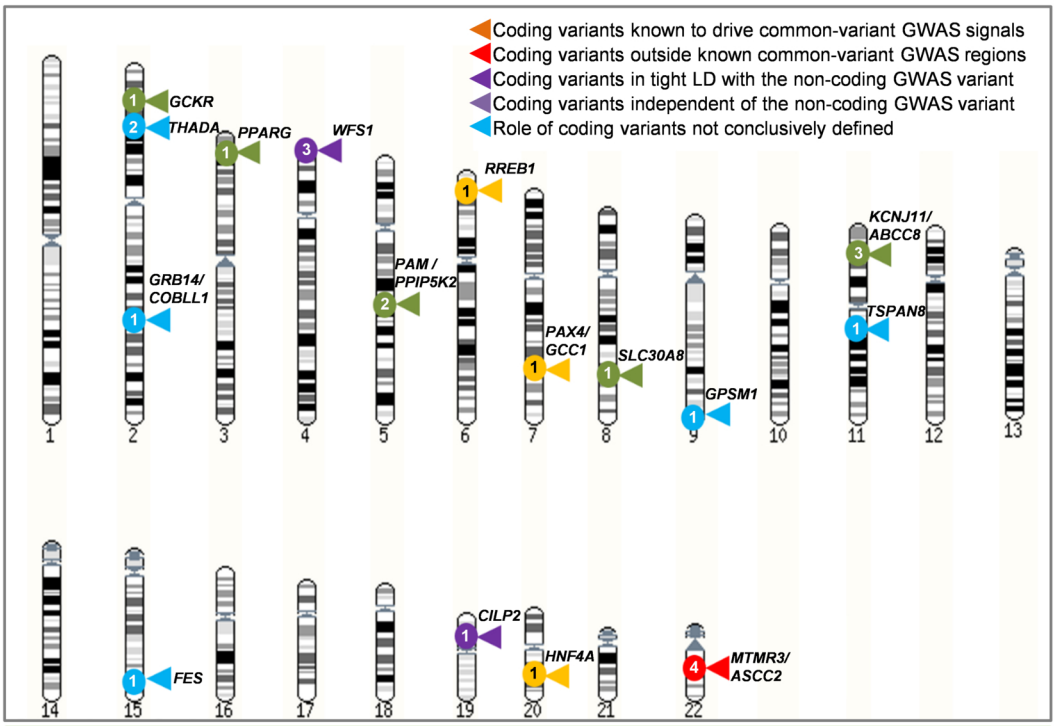
Exome sequence data

	All samples	African-American	East-Asian	European	Hispanic	South-Asian
Samples:	13,008	2,086	2,165	4,579	1,959	2,219
T2D cases	6,504	1,018	1,012	2,359	1,021	1,094
T2D controls	6,436	1,056	1,153	2,182	922	1,123
Excluded from association analysis	*68*	*12*	*0*	*38*	*16*	*2*
Coverage:						
Coding:						
Mean (*Mc*) per gene	81.7 ±23.7	83.2 ±24.0	84.6 ±23.8	78.6 ±23.3	83.8 ±24.1	78.2 ±23.2
# of genes with *Mc*<20	368	302	302	351	269	325
Non-coding:						
Mean per gene	59.0 ±21.0	60.9 ±21.5	62.2 ±21.6	57.5 ±20.6	59.2 ±21.2	55.4 ±20.3
# of genes with *Mc*<20	1,150	738	731	1,102	804	945
Variant annotations:						
Synonymous SNV	627,630	237,430	178,232	192,282	156,231	211,218
Missense SNV	1,110,897	354,797	296,707	327,049	231,351	344,191
Start SNV	2,055	593	523	639	384	583
Nonsense SNV	26,321	7,188	6,668	8,030	4,660	7,339
Frameshift INDEL	26,901	6,605	6,159	7,515	4,155	6,609
Inframe INDEL	11,090	3,471	2,963	3,145	2,068	3,165
3'UTR SNV, INDEL	65,013	24,583	19,149	21,102	16,959	22,177
5'UTR SNV, INDEL	43,965	16,920	13,520	15,562	11,634	15,595
Intron SNV, INDEL	931,449	352,398	270,564	296,970	243,139	314,810
Essential splicing SNV, INDEL	14,286	3,648	3,454	4,108	2,301	3,744
Other splicing SNV, INDEL	128,644	45,876	35,413	38,263	30,301	41,122
Non-coding RNA SNV, INDEL	18,113	7,247	5,996	6,715	5,084	6,706
Intergenic SNV, INDEL	37,345	14,335	11,498	13,614	10,700	12,937
All	**3,043,709**	**1,075,091**	**850,846**	**934,994**	**718,967**	**990,196**
Coding frequency spectrum:						
Rare (MAF<0.5%)	95.79%	83.30%	90.06%	89.19%	84.56%	89.89%
private	*77.93%*	*53.79%*	*65.47%*	*51.80%*	*37.26%*	*61.55%*
cosmopolitan	*0.35%*	*1.80%*	*3.02%*	*1.88%*	*2.24%*	*1.73%*
Low frequency (0.5<MAF<5%)	2.57%	10.36%	4.61%	5.52%	8.21%	5.10%
private	*0.17%*	*1.43%*	*1.10%*	*0.26%*	*0.52%*	*1.02%*
cosmopolitan	*0.60%*	*1.50%*	*1.54%*	*1.94%*	*2.74%*	*1.62%*
Common (MAF>5%)	1.65%	6.35%	5.33%	5.29%	7.23%	5.00%
private	*0.09%*	*0.00%*	*0.00%*	*0.00%*	*0.01%*	*0.00%*
cosmopolitan	*1.50%*	*4.35%*	*5.17%*	*4.97%*	*6.88%*	*4.86%*
Intron/UTR frequency spectrum:						
Rare (MAF<0.5%)	94.09%	78.68%	86.91%	86.17%	81.43%	86.68%
private	*74.76%*	*49.81%*	*61.36%*	*45.26%*	*31.03%*	*56.96%*
cosmopolitan	*0.46%*	*2.07%*	*3.98%*	*2.49%*	*2.66%*	*2.19%*
Low frequency (0.5<MAF<5%)	3.52%	12.57%	5.63%	6.51%	9.43%	6.32%
private	*0.25%*	*1.74%*	*1.25%*	*0.29%*	*0.47%*	*1.18%*
cosmopolitan	*0.80%*	*1.81%*	*2.11%*	*2.53%*	*3.30%*	*2.17%*
Common (MAF>5%)	2.39%	8.76%	7.46%	7.32%	9.14%	7.00%
private	*0.15%*	*0.00%*	*0.00%*	*0.01%*	*0.00%*	*0.00%*
cosmopolitan	*2.17%*	*5.94%*	*7.26%*	*6.93%*	*8.77%*	*6.81%*


a
Neff = 2657
b
Neff = 28350
c
Neff = 31007
Allelic odds ratio
Risk allele frequency
Power
d
Neff = 2000
e
Neff = 5000
f
Neff = 12940
g
Neff = 82758
h
50% risk variants
i
100% risk variants
12940
2000
α = 2.5× 10−6
α = 0.001
Liability variance explained
(%)

Locus	Variant	MAF	Unconditional and conditional association p-values				Interpretation
Non-coding variant associations characterized in 38,738 samples in GoT2D genome wide imputed meta analysis							
IRS1			**rs78124264**	**rs7578326**	**rs2943640**	**rs2943641**	The association signal rs78124264 and the GWAS SNPs at this locus are distinct. Signals are not extinguished in reciprocal conditional analysis. Previous GWAS signals are not mediated through rs78124264, which represents a distinct association signal at this locus.
	rs78124264	0.022	*$8.5x10^{-5}$*	$2.5x10^{-7}$ *	$2.5x10^{-7}$ *	$2.5x10^{-7}$ *	
	rs7578326 ¶	0.35	$1.2x10^{-7}$	*$1.1x10^{-5}$*	n.d.	n.d.	
	rs2943640 ¶	0.35	$2.5x10^{-11}$	n.d.	*$4.5x10^{-10}$*	n.d.	
	rs2943641 ¶	0.36	$9.0x10^{-12}$	n.d.	n.d.	*$1.5x10^{-10}$*	
PPARG			**rs79856023**	**rs1801282**			The association signal rs79856023 and the GWAS SNP at this locus are distinct. Signals are not extinguished in reciprocal conditional analysis. Previous GWAS signal is not mediated through rs79856023, which represents a distinct association signal at this locus.
	rs79856023	0.022	*$1.2x10^{-4}$*	$9.2x10^{-7}$			
	rs1801282 ¶	0.13	$1.6x10^{-6}$	*$1.2x10^{-5}$*			
Coding variant associations characterized in 28,305 cases and 51,549 controls typed on exome array							
PAM			**Asp563Gly (*PAM*)**	**Ser1207Gly (*PPIP5K2*)**			Association signals for *PAM* Asp563Gly and *PPIP5K2* Ser1207Gly are indistinguishable in reciprocal conditional analysis. Gene biology, as well as previous reports of additional *PAM* variants associated with T2D in Icelandic cohorts, highlights *PAM* as the probable transcript at this locus.
	Asp563Gly	0.054	*$1.7x10^{-7}$*	0.24			
	Ser1207Gly	0.054	0.30	*$1.0x10^{-6}$*			
MTMR3-ASCC2			**Asn960Ser (*MTMR3*)**	**Val123Ile (*ASCC2*)**	**Asp407His (*ASCC2*)**	**Pro423Ser (*ASCC2*)**	Association signals for the *MTMR3* and *ASCC2* coding variants are indistinguishable in reciprocal conditional analysis. The *MTMR3* Asn960Ser variant has the strongest signal, and highlights *MTMR3* as the most likely effector transcript at this locus.
	Asn960Ser	0.083	*$3.2x10^{-6}$*	0.022	0.027	0.022	
	Val123Ile	0.083	0.15	*$2.0x10^{-5}$*	0.066	0.76	
	Asp407His	0.083	0.18	0.99	*$1.9x10^{-5}$*	0.88	
	Pro423Ser	0.083	0.18	0.67	0.98	*$2.0x10^{-5}$*	
KCNJ11-ABCC8			**Val337Ile (*KCNJ11*)**	**Lys23Glu (*KCNJ11*)**	**Ala1369Ser (*ABCC8*)**		Association signals for *KCNJ11* Val337Ile and Lys23Glu and *ABCC8* Ala1369Ser are indistinguishable in reciprocal conditional analysis. The relative causal contributions of the two genes, making up the two components of the sulfonlylurea-responsive potassium channel, are indistinguishable on statistical grounds.
	Val337Ile	0.40	*$3.4x10^{-9}$*	0.17	0.049		
	Lys23Glu	0.40	0.48	*$5.1x10^{-9}$*	0.082		
	Ala1369Ser	0.40	0.68	0.84	*$2.3x10^{-8}$*		
WFS1			**Val333Ile**	**Asn500Asn**	**Arg611His**	**rs4689388**	Association signals for the *WFS1* coding variants are indistinguishable from each other and the previously reported non-coding GWAS SNP at this locus in reciprocal conditional analysis. *WFS1* is the likely effector transcript for the non-coding GWAS signal at this locus, although the causal variant in the gene is unclear.
	Val333Ile	0.30	*$9.3x10^{-12}$*	0.024	0.00070	0.0030	
	Asn500Asn	0.41	0.0070	*$2.0x10^{-12}$*	0.0049	0.027	
	Arg611His	0.47	0.020	0.62	*$1.3x10^{-10}$*	0.19	
	rs4689388 ¶	0.43	0.011	0.62	0.024	*$2.3x10^{-11}$*	
CILP2-TM6SF2			**Glu167Lys (*TM6SF2*)**	**rs10401969**			Association signals for *TM6SF2* Glu167Lys and the previously reported non-coding GWAS SNP at this locus are indistinguishable from each other in reciprocal conditional analysis. *TM6SF2* is the probable effector transcript for the non-coding GWAS signal at this locus, with the effect mediated through Glu167Lys.
	Glu167Lys	0.082	*$1.9x10^{-7}$*	0.52			
	rs10401969 ¶	0.083	0.62	*$4.2x10^{-7}$*			
GRB14-COBLL1			**Asn939Asp (*COBLL1*)**	**rs13389219**			Association signals for *COBLL1* Asn939Asp and the previously reported non-coding GWAS SNP at this locus are partially correlated. The association signal for the GWAS signal is not entirely extinguished in reciprocal conditional analysis. *COBLL1* is a candidate effector transcript for the GWAS signal at this locus.
	Asn939Asp	0.12	*$4.7x10^{-11}$*	$3.00x10^{-5}$			
	rs13389219 ¶	0.39	$7.0x10^{-5}$	*$1.9x10^{-10}$*			
Coding variant associations characterized in 44,414 samples in GoT2D genome wide imputed meta analysis							
THADA			**Cys1605Tyr**	**rs10203174**			Association signals *THADA* Cys1605Tyr and the GWAS SNP are partially correlated. The association signal for the GWAS SNP is not entirely extinguished in reciprocal conditional analysis. *THADA* is a candidate effector transcript for the GWAS signal at this locus.
	Cys1605Tyr	0.10	*0.00035*	0.92			
	rs10203174 ¶	0.10	0.0063	*$5.7x10^{-6}$*			
RREB1			**Asp1171Asn**	**rs9502570**			The association signals of *RREB1* Asp1171Asn and the GWAS SNP at this locus are distinct. The association signal is not extinguished in reciprocal conditional analysis. Previous GWAS signal is not mediated through *RREB1* Asp1171Asn. *RREB1* Asp1171Asn represents a distinct association signal at this locus.
	Asp1171Asn	0.11	*0.0018*	0.0017			
	rs9502570 ¶	0.28	0.0037	*0.0042*			


a
Coding variants known to drive common-variant GWAS signals
Coding variants outside known common-variant GWAS regions
Coding variants in tight LD with the non-coding GWAS variant
Coding variants independent of the non-coding GWAS variant
Role of coding variants not conclusively defined
Exome Sequence
-log10 P value
PAX4
b
Exome-array
-log10 P value
SLC30A8
WFS1
COBL11
GCKR
THADA
GPSM1
KCNJ11
c
Combined
-log10 P value
SLC30A8
COBL11
WFS1
GCKR
THADA
PPARG
PAM
RREB1
PAX4
GPSM1
KCNJ11
HNF1A
TM6SF2
MTMR3
Chromosome



Coding variants known to drive GWAS signals

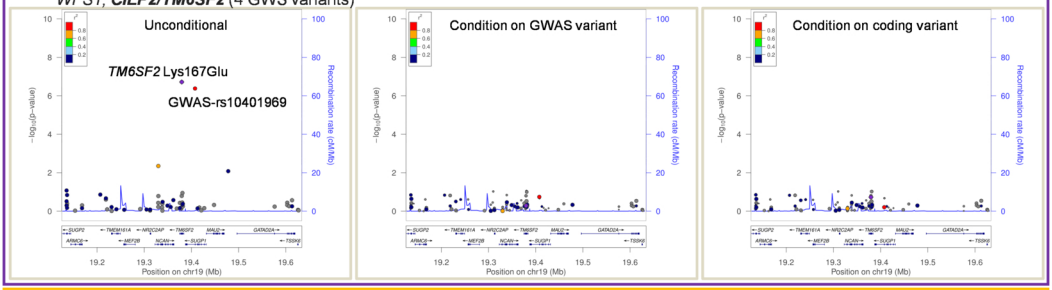
GCKR, *PPARG*, *PAMIPPIP5K2*, *SLC30A8*, *KCNJ11/ABCC8* (8 GWS variants)

Coding variants outside known GWAS regions

*MTMR3/ASCC2** (1 GWS variant and 3 highly correlated variants within the same region)

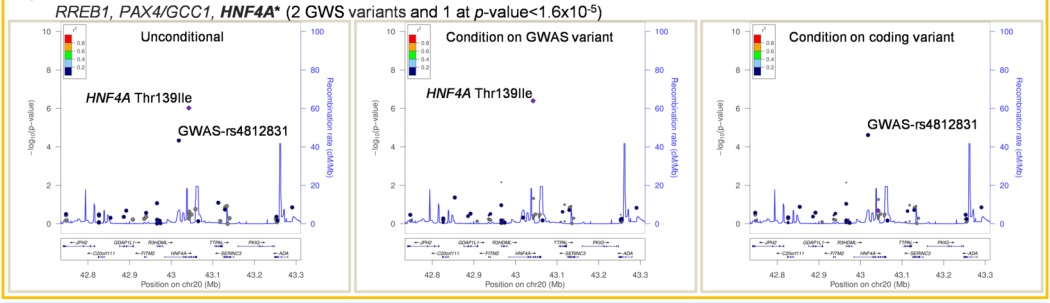
Coding variants in tight LD with the non-coding GWAS variant

WFS1, ***CILP2/TM6SF2*** (4 GWS variants)



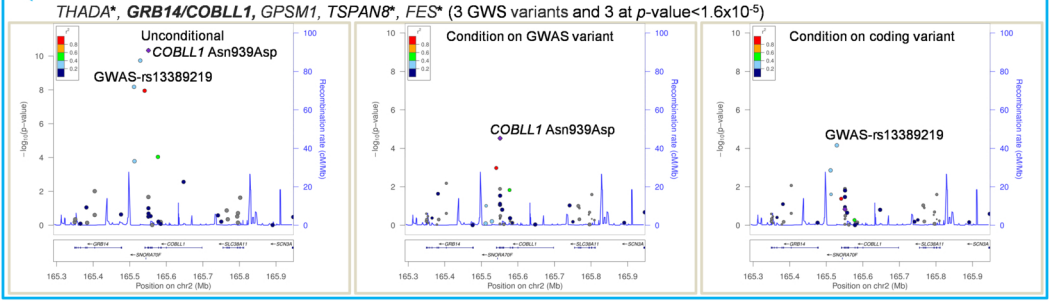
Coding variants independent of the non-coding GWAS variant

RREB1, *PAX4/GCC1*, ***HNF4A**** (2 GWS variants and 1 at p-value<1.6x10^{-5})



Role of coding variants not conclusively defined

*THADA**, ***GRB14/COBLL1***, *GPSM1*, *TSPAN8**, *FES** (3 GWS variants and 3 at p-value<1.6x10^{-5})



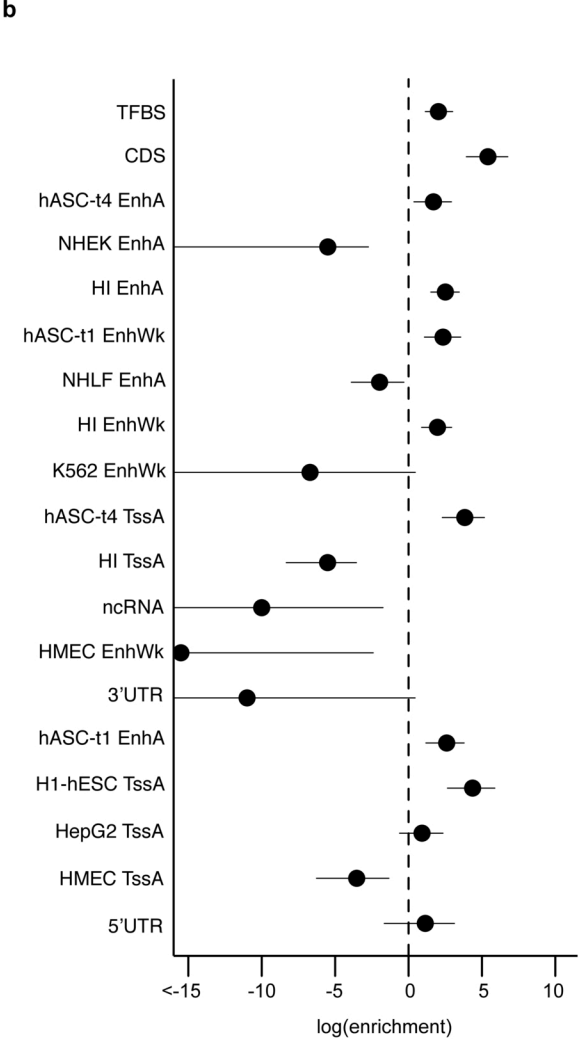
			Index SNV association signal		Synthetic association by missense variants				Synthetic association by all low-frequency and rare variants across 5Mb region					
			Index SNV association signal before inclusion of missense variants			Index SNV association signal after inclusion of missense variants					Index SNV association after inclusion of single best variant		Testing groups of low-frequency and rare variants	
Gene	Index SNV	MAF	OR [95% interval]	*p*-value	Number Missense Variants	OR [95% interval]	*p*-value	Relative likelihood of LF model	Best LF Variant	MAF	OR [95% interval]	*p*-value	n_1	n_2
TCF7L2	10:114758349	0.27	1.75 [1.54-1.99]	2.80×10^{-18}	6	1.73 [1.52-1.97]	2.33×10^{-17}	1.8×10^{-17}	10:114787948	1.6%	1.72 [1.51-1.95]	1.62×10^{-16}	>50	35
ADCY5	3:123065778	0.19	0.69 [0.60-0.79]	1.12×10^{-7}	13	0.70 [0.61-0.81]	9.00×10^{-7}	9.7×10^{-8}	3:123096056	2.5%	0.71 [0.61-0.82]	3.04×10^{-6}	13	6
IRS1	2:227093745	0.36	0.76 [0.68-0.86]	2.80×10^{-6}	5	0.77 [0.69-0.86]	4.30×10^{-6}	4.5×10^{-7}	2:226993370	1.7%	0.78 [0.70-0.88]	2.19×10^{-5}	12	6
KCNQ1	11:2847069	0.45	0.78 [0.70-0.87]	1.22×10^{-5}	>50	0.84 [0.75-0.94]	2.07×10^{-3}	1.0×10^{-7}	11:2825279	4.7%	0.81 [0.71-0.91]	3.19×10^{-4}	16	6
CDC123-CAMK1D	10:12307894	0.25	1.33 [1.17-1.52]	1.19×10^{-5}	4	1.30 [1.13-1.50]	2.06×10^{-4}	7.1×10^{-5}	10:12325477	3.8%	1.29 [1.12-1.48]	3.03×10^{-4}	10	5
CDKN2A-CDKN2B	9:22137685	0.28	1.28 [1.14-1.45]	4.52×10^{-5}	4	1.27 [1.13-1.43]	9.28×10^{-5}	4.3×10^{-5}	9:22133773	3.5%	1.25 [1.10-1.41]	5.98×10^{-4}	22	7
IGF2BP2	3:185511687	0.32	1.25 [1.11-1.41]	1.65×10^{-4}	14	1.21 [1.07-1.36]	2.12×10^{-3}	3.0×10^{-4}	3:185550500	4.1%	1.20 [1.07-1.36]	2.91×10^{-3}	8	3
KLHDC5	12:27965150	0.17	0.76 [0.66-0.88]	2.19×10^{-4}	3	0.77 [0.66-0.89]	4.45×10^{-4}	1.2×10^{-3}	12:27832062	2.0%	0.80 [0.68-0.92]	3.04×10^{-3}	10	4
SLC30A8	8:118184783	0.33	0.81 [0.72-0.91]	2.95×10^{-4}	2	0.81 [0.72-0.91]	3.73×10^{-4}	0.02	8:117964024	2.2%	0.83 [0.73-0.93]	1.23×10^{-3}	17	6
CDKAL1	6:20694884	0.18	1.28 [1.11-1.48]	6.05×10^{-4}	1	1.28 [1.11-1.48]	7.57×10^{-4}	0.007	6:20718780	2.8%	1.23 [1.06-1.43]	7.71×10^{-3}	9	3

a

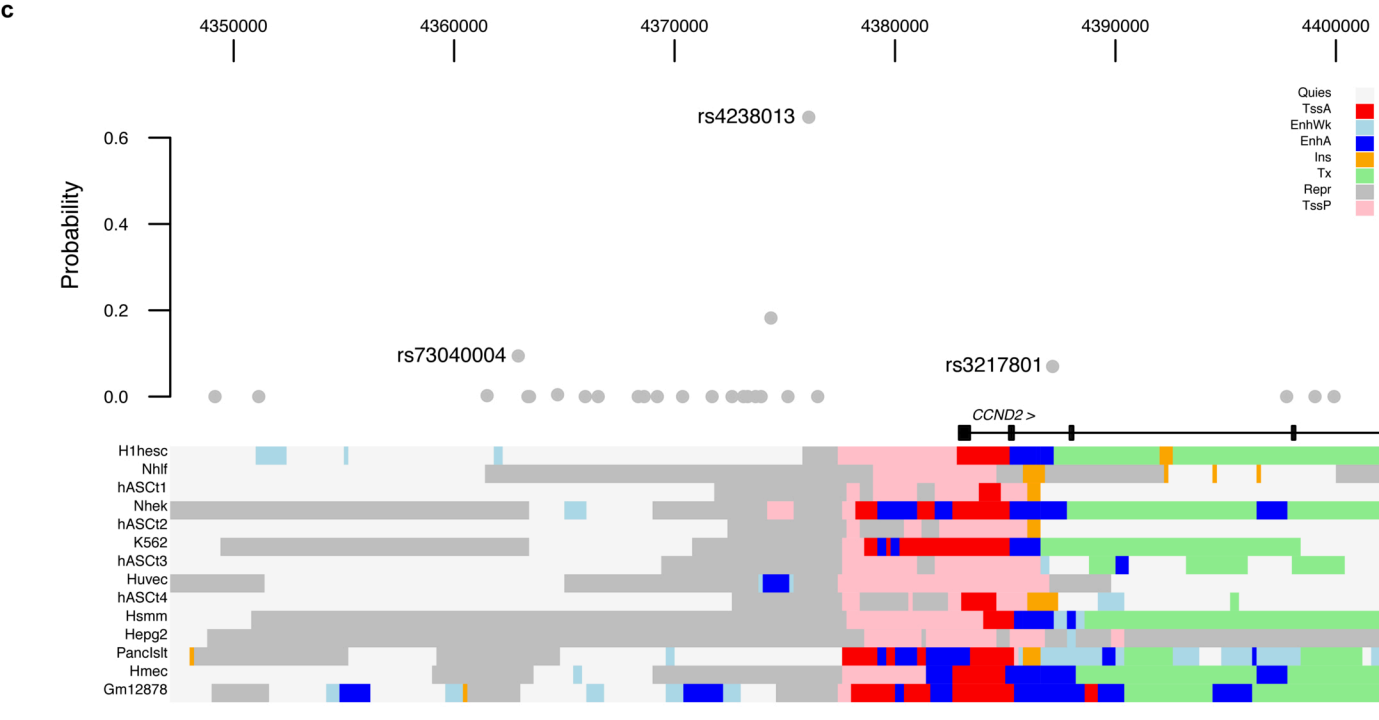
Annotations included in genome enrichment analysis

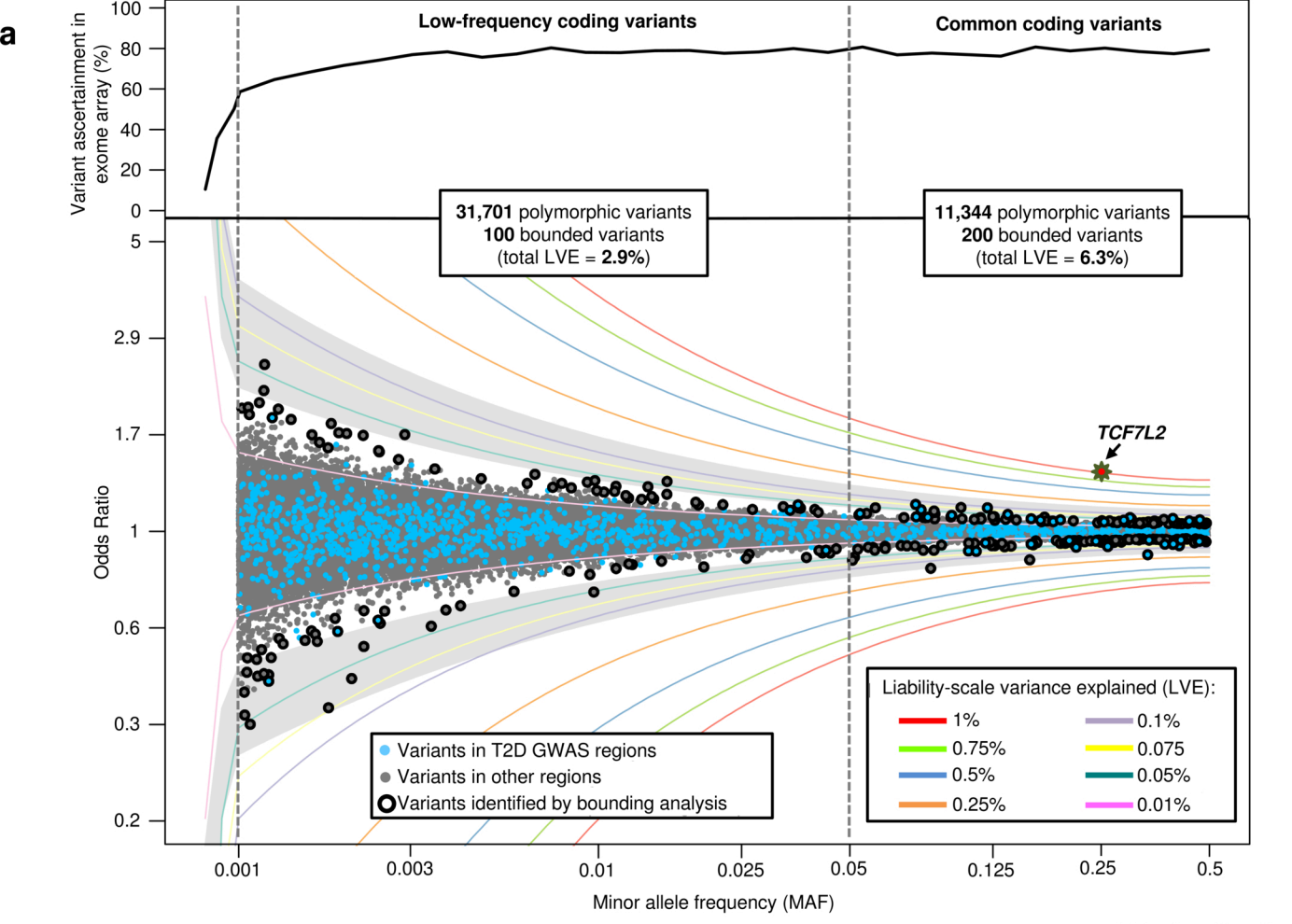
Category	Source	Description
CDS	GENCODE	protein-coding gene exons
5' UTR	GENCODE	protein-coding gene 5' UTR
3' UTR	GENCODE	protein-coding gene 3' UTR
ncRNA	GENCODE	non-coding RNA
TFBS	ENCODE, Pasquali et al	transcription factor binding
Gm12878 TssA	ENCODE	Lymphoblastoid cell line active promoter
Gm12878 EnhA	ENCODE	Lymphoblastoid cell line active enhancer
Gm12878 EnhWk	ENCODE	Lymphoblastoid cell line weak enhancer
H1-hESC TssA	ENCODE	Embryonic stem cell active promoter
H1-hESC EnhA	ENCODE	Embryonic stem cell active enhancer
H1-hESC EnhWk	ENCODE	Embryonic stem cell weak enhancer
hASC-t1 TssA	Mikkelsen et al	Pre-adipose stromal cell active promoter
hASC-t1 EnhA	Mikkelsen et al	Pre-adipose stromal cell active enhancer
hASC-t1 EnhWk	Mikkelsen et al	Pre-adipose stromal cell weak enhancer
hASC-t4 TssA	Mikkelsen et al	Mature adipose stromal cell active promoter
hASC-t4 EnhA	Mikkelsen et al	Mature adipose stromal cell active enhancer
hASC-t4 EnhWk	Mikkelsen et al	Mature adipose stromal cell weak enhancer
HepG2 TssA	ENCODE	Hepatocellular carcinoma cell line active promoter
HepG2 EnhA	ENCODE	Hepatocellular carcinoma cell line active enhancer
HepG2 EnhWk	ENCODE	Hepatocellular carcinoma cell line weak enhancer
HI TssA	Pasquali et al	Pancreatic islet cell active promoter
HI EnhA	Pasquali et al	Pancreatic islet cell active enhancer
HI EnhWk	Pasquali et al	Pancreatic islet cell weak enhancer
HMEC TssA	ENCODE	Mammary epithelial cell active promoter
HMEC EnhA	ENCODE	Mammary epithelial cell active enhancer
HMEC EnhWk	ENCODE	Mammary epithelial cell weak enhancer
Hsmm TssA	ENCODE	Skeletal muscle myoblast cell active promoter
Hsmm EnhA	ENCODE	Skeletal muscle myoblast cell active enhancer
Hsmm EnhWk	ENCODE	Skeletal muscle myoblast cell weak enhancer
HUVEC TssA	ENCODE	Vascular epithelial cell active promoter
HUVEC EnhA	ENCODE	Vascular epithelial cell active enhancer
HUVEC EnhWk	ENCODE	Vascular epithelial cell weak enhancer
K562 TssA	ENCODE	Erythroleukemia cell line active promoter
K562 EnhA	ENCODE	Erythroleukemia cell line active enhancer
K562 EnhWk	ENCODE	Erythroleukemia cell line weak enhancer
NHEK TssA	ENCODE	Epidermal keratinocyte cell active promoter
NHEK EnhA	ENCODE	Epidermal keratinocyte cell active enhancer
NHEK EnhWk	ENCODE	Epidermal keratinocyte cell weak enhancer
NHLF TssA	ENCODE	Lung fibroblast cell active promoter
NHLF EnhA	ENCODE	Lung fibroblast cell active enhancer
NHLF EnhWk	ENCODE	Lung fibroblast cell weak enhancer

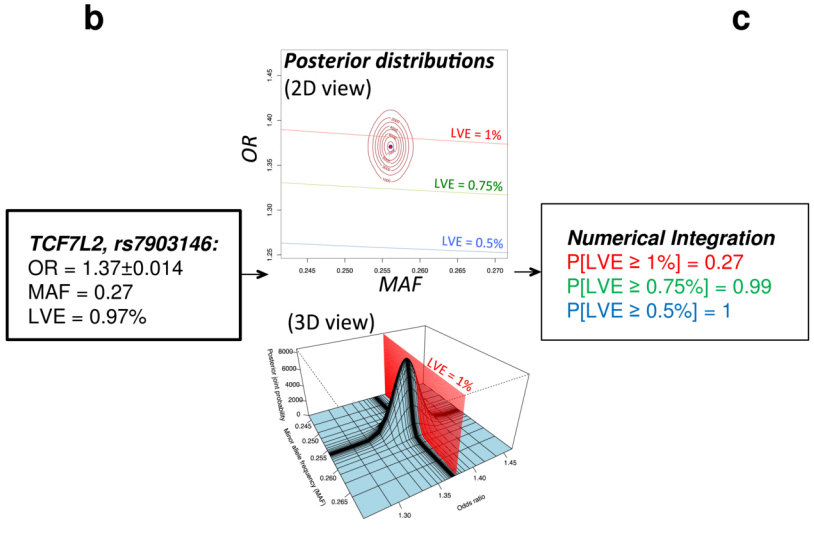
b

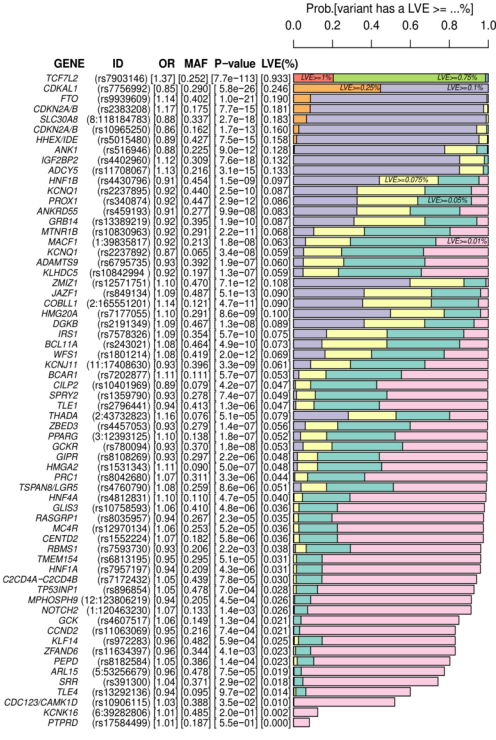


c




a
Low-frequency coding variants
Common coding variants
Variant ascertainment in exome array (%)
31,701 polymorphic variants
100 bounded variants
(total LVE = 2.9%)
11,344 polymorphic variants
200 bounded variants
(total LVE = 6.3%)
TCF7L2
Odds Ratio
Liability-scale variance explained (LVE):
1%
0.75%
0.5%
0.25%
0.1%
0.075
0.05%
0.01%
Variants in T2D GWAS regions
Variants in other regions
Variants identified by bounding analysis
Minor allele frequency (MAF)


b
TCF7L2, rs7903146:
OR = 1.37±0.014
MAF = 0.27
LVE = 0.97%
Posterior distributions
(2D view)
LVE = 1%
LVE = 0.75%
LVE = 0.5%
OR
MAF
(3D view)
Numerical Integration
P[LVE ≥ 1%] = 0.27
P[LVE ≥ 0.75%] = 0.99
P[LVE ≥ 0.5%] = 1


c
Prob.[variant has a LVE >= ...%]
GENE ID OR MAF P-value LVE(%)