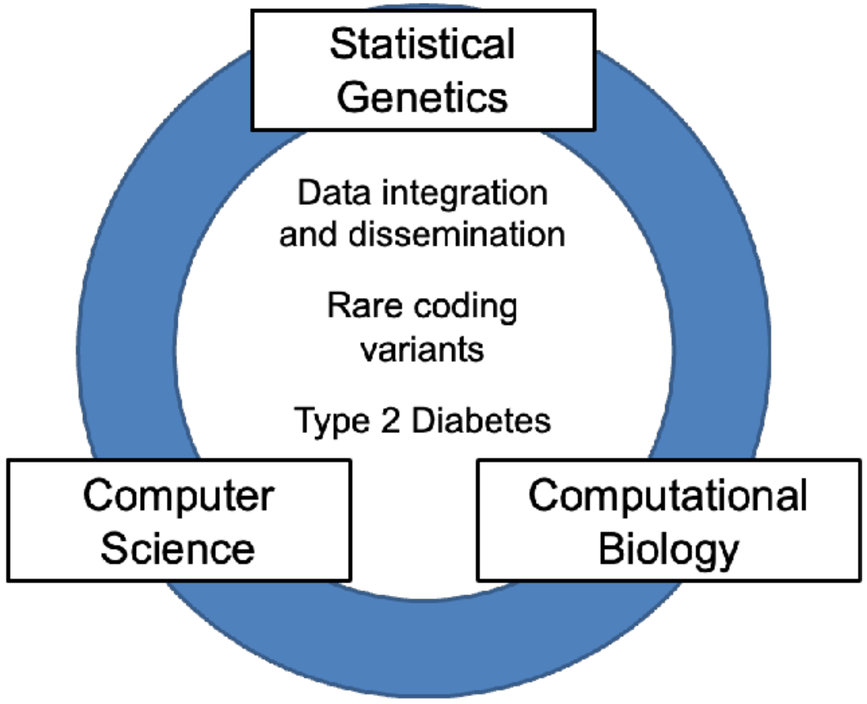
Insights into type 2 diabetes from rare coding variants

Jason Flannick
flannick@broadinstitute.org
flannicklab.org



BROAD INSTITUTE

Boston Children's Hospital
Until every child is well™

HARVARD MEDICAL SCHOOL

FlannickLab



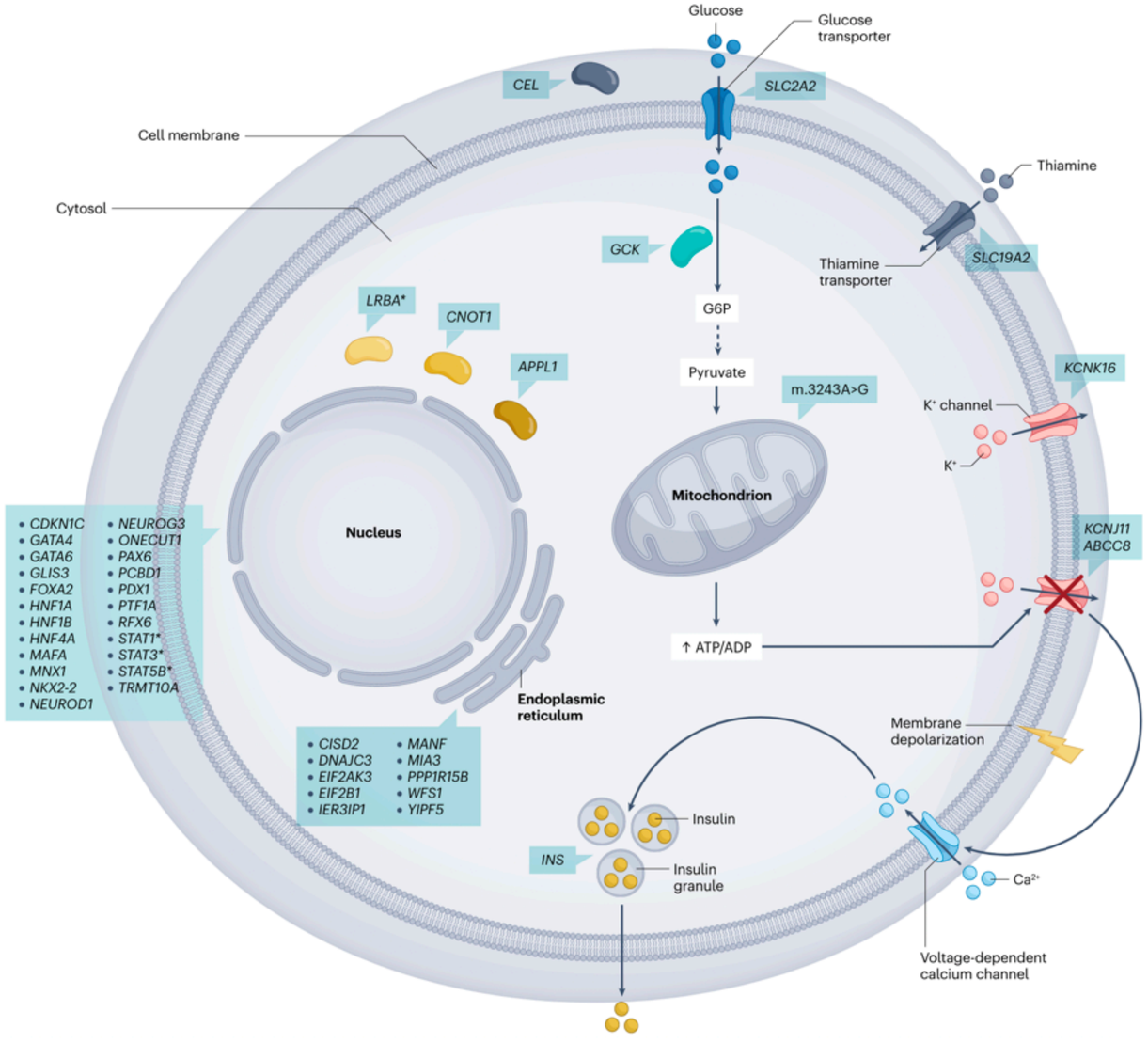


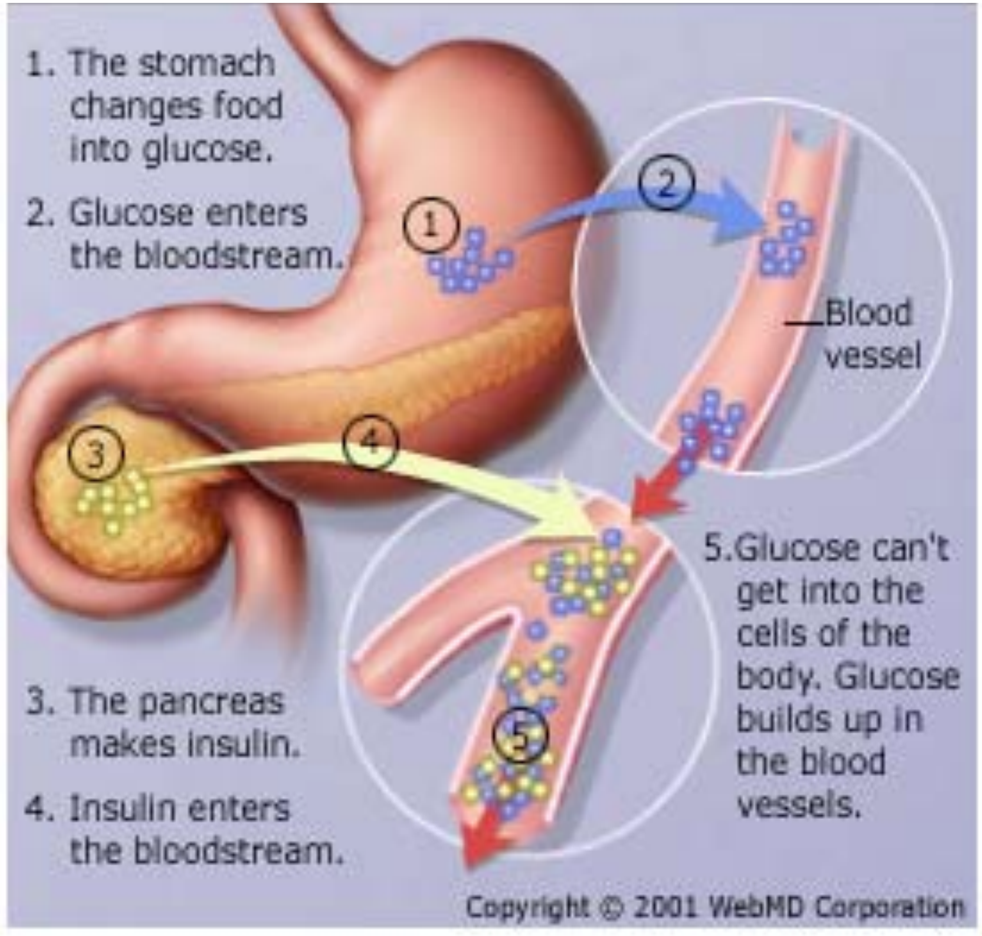


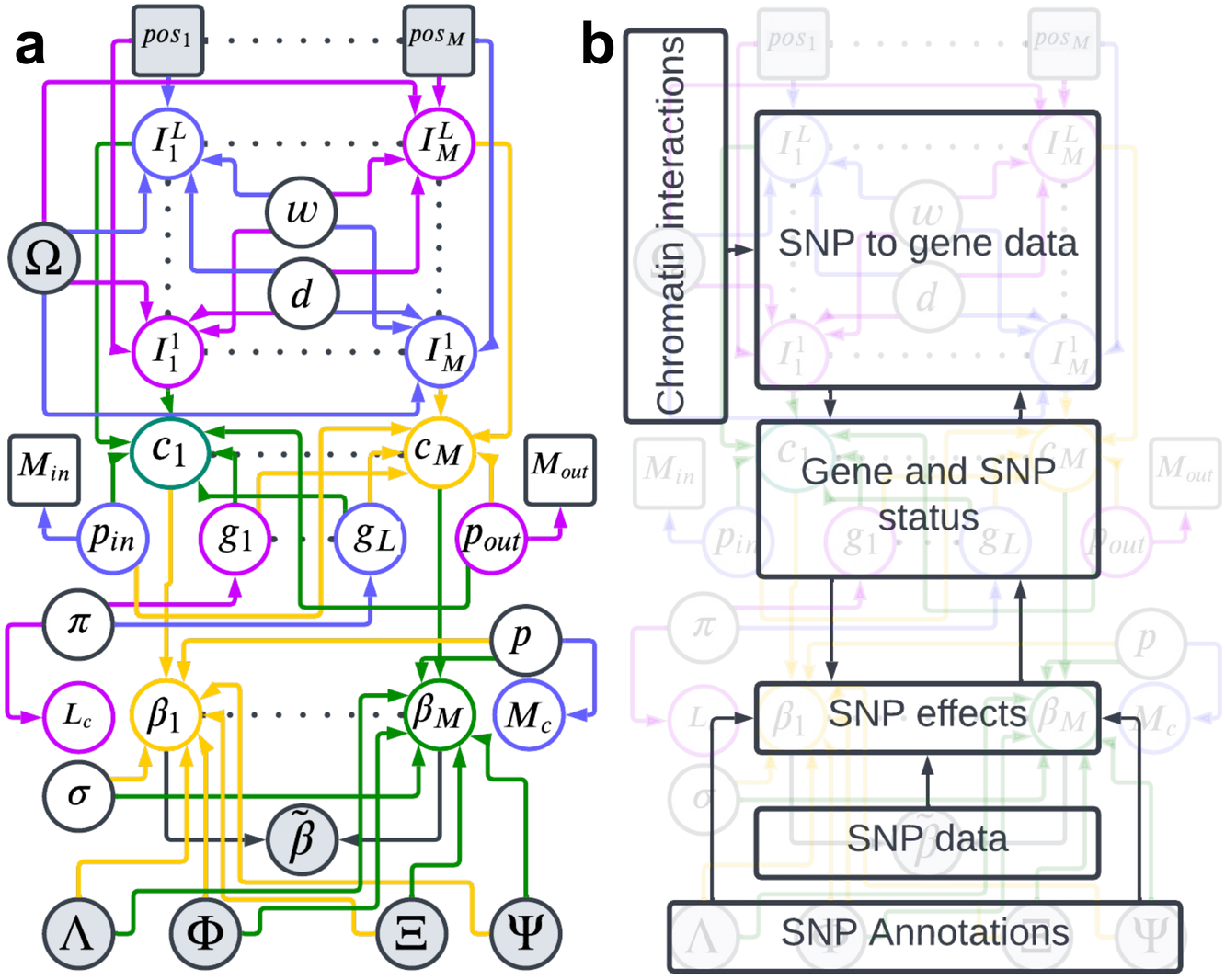








Glucose
Glucose transporter
CEL
SLC2A2
Cell membrane
Cytosol
Thiamine
Thiamine transporter
SLC19A2
GCK
G6P
LRBA*
CNOT1
APPL1
Pyruvate
m.3243A>G
KCNK16
K+ channel
K+
Mitochondrion
Nucleus
• CDKN1C • GATA4 • GATA6 • GLIS3 • FOXA2 • HNF1A • HNF1B • HNF4A • MAFA • MNX1 • NKX2-2 • NEUROD1
• NEUROG3 • ONECUT1 • PAX6 • PCBD1 • PDX1 • PTF1A • RFX6 • STAT1* • STAT3* • STAT5B* • TRMT10A
KCNJ11 ABCC8
↑ ATP/ADP
Endoplasmic reticulum
• CISD2 • DNAJC3 • EIF2AK3 • EIF2B1 • IER3IP1
• MANF • MIA3 • PPP1R15B • WFS1 • YIPF5
Membrane depolarization
Insulin
INS
Insulin granule
Ca2+
Voltage-dependent calcium channel


1. The stomach changes food into glucose.
2. Glucose enters the bloodstream.
3. The pancreas makes insulin.
4. Insulin enters the bloodstream.
Blood vessel
5. Glucose can't get into the cells of the body. Glucose builds up in the blood vessels.
Copyright © 2001 WebMD Corporation


a
b
Chromatin interactions
SNP to gene data
Gene and SNP status
SNP effects
SNP data
SNP Annotations

Rare coding variants

Back in 2008…

The case of the missing heritability

When scientists opened up the human genome, they expected to find the genetic components of common traits and diseases. But they were nowhere to be seen. **Brendan Maher** shines a light on six places where the missing loot could be stashed away.

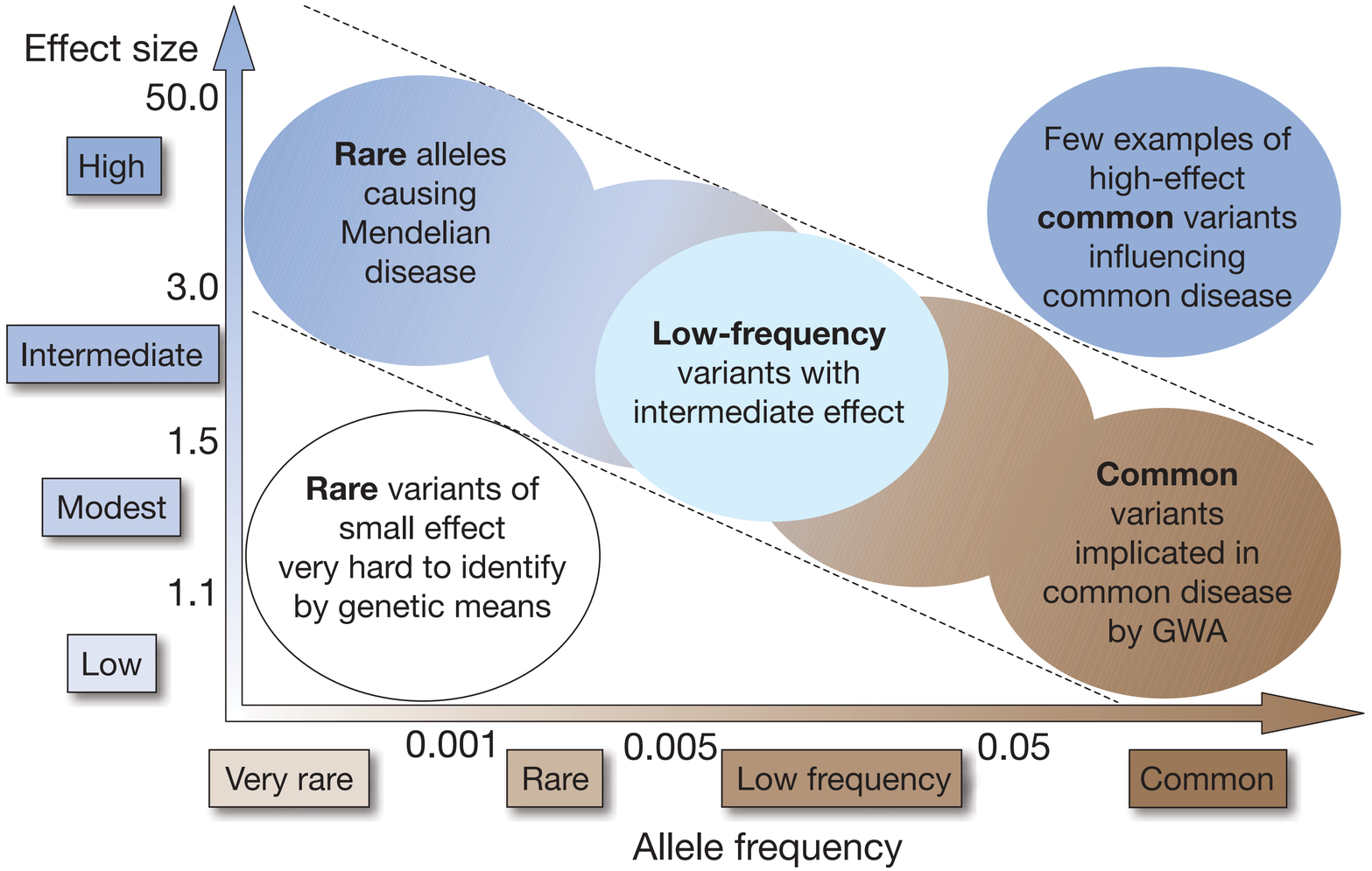
Table 1 | Estimates of heritability and number of loci for several complex traits

Disease	Number of loci	Proportion of heritability explained	Heritability measure
Age-related macular degeneration[72]	5	50%	Sibling recurrence risk
Crohn's disease[21]	32	20%	Genetic risk (liability)
Systemic lupus erythematosus[73]	6	15%	Sibling recurrence risk
Type 2 diabetes[74]	18	6%	Sibling recurrence risk
HDL cholesterol[75]	7	5.2%	Residual* phenotypic variance
Height[15]	40	5%	Phenotypic variance
Early onset myocardial infarction[76]	9	2.8%	Phenotypic variance
Fasting glucose[77]	4	1.5%	Phenotypic variance

* Residual is after adjustment for age, gender, diabetes.

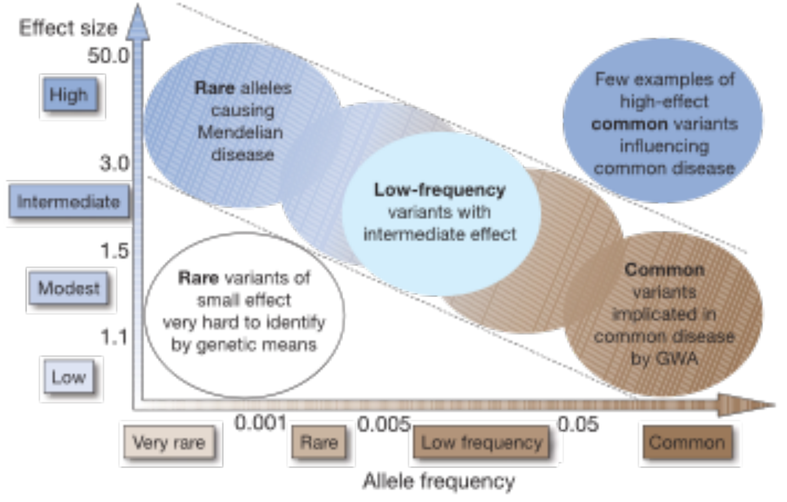
Manolio et al, Nature 2009

Classes of genetic variation



The case of the missing heritability

When scientists opened up the human genome, they expected to find the genetic components of common traits and diseases. But they were nowhere to be seen. **Brendan Maher** shines a light on six places where the missing loot could be stashed away.



Are Rare Variants Responsible for Susceptibility to Complex Diseases?

Jonathan K. Pritchard

Department of Statistics, University of Oxford, Oxford

Common Genetic Variation and Human Traits

David B. Goldstein, Ph.D.

Genetic Heterogeneity in Human Disease

Jon McClellan[1,*] and Mary-Claire King[2,*]
[1]Department of Psychiatry
[2]Departments of Medicine and Genome Sciences
University of Washington, Seattle, WA 98195-7720, USA
*Correspondence: drjack@uw.edu (J.M.), mcking@uw.edu (M.-C.K.)
DOI 10.1016/j.cell.2010.03.032

Clan Genomics and the Complex Architecture of Human Disease

James R. Lupski,[1,2,3,*] John W. Belmont,[1,2] Eric Boerwinkle,[4,5] and Richard A. Gibbs[1,5,*]

OPEN ACCESS Freely available online

PLoS BIOLOGY

Rare Variants Create Synthetic Genome-Wide Associations

Samuel P. Dickson[1,2], Kai Wang[3], Ian Krantz[3,4,5], Hakon Hakonarson[3,4,5], David B. Goldstein[1]*

Common and rare variants in multifactorial susceptibility to common diseases

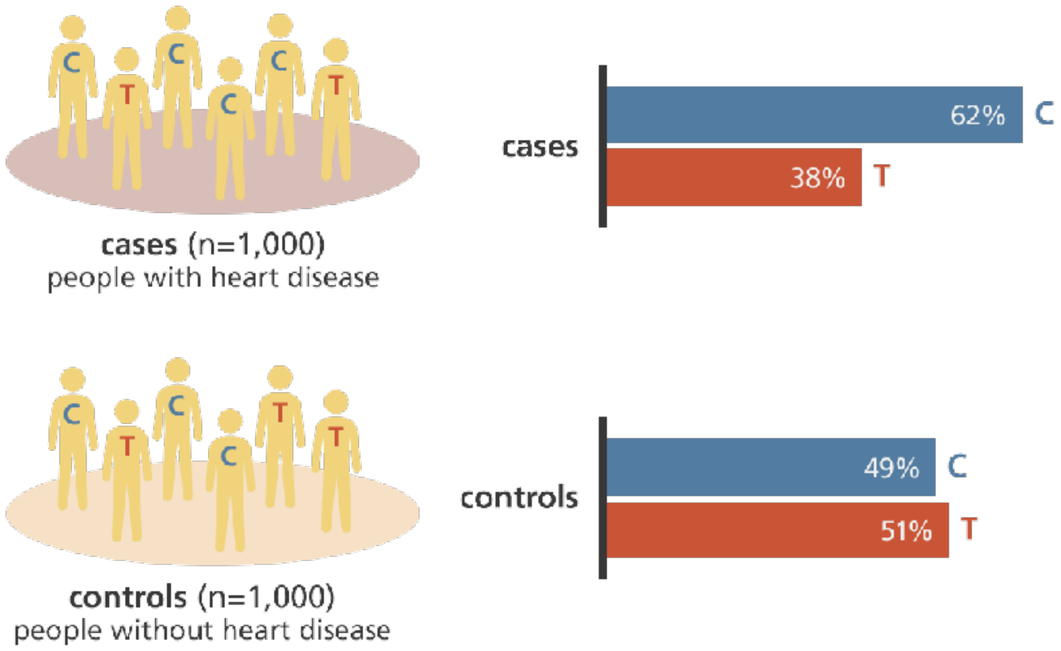
Walter Bodmer & Carolina Bonilla

APPLICATIONS OF NEXT-GENERATION SEQUENCING

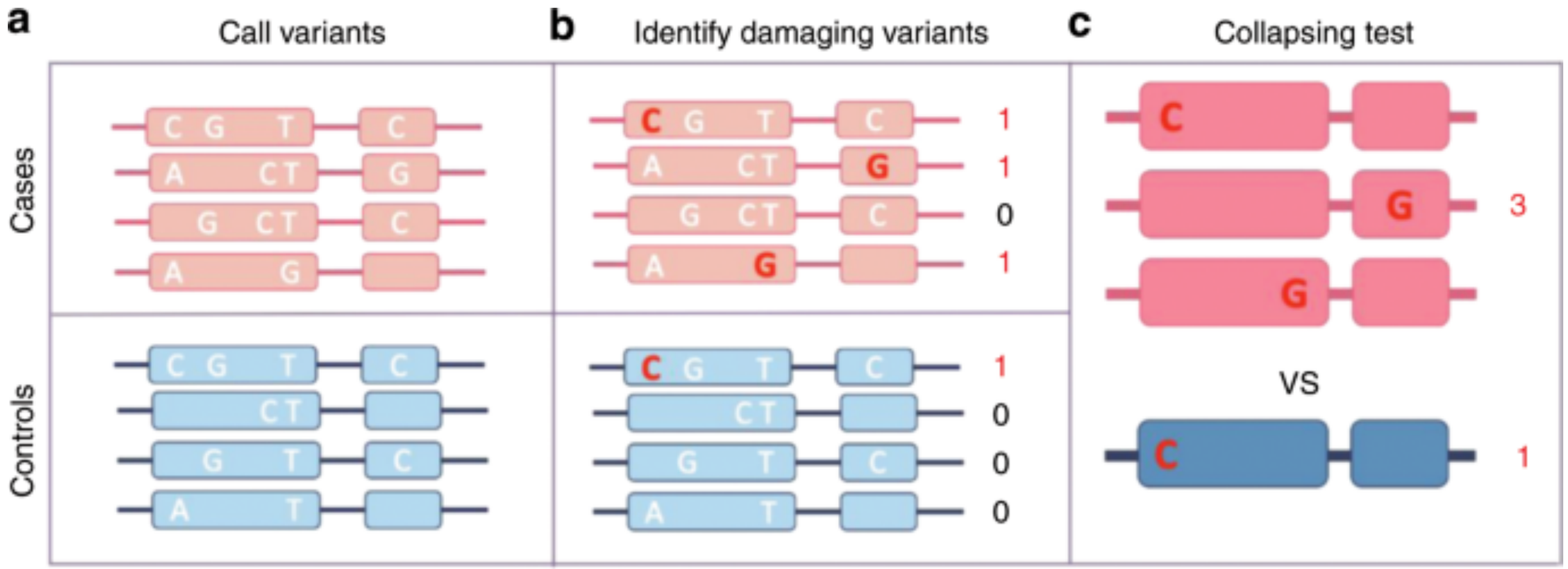
Uncovering the roles of rare variants in common disease through whole-genome sequencing

Elizabeth T. Cirulli and David B. Goldstein

GWAS: individual common variant associations

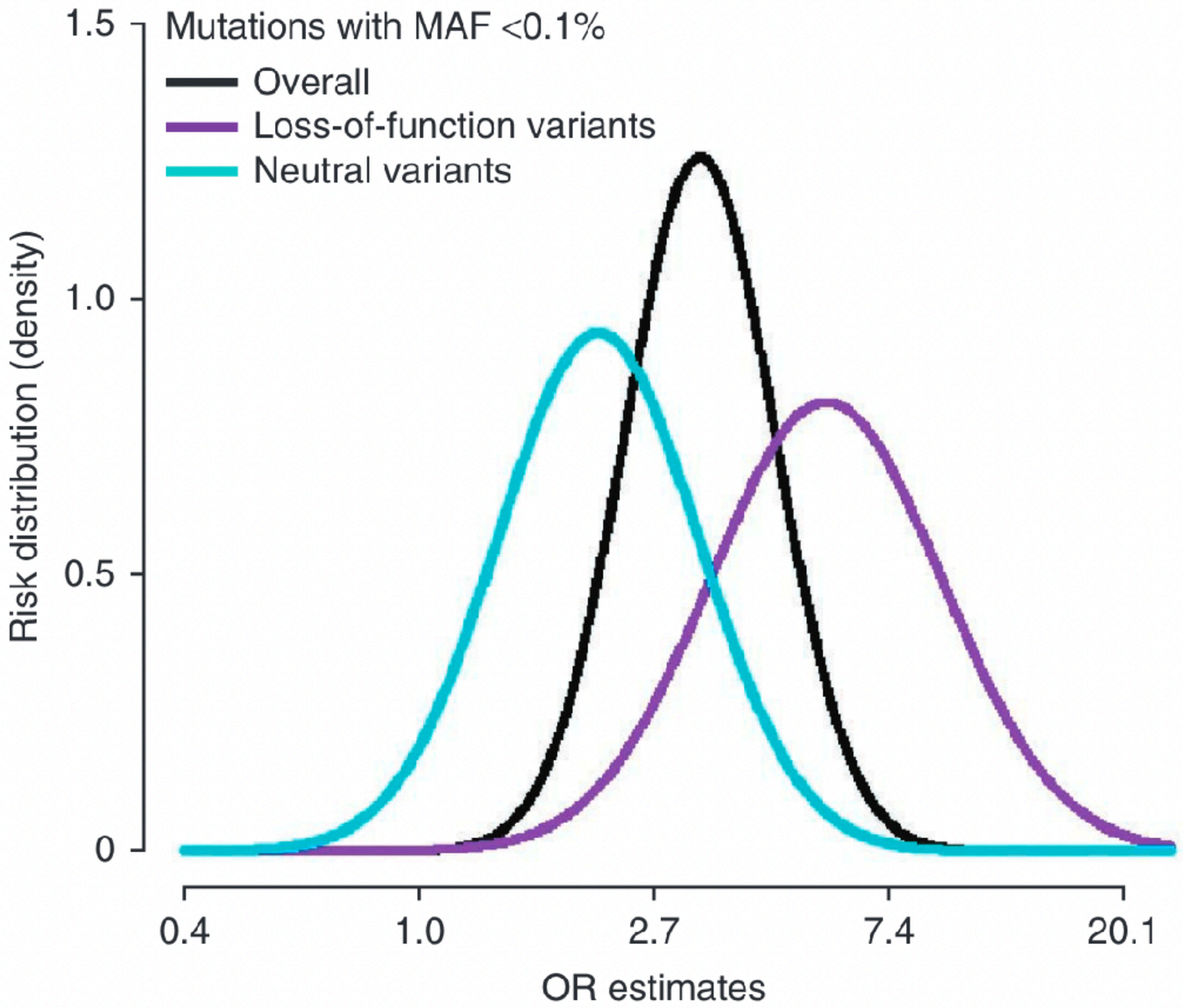


Rare variants: aggregate gene-level associations



Cirulli et al, Nature Communications, 2020

Early successes from targeted sequencing

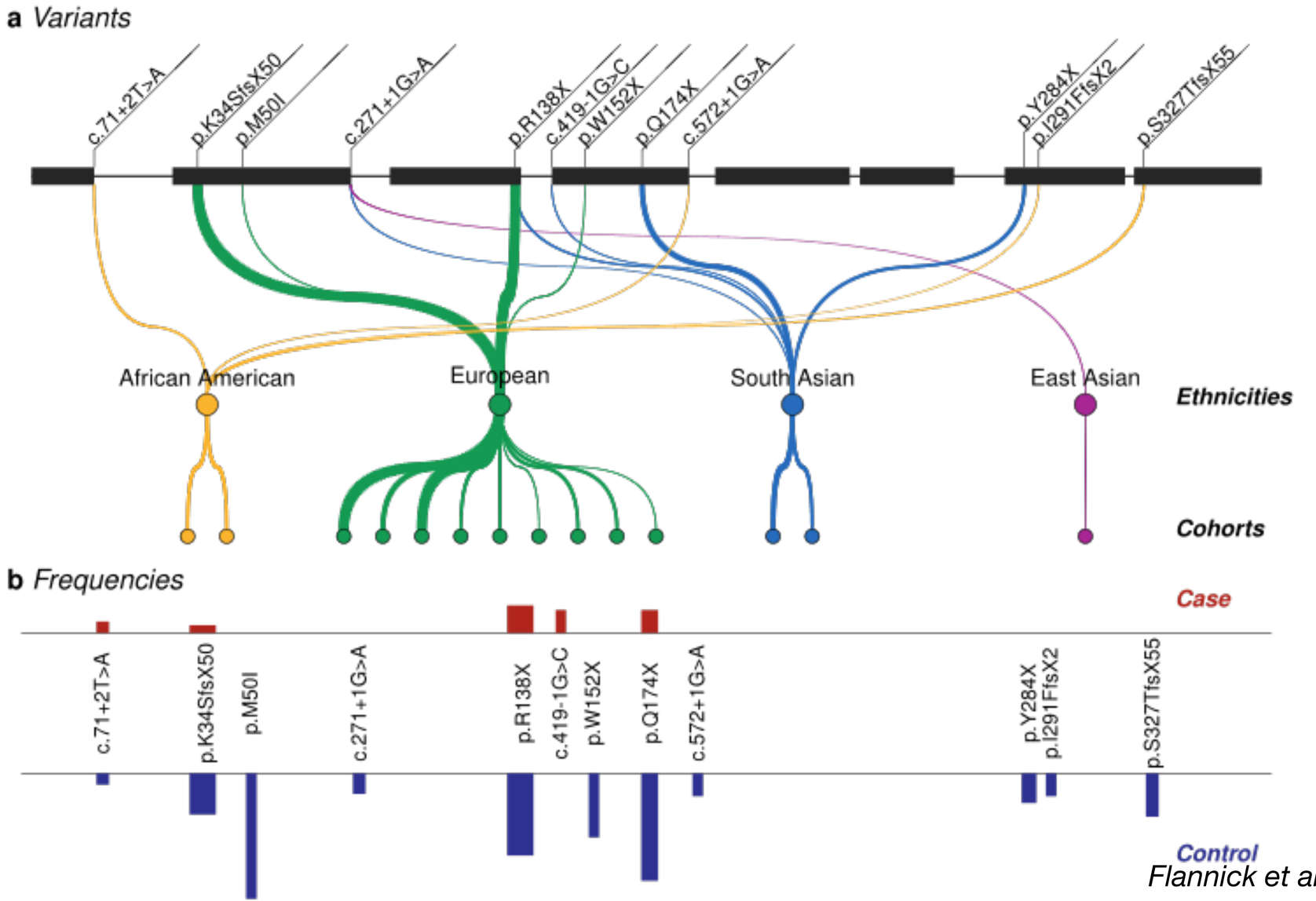


Bonnefond et al. Nature Genetics, 2012

Early successes from targeted sequencing

12 loss-of-function *SLC30A8* mutations in **149,134** individuals

Aggregate odds ratio: **0.34** p=**1.7 x 10^{-6}**

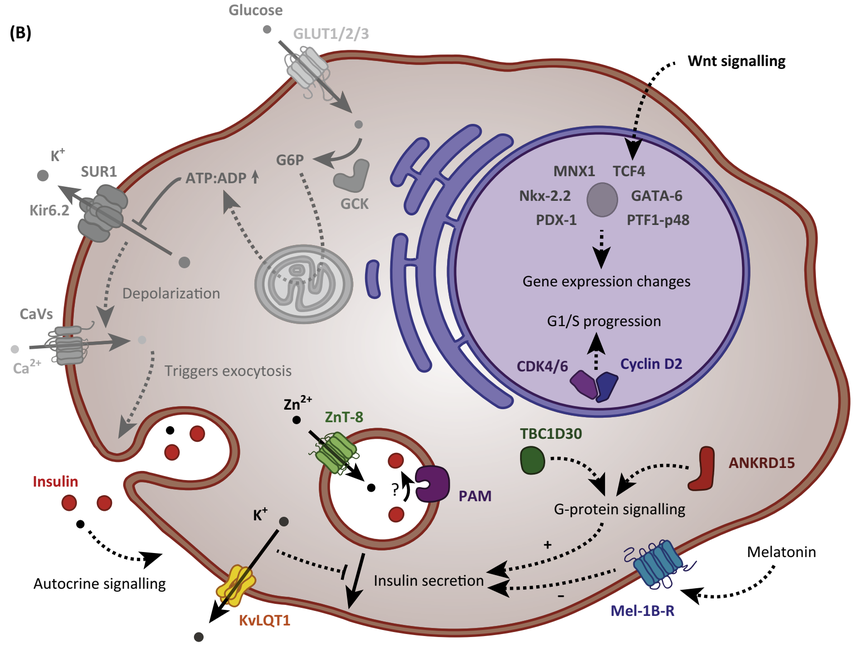


Flannick et al. Nature Genetics, 2014

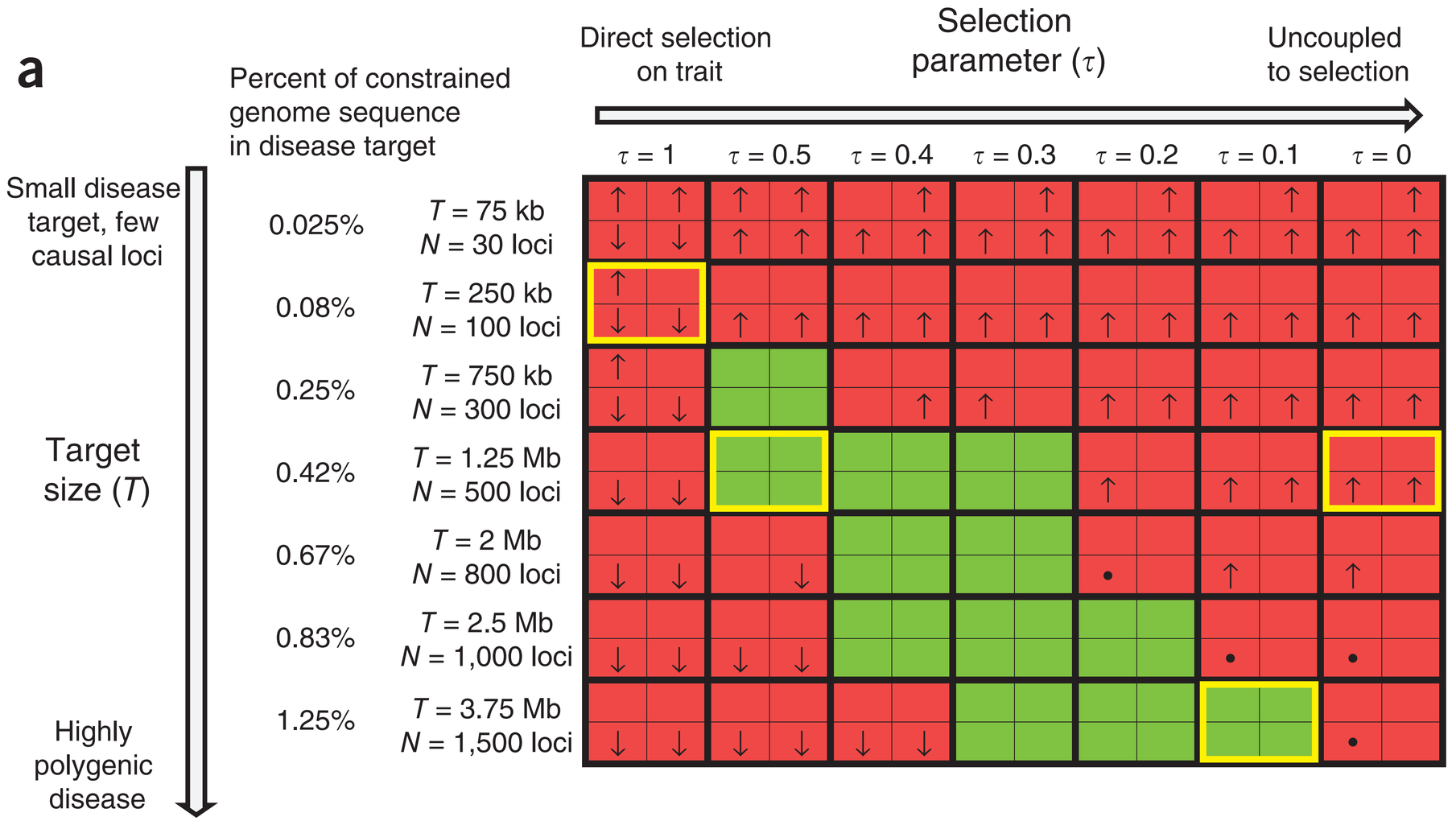
Biological insights

vs.

Genetic architecture



Are rare variants responsible for a significant fraction of disease heritability?



Agarwala, Flannick et al, Nature Genetics 2013

Coding variation

13,000
samples from
diverse ancestries

Noncoding variation

2,700
samples of
European ancestry

Follow-up

82,000 samples

Exome array genotyping

WGS imputation

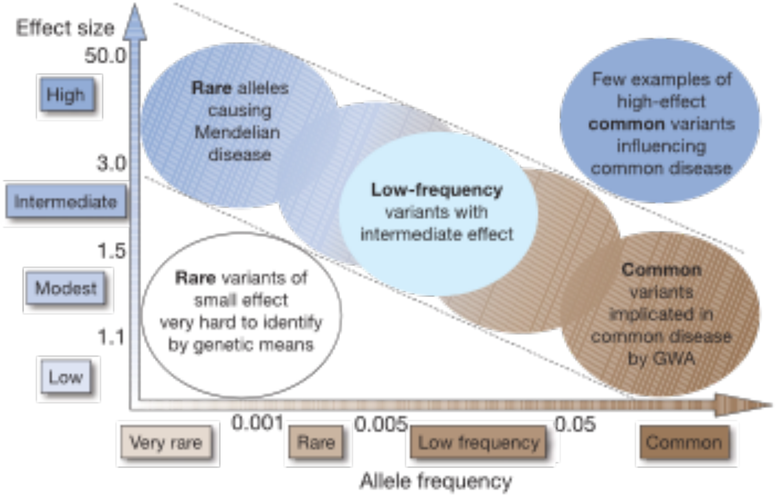
Main finding:

a polygenic, common variant model for T2D

Fuchsberger, Flannick, Teslovich, Mahajan, Gaulton, Agarwala et al, Nature 2016

The case of the missing heritability

When scientists opened up the human genome, they expected to find the genetic components of common traits and diseases. But they were nowhere to be seen. **Brendan Maher** shines a light on six places where the missing loot could be stashed away.



Are Rare Variants Responsible for Susceptibility to Complex Diseases?

Jonathan K. Pritchard

Department of Statistics, University of Oxford, Oxford

Common Genetic Variation and Human Traits

David B. Goldstein, Ph.D.

Genetic Heterogeneity in Human Disease

Jon McClellan[1,*] and Mary-Claire King[2,*]
[1]Department of Psychiatry
[2]Departments of Medicine and Genome Sciences
University of Washington, Seattle, WA 98195-7720, USA
*Correspondence: drjack@uw.edu (J.M.), mcking@uw.edu (M.-C.K.)
DOI 10.1016/j.cell.2010.03.032

Clan Genomics and the Complex Architecture of Human Disease

James R. Lupski,[1,2,3,*] John W. Belmont,[1,2] Eric Boerwinkle,[4,5] and Richard A. Gibbs[1,5,*]

OPEN ACCESS Freely available online

PLoS BIOLOGY

Rare Variants Create Synthetic Genome-Wide Associations

Samuel P. Dickson[1,2], Kai Wang[3], Ian Krantz[3,4,5], Hakon Hakonarson[3,4,5], David B. Goldstein[1,*]

Common and rare variants in multifactorial susceptibility to common diseases

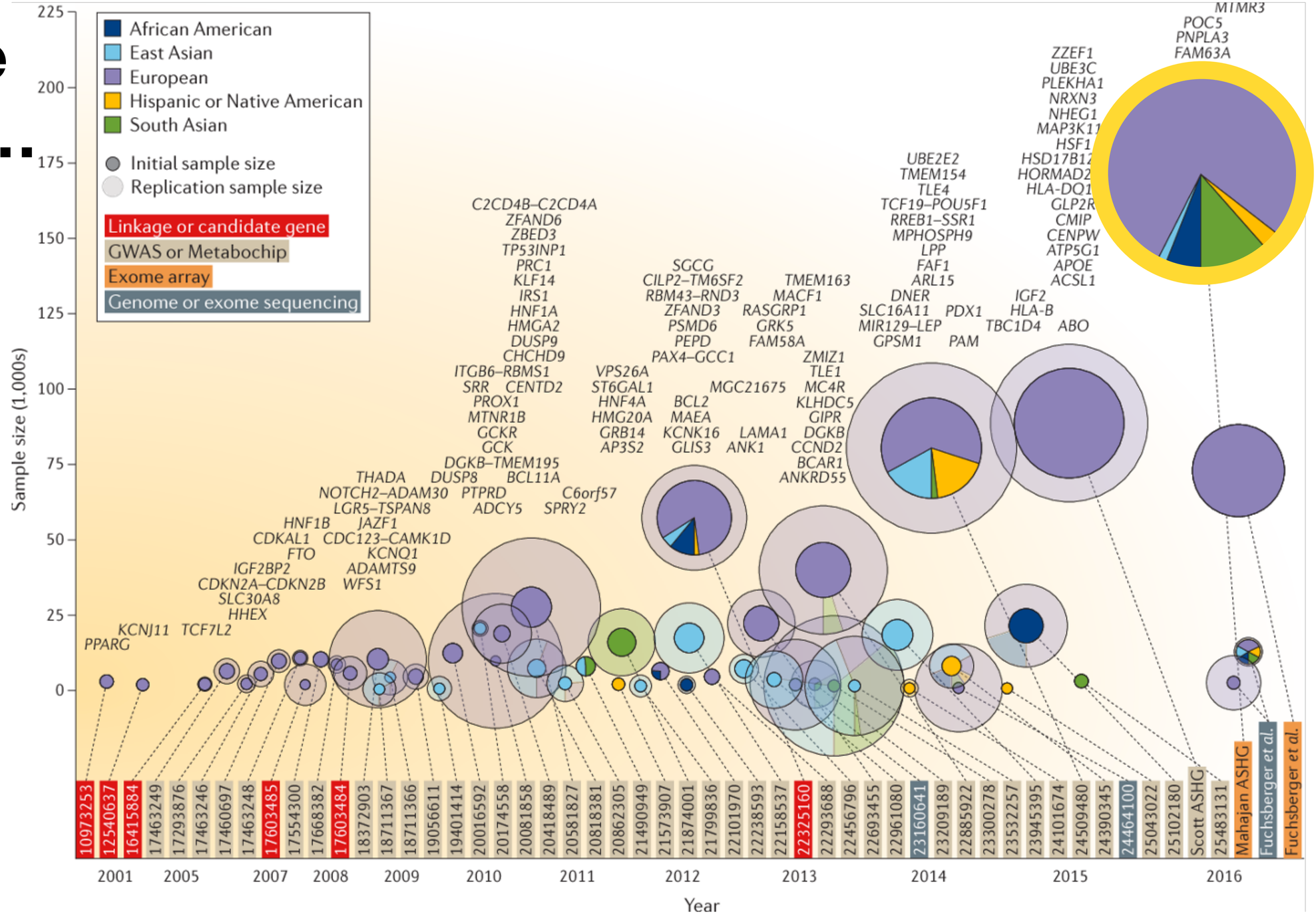
Walter Bodmer & Carolina Bonilla

APPLICATIONS OF NEXT-GENERATION SEQUENCING

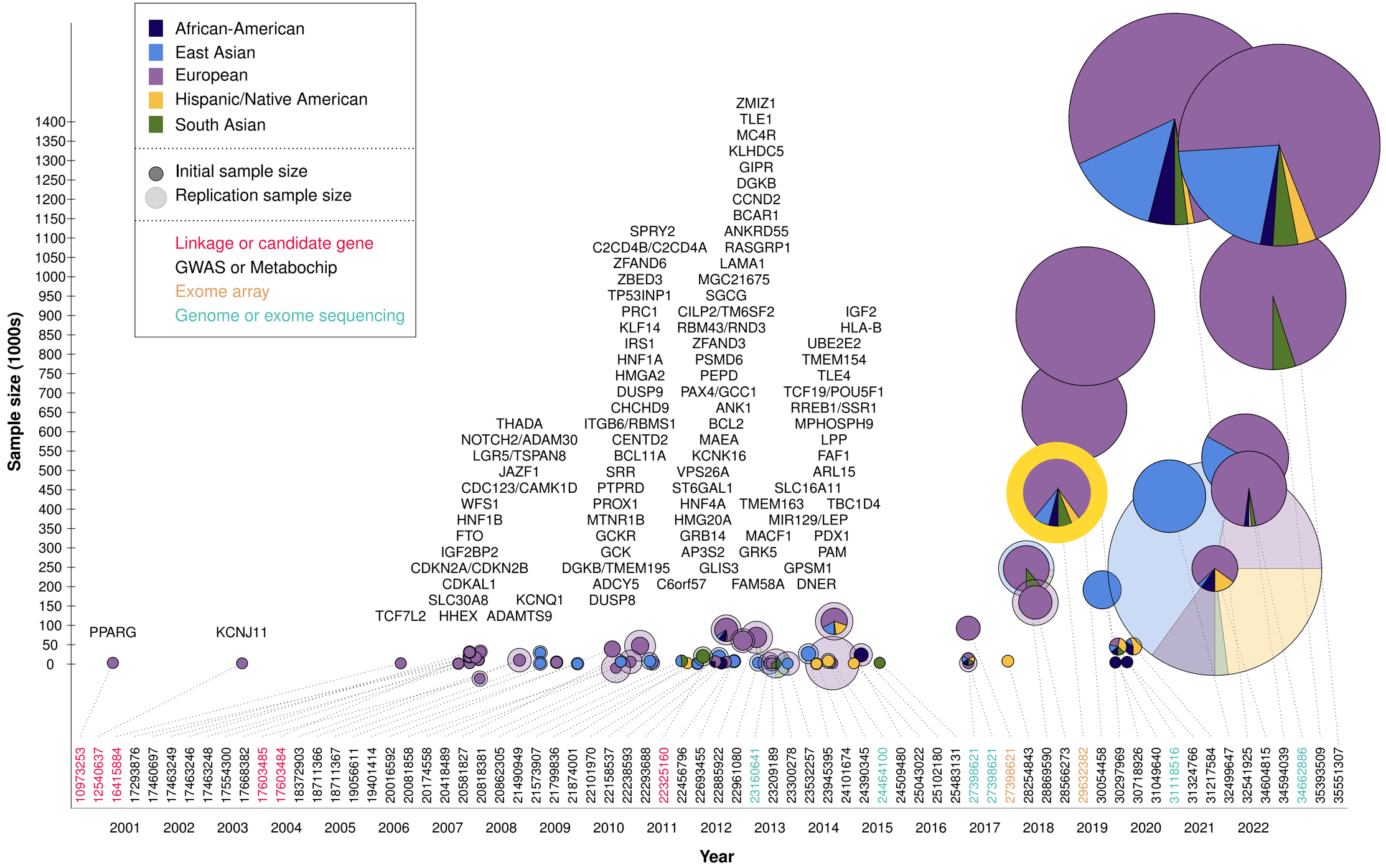
Uncovering the roles of rare variants in common disease through whole-genome sequencing

Elizabeth T. Cirulli and David B. Goldstein

Since then…

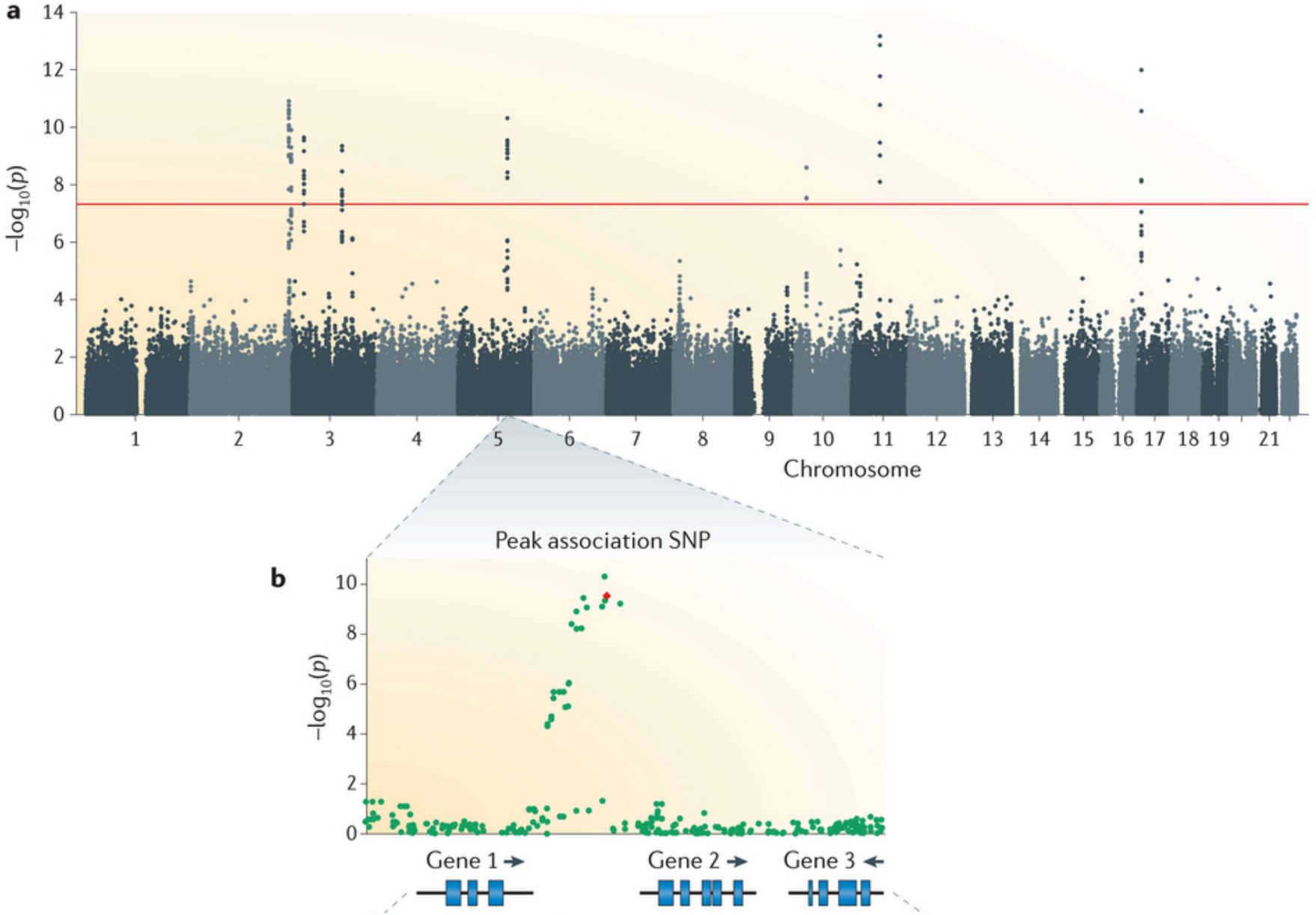


Flannick and Florez, Nature Reviews Genetics, 2016

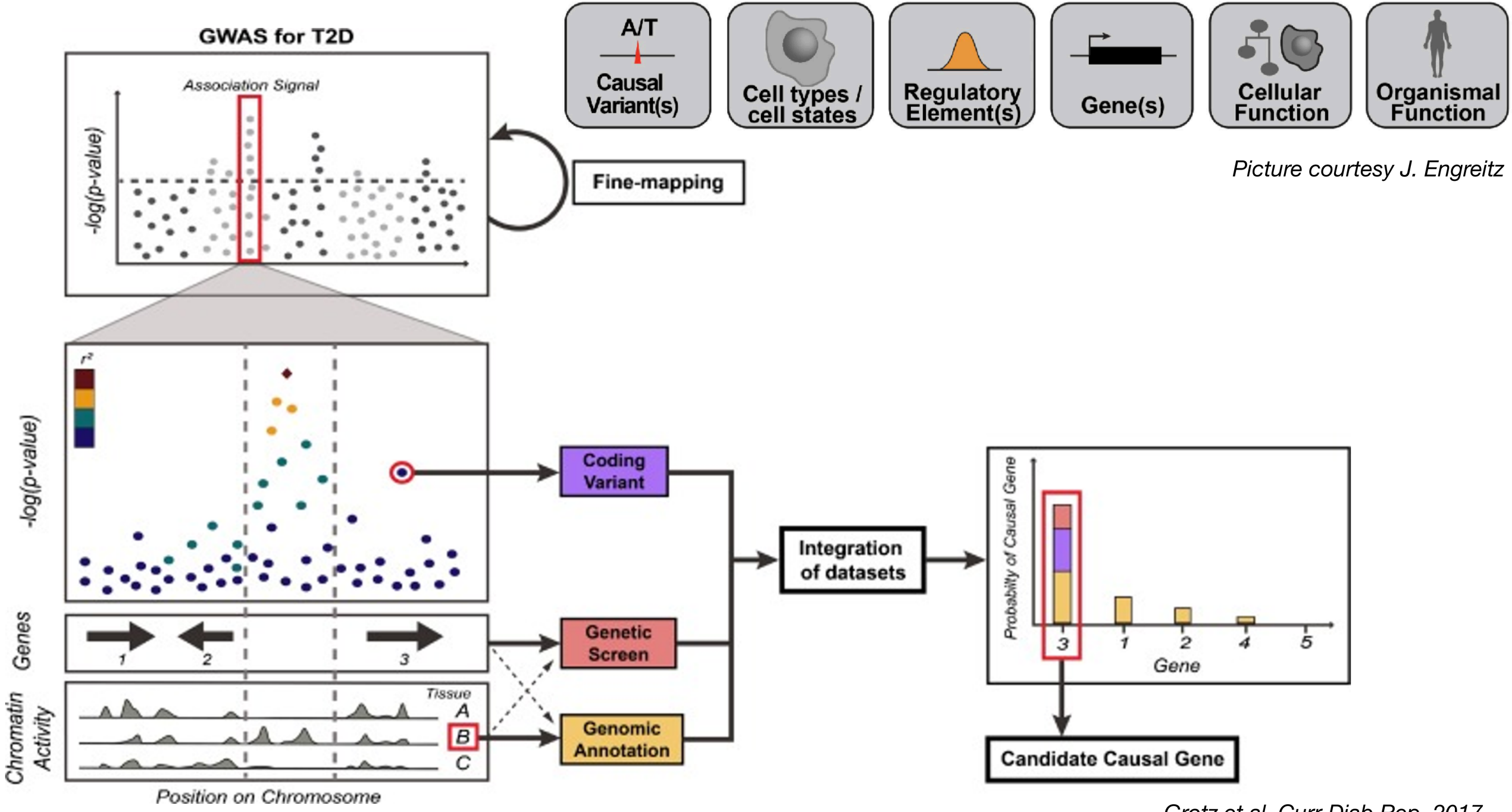

African-American
East Asian
European
Hispanic/Native American
South Asian
Initial sample size
Replication sample size
Linkage or candidate gene
GWAS or Metabochip
Exome array
Genome or exome sequencing
Sample size (1000s)
Year
PPARG
KCNJ11
TCF7L2
HHEX
ADAMTS9
SLC30A8
KCNQ1
DUSP8
CDKAL1
ADCY5
C6orf57
FAM58A
DNER
CDKN2A/CDKN2B
DGKB/TMEM195
GLIS3
GPSM1
IGF2BP2
GCK
AP3S2
GRK5
PAM
FTO
GCKR
GRB14
MACF1
PDX1
HNF1B
MTNR1B
HMG20A
MIR129/LEP
WFS1
PROX1
HNF4A
TMEM163
TBC1D4
CDC123/CAMK1D
PTPRD
ST6GAL1
SLC16A11
JAZF1
SRR
VPS26A
ARL15
LGR5/TSPAN8
BCL11A
KCNK16
FAF1
NOTCH2/ADAM30
CENTD2
MAEA
LPP
THADA
ITGB6/RBMS1
BCL2
MPHOSPH9
CHCHD9
ANK1
RREB1/SSR1
DUSP9
PAX4/GCC1
TCF19/POU5F1
HMGA2
PEPD
TLE4
HNF1A
PSMD6
TMEM154
IRS1
ZFAND3
UBE2E2
KLF14
RBM43/RND3
HLA-B
PRC1
CILP2/TM6SF2
IGF2
TP53INP1
SGCG
ZBED3
MGC21675
ZFAND6
LAMA1
C2CD4B/C2CD4A
RASGRP1
SPRY2
ANKRD55
BCAR1
CCND2
DGKB
GIPR
KLHDC5
MC4R
TLE1
ZMIZ1

Problem: what are the genes?

- Usually only one (or a few) variants are causal



Civelek and Lusis, 2014

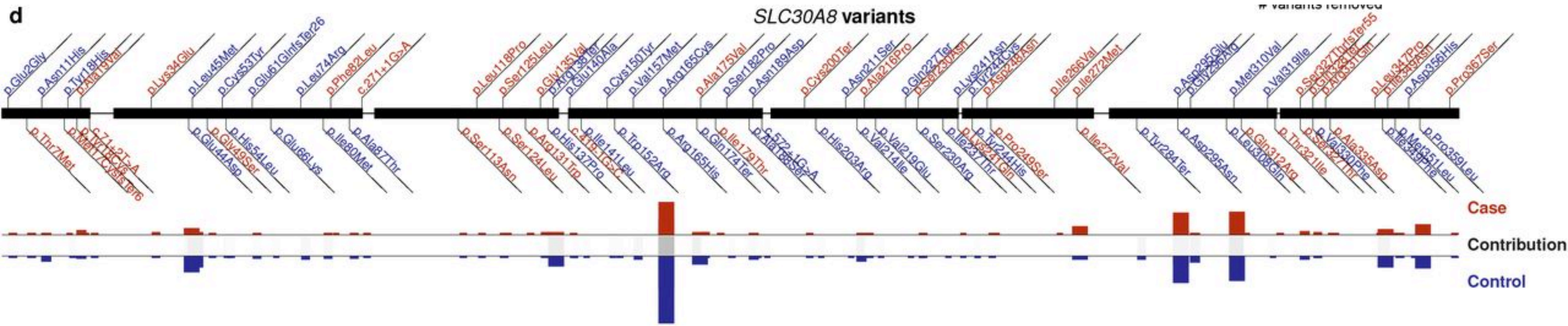


Picture courtesy J. Engreitz

Grotz et al, Curr Diab Rep, 2017

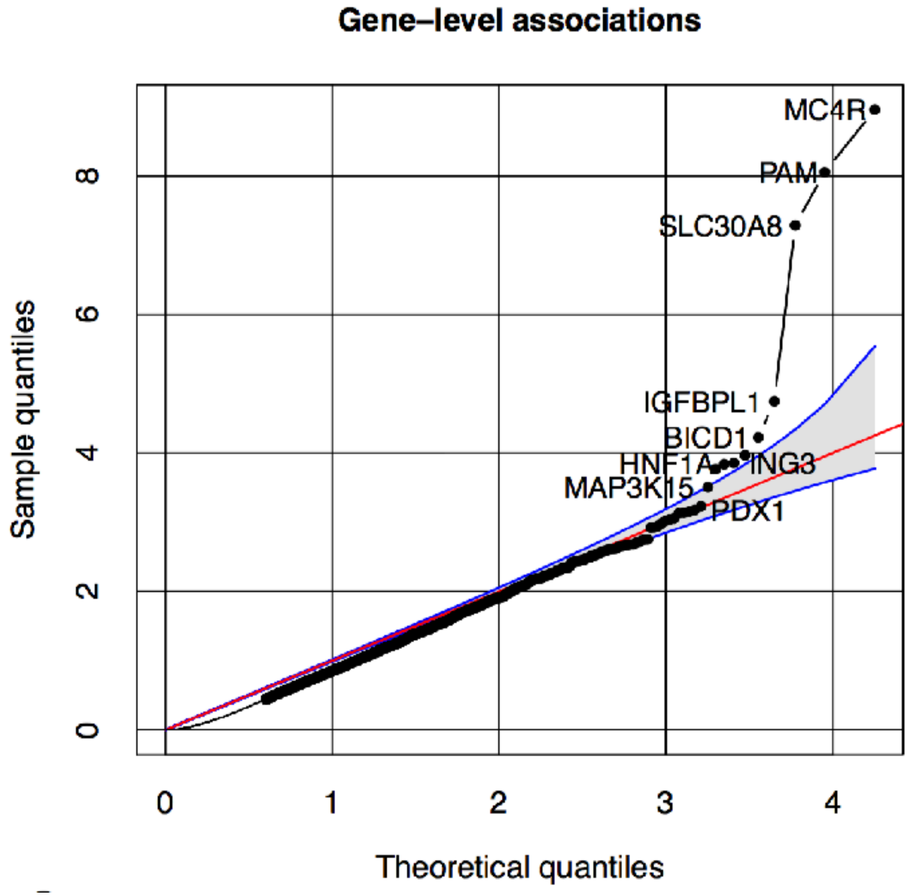
Meanwhile, rare variants…

- When significant: **direct links** from disease to genes
- When variants inactivate protein: **directional link** with disease risk



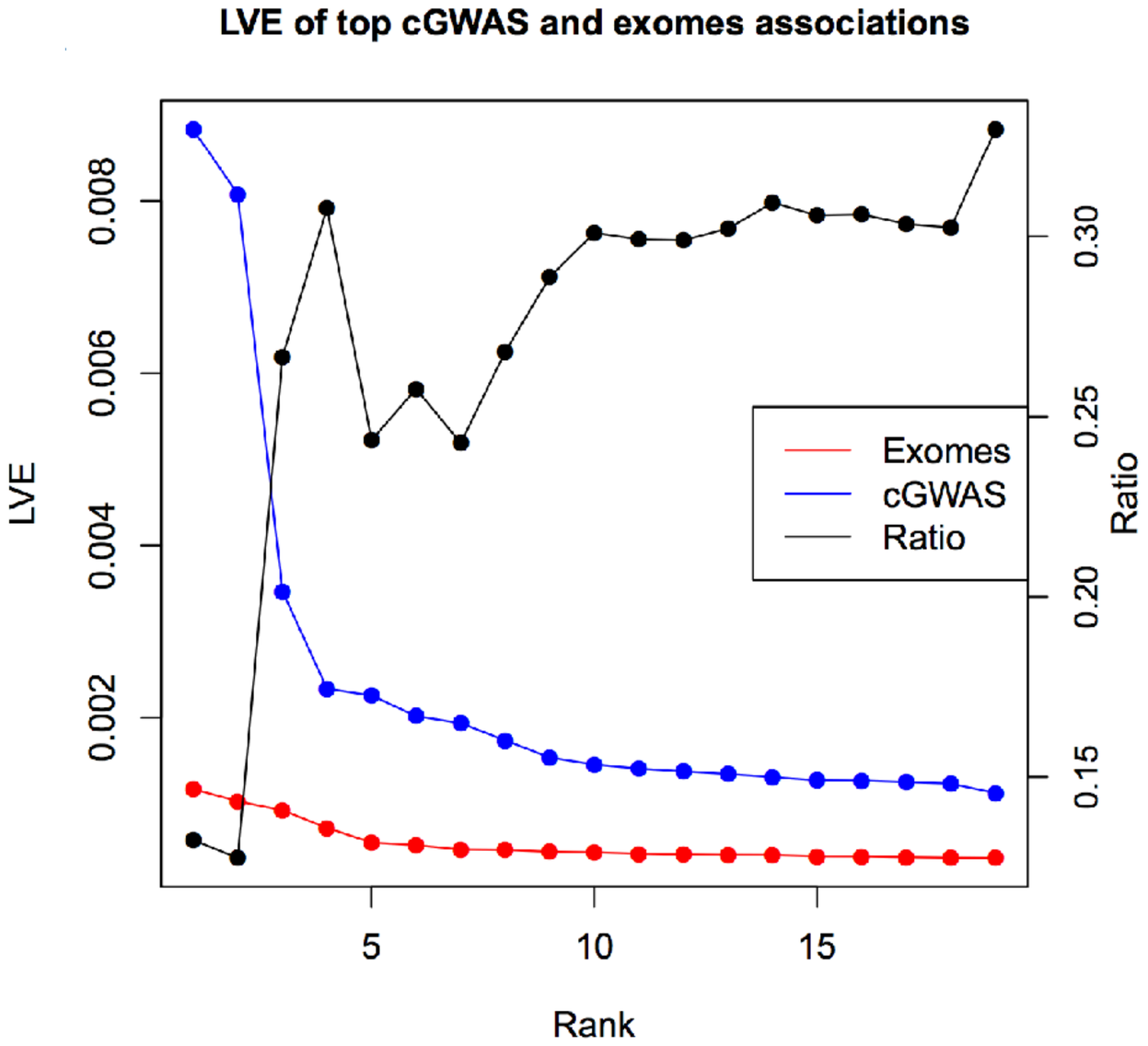
More recent (45,000 sample) exome sequencing

- Three exome-wide significant gene-level associations

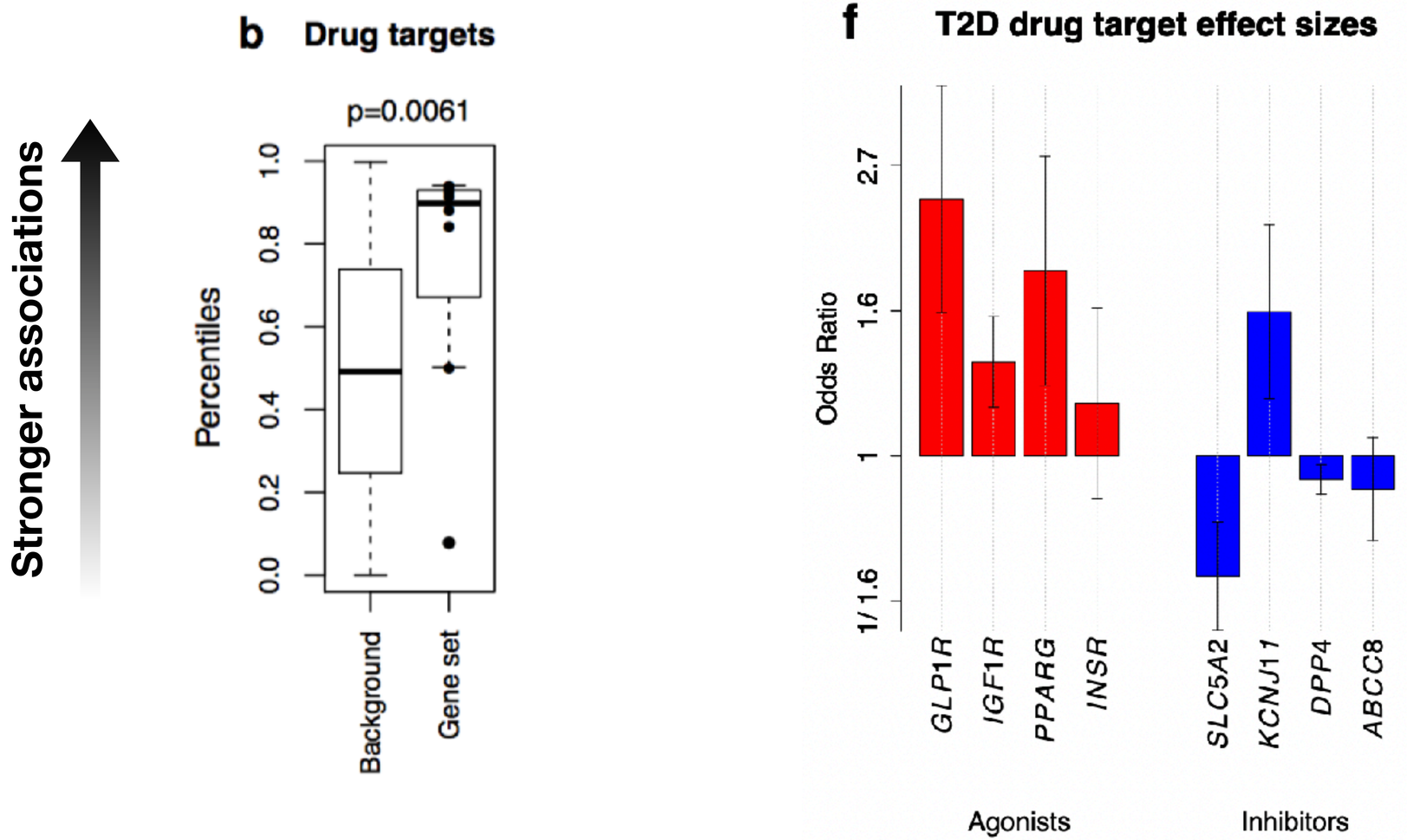


Flannick et al, Nature 2019

Explaining minimal heritability

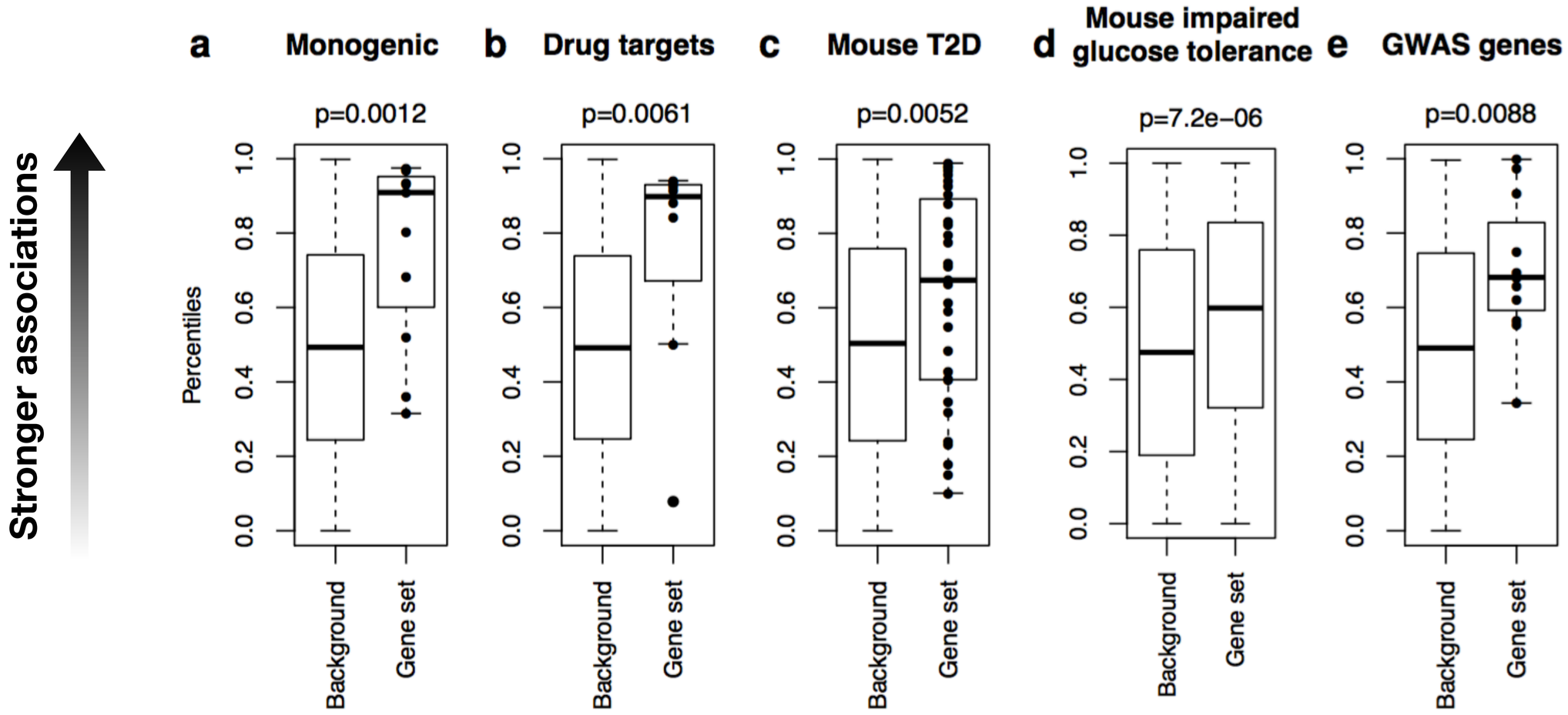


But: many signals beyond



Flannick et al, Nature 2019

But: many signals beyond



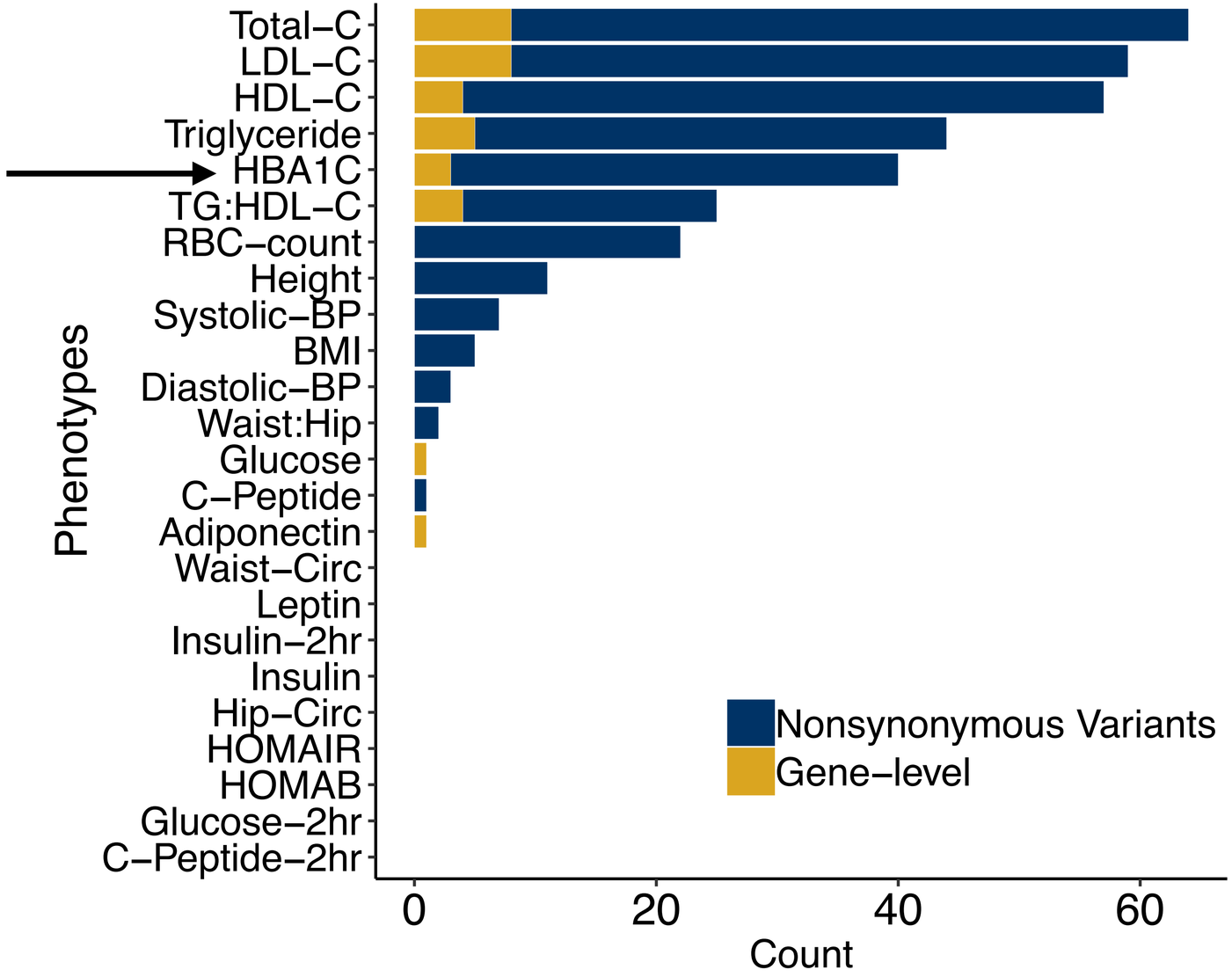
Flannick et al, Nature 2019

Further support for this model

Beyond type 2 diabetes

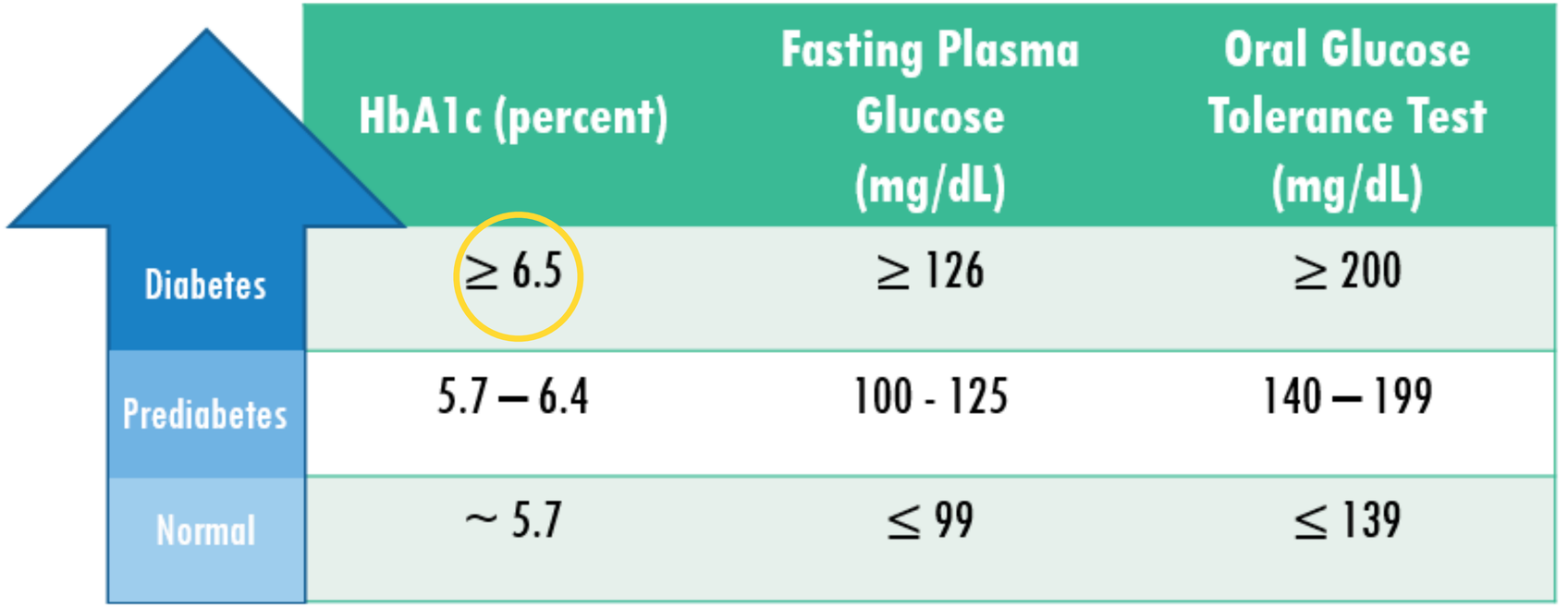
- **Study design:**
  - Exome sequencing of 85,474 non-diabetic individuals
    - UKBB exomes (n=40,151; 100% European)
    - AMP-T2D exomes (n=45,323; 15.8% African American; 25.6% East Asian; 18.7% European; 18.0% Hispanic; and 22.2% South Asian)
  - 24 quantitative traits
  - Single variant analysis (mostly for common variants)
  - Gene-level analysis (for rare variants)

Dornbos et al, Nature Genetics 2022



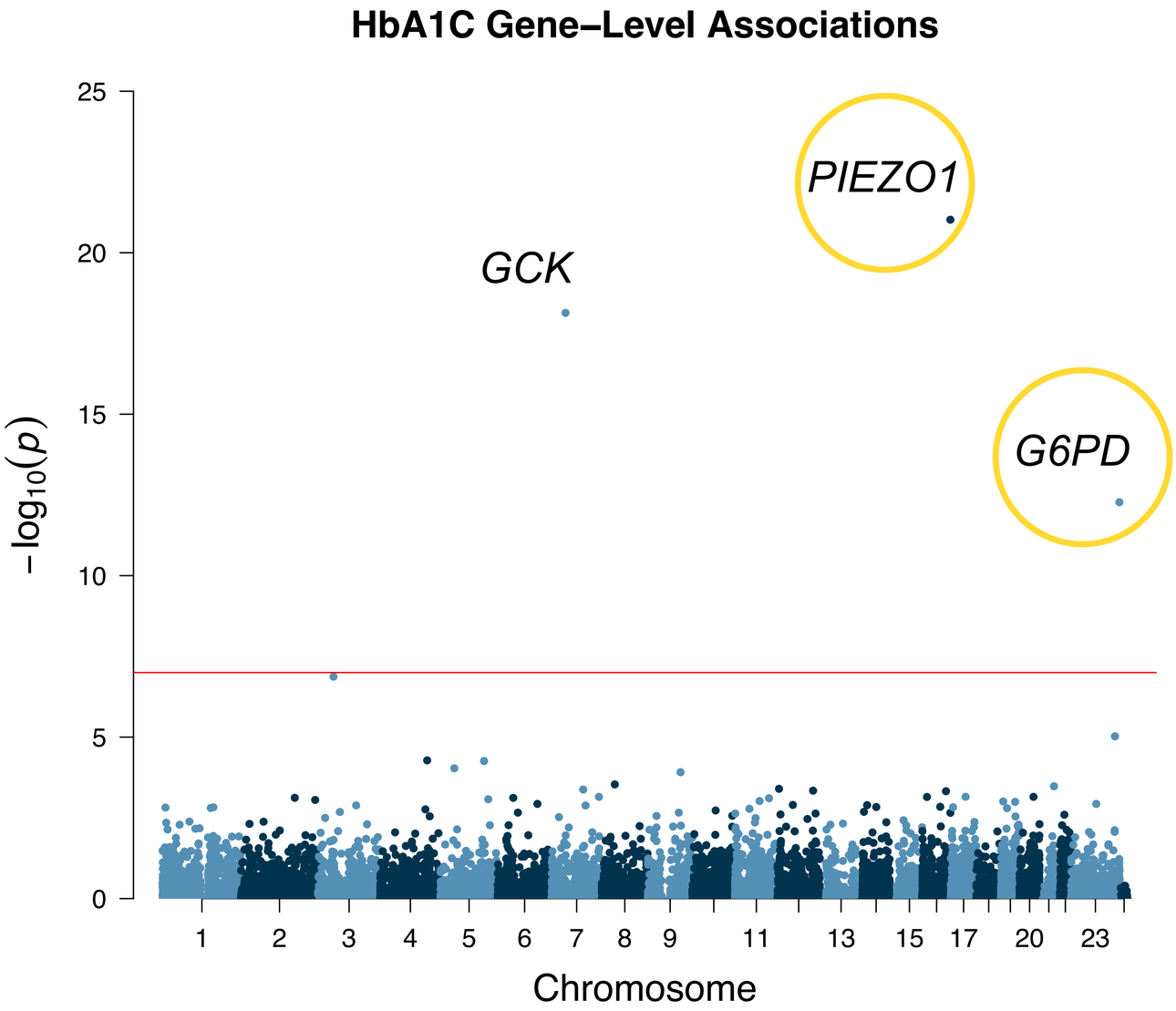
Dornbos et al, Nature Genetics 2022

HbA1C is commonly used in T2D diagnosis



	HbA1c (percent)	Fasting Plasma Glucose (mg/dL)	Oral Glucose Tolerance Test (mg/dL)
Diabetes	≥ 6.5	≥ 126	≥ 200
Prediabetes	5.7 – 6.4	100 - 125	140 – 199
Normal	~ 5.7	≤ 99	≤ 139

Dornbos et al, Nature Genetics 2022

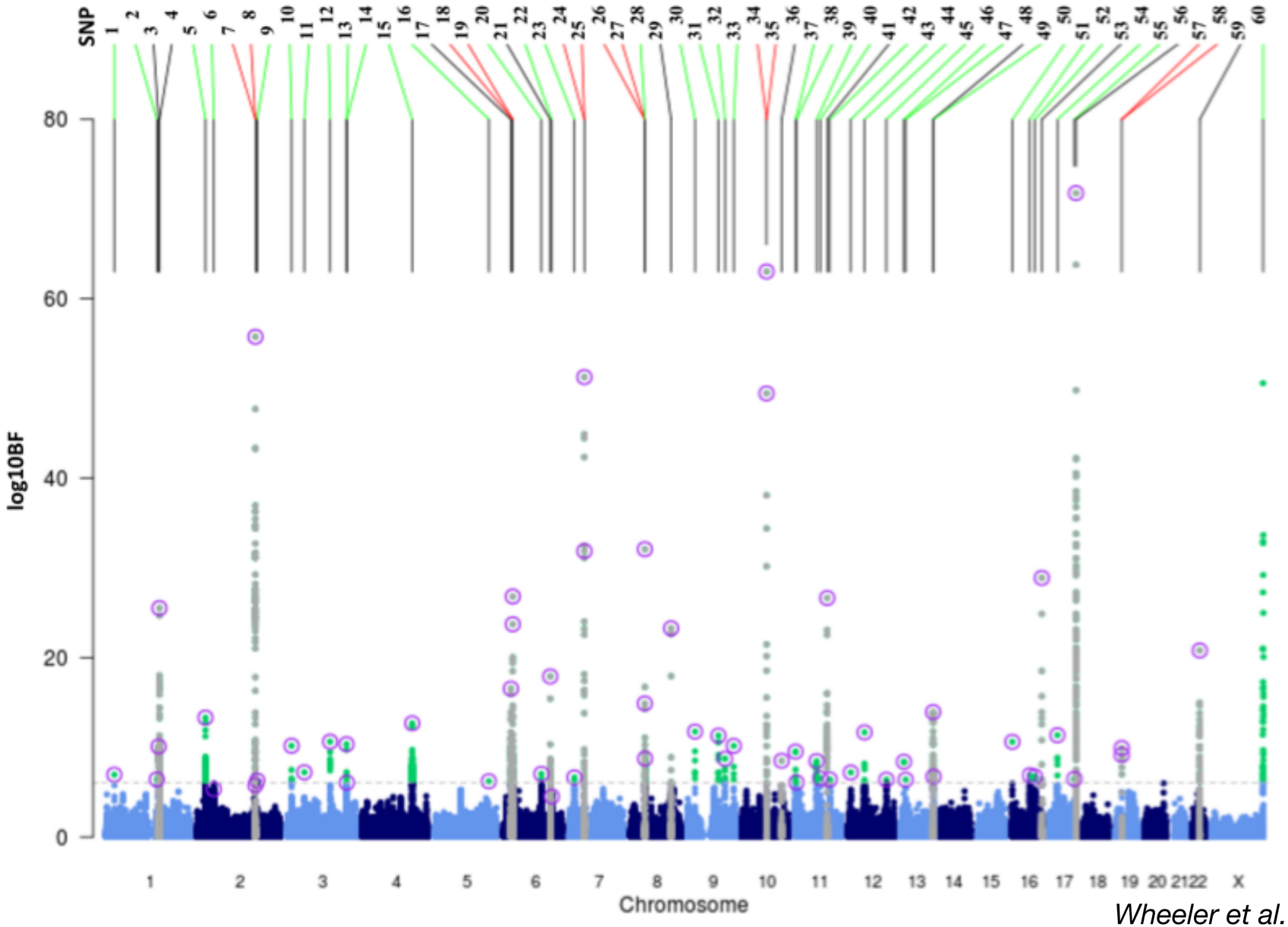


Dornbos et al, Nature Genetics 2022

Common variants and HbA1C

~150,000 individuals

60 common variant associations



G6PD variant 11% of African-Americans

0.81% **reduction** in HbA1c

Wheeler et al. Nature Genetics, 2017

Common variants and HbA1c

- Associations can be grouped into two classes

Associated with red blood cell traits

Associated with other glycemic traits

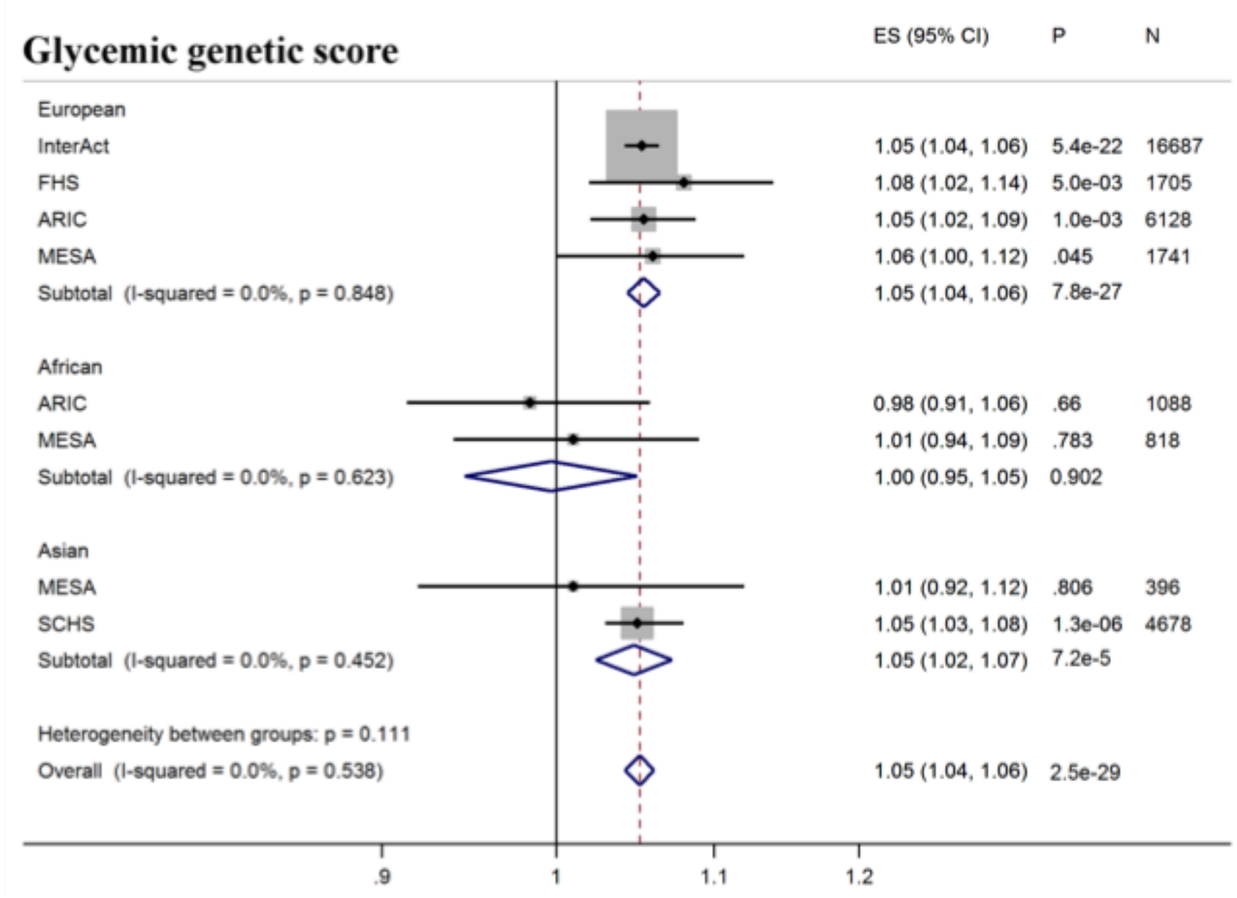
Table 1. (*Continued*)

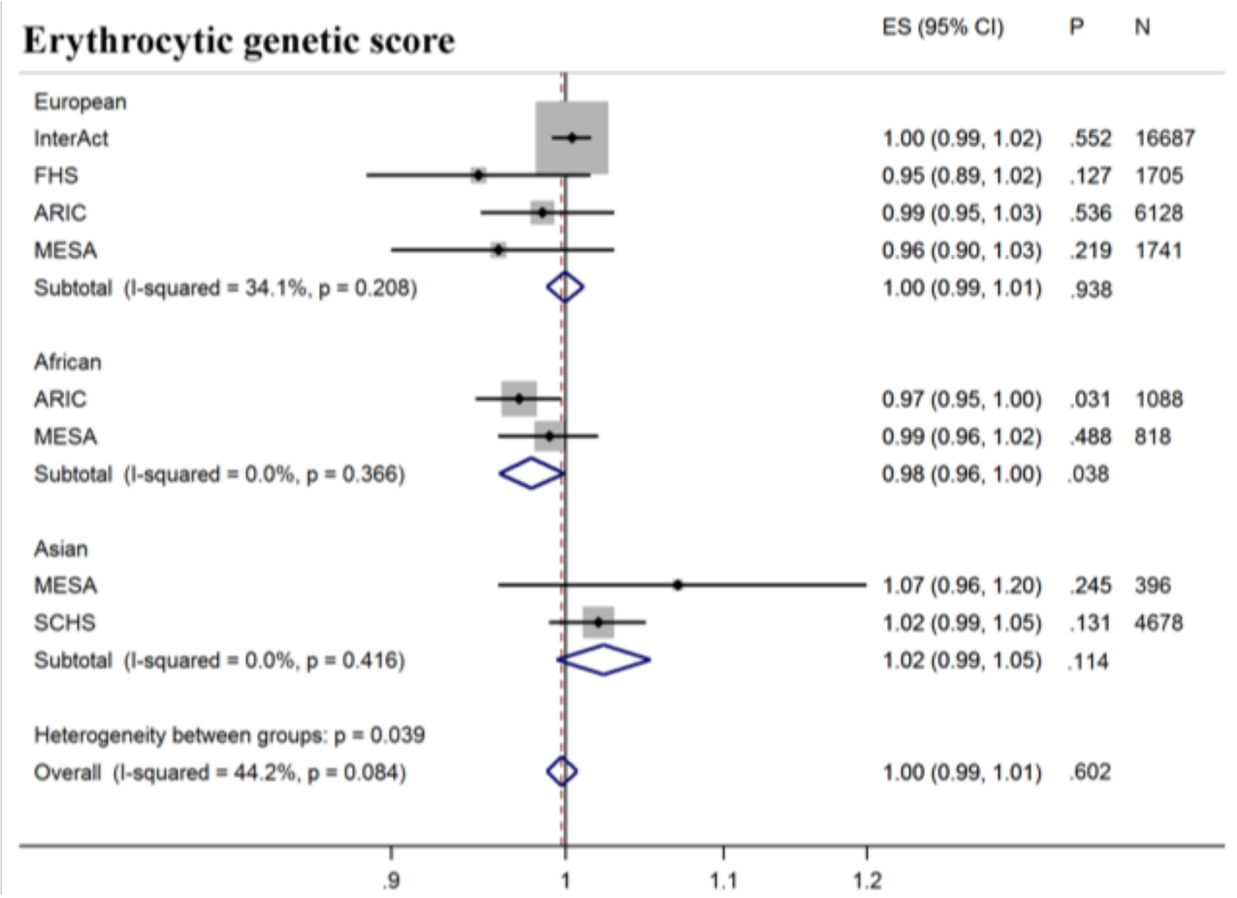
SNP	Markername	Chr.	Position (bp)	Effect Allele	Other Allele	Gene	Status	Signals	Classification	European ancestry METAL *p*-value	Trans-ethnic MANTRA log10BF
45	**rs10774625**	12	110394602	G	A	ATXN2	Novel	Single	**Erythrocytic**	1.46×10^{-8}	6.38
46	**rs11619319**	13	27385599	G	A	PDX1	Novel	Single	**Glycemic**	4.58×10^{-7}	8.38
47	**rs576674**	13	32452302	G	A	KL	Novel	Single	**Glycemic**	1.39×10^{-5}	6.38
48	**rs282587**	13	112399663	G	A	ATP11A	Known	Single	Unclassified	1.70×10^{-12}	13.92
49	**rs9604573**	13	113571085	T	C	GAS6	Novel	Single	Unclassified	9.60×10^{-9}	6.72
50	**rs11248914**	16	233563	T	C	ITFG3	Novel	Single	**Erythrocytic**	2.56×10^{-14}	10.60
51	**rs1558902**	16	52361075	A	T	FTO	Novel	Single	Unclassified	3.27×10^{-8}	6.88
52	**rs4783565**	16	67307691	A	G	CDH3	Novel	Single	**Erythrocytic**	1.73×10^{-7}	6.73
53	**rs837763**	16	87381230	T	C	CDT1	Known	Single	**Erythrocytic**	1.68×10^{-28}	28.89
54	**rs9914988**	17	24207230	A	G	ERAL1	Novel	Single	**Erythrocytic**	2.77×10^{-11}	11.34
55	**rs2073285**	17	73628956	C	T	TMC6	Novel	Single	Unclassified	1.27×10^{-4}	6.47
56	**rs1046896**	17	78278822	T	C	FN3KRP	Known	Single	Unclassified	4.46×10^{-64}	71.79
57	**rs11086054**	19	17107737	A	T	MYO9B	Novel	Multiple	Unclassified	8.16×10^{-6}	9.12
58	**rs17533903**	19	17117523	A	G	MYO9B	Known	Multiple	**Erythrocytic**	5.27×10^{-12}	9.912
59	**rs4820268**	22	35799537	G	A	TMPRSS6	Known	Single	**Erythrocytic**	1.40×10^{-22}	20.79
60	**rs1050828**	X	153417411	T	C	G6PD	Novel	Single	**Erythrocytic**	NA*	NA

Wheeler et al. Nature Genetics, 2017

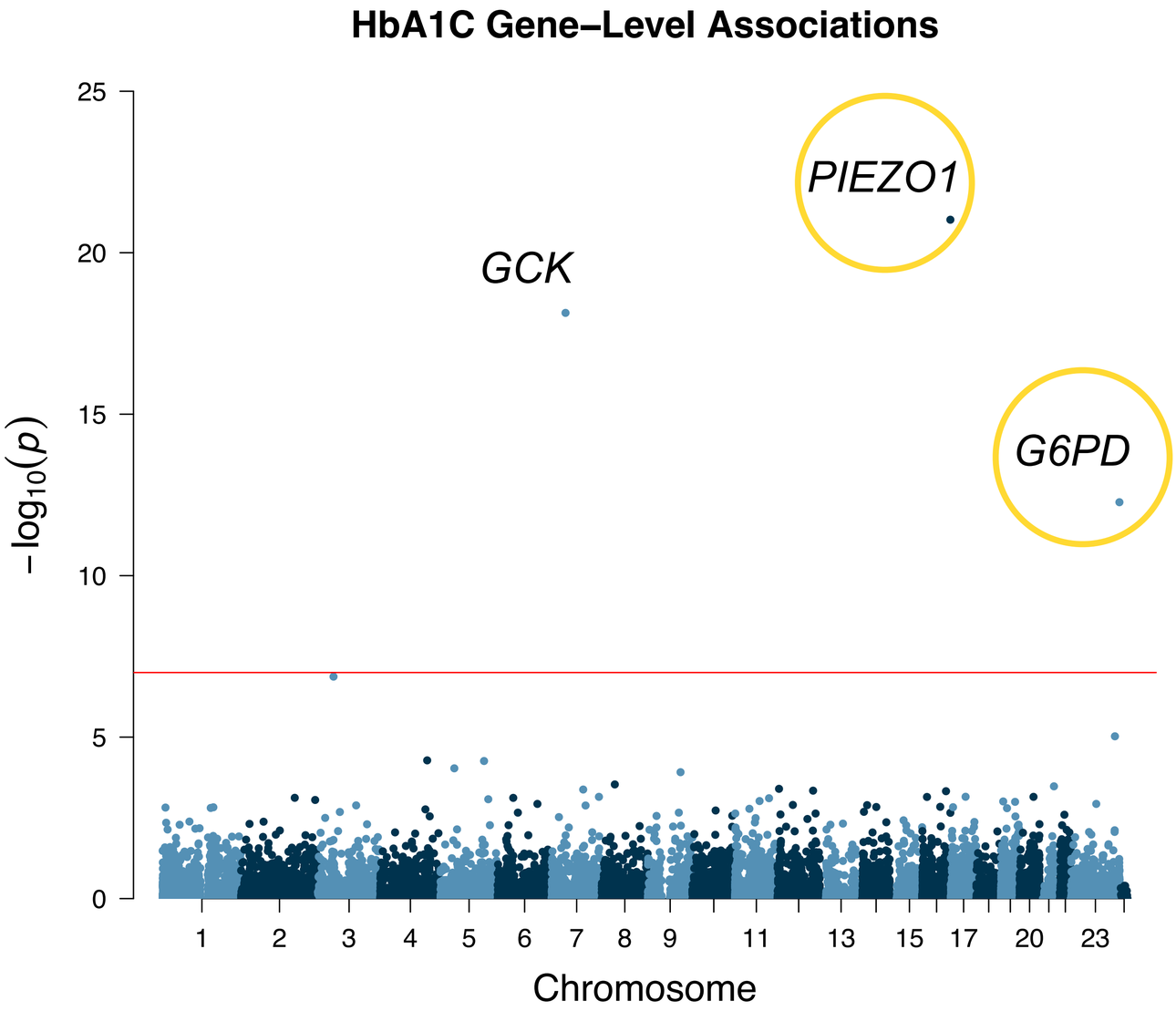
Different biological effects

- Glycemic associations, but not erythrocytic associations, predict future development of T2D
  - ~2% of African-Americans could be misclassified due to *G6PD* variant



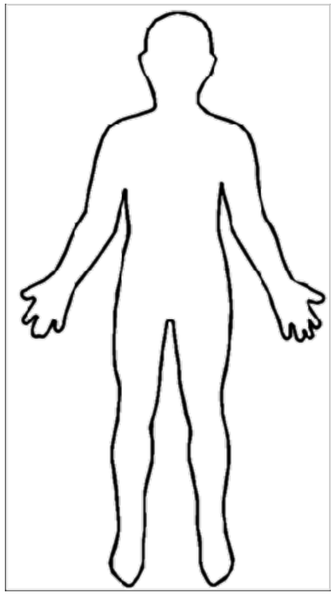


Wheeler et al. Nature Genetics, 2017



Dornbos et al, Nature Genetics 2022

How do these variants affect HbA1c diagnosis?

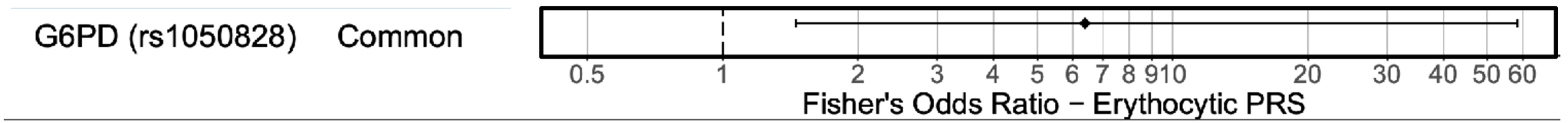


HbA1C 6.4%

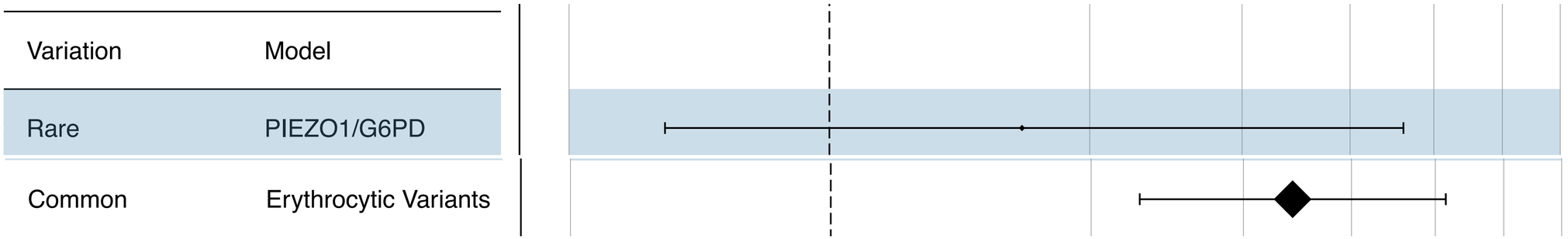
Carries variants with HbA1C effect -0.2%

Evaluate via independent diagnosis

	HbA1c (percent)	Fasting Plasma Glucose (mg/dL)	Oral Glucose Tolerance Test (mg/dL)
Diabetes	≥ 6.5	≥ 126	≥ 200
Prediabetes	5.7 – 6.4	100 - 125	140 – 199
Normal	~ 5.7	≤ 99	≤ 139



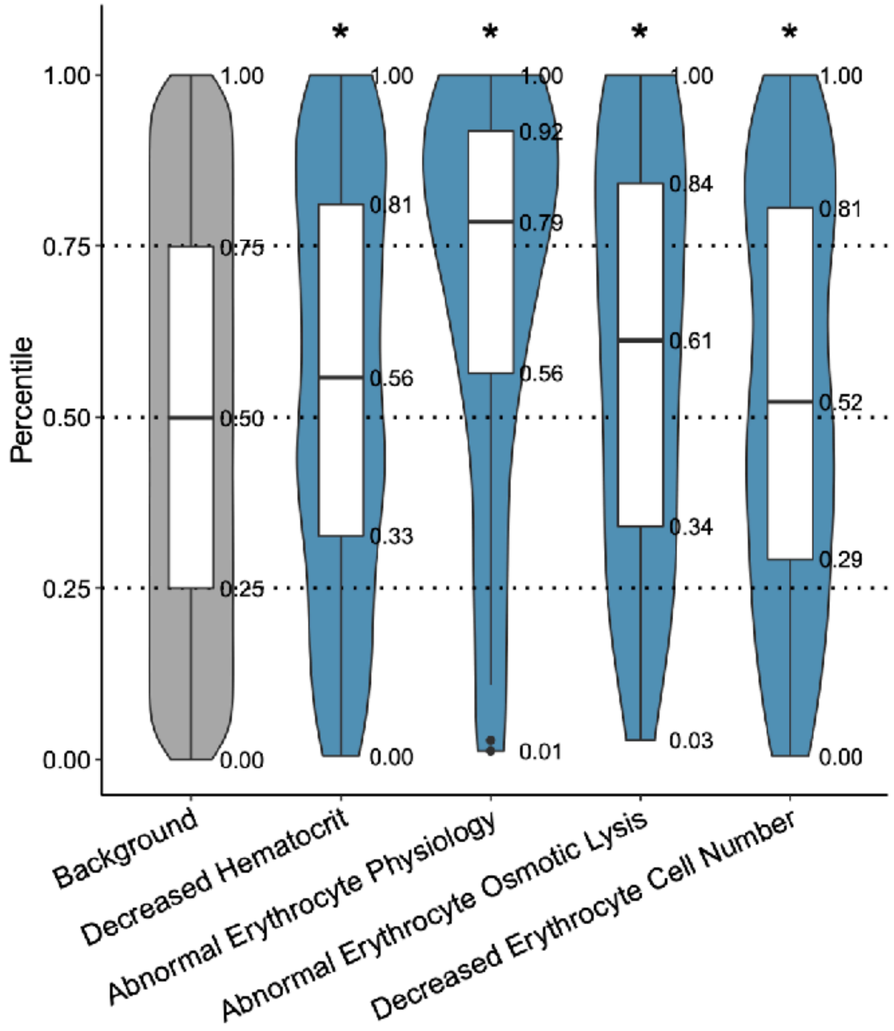
How do these variants affect HbA1c diagnosis?



Dornbos et al, Nature Genetics 2022

Expanding the model

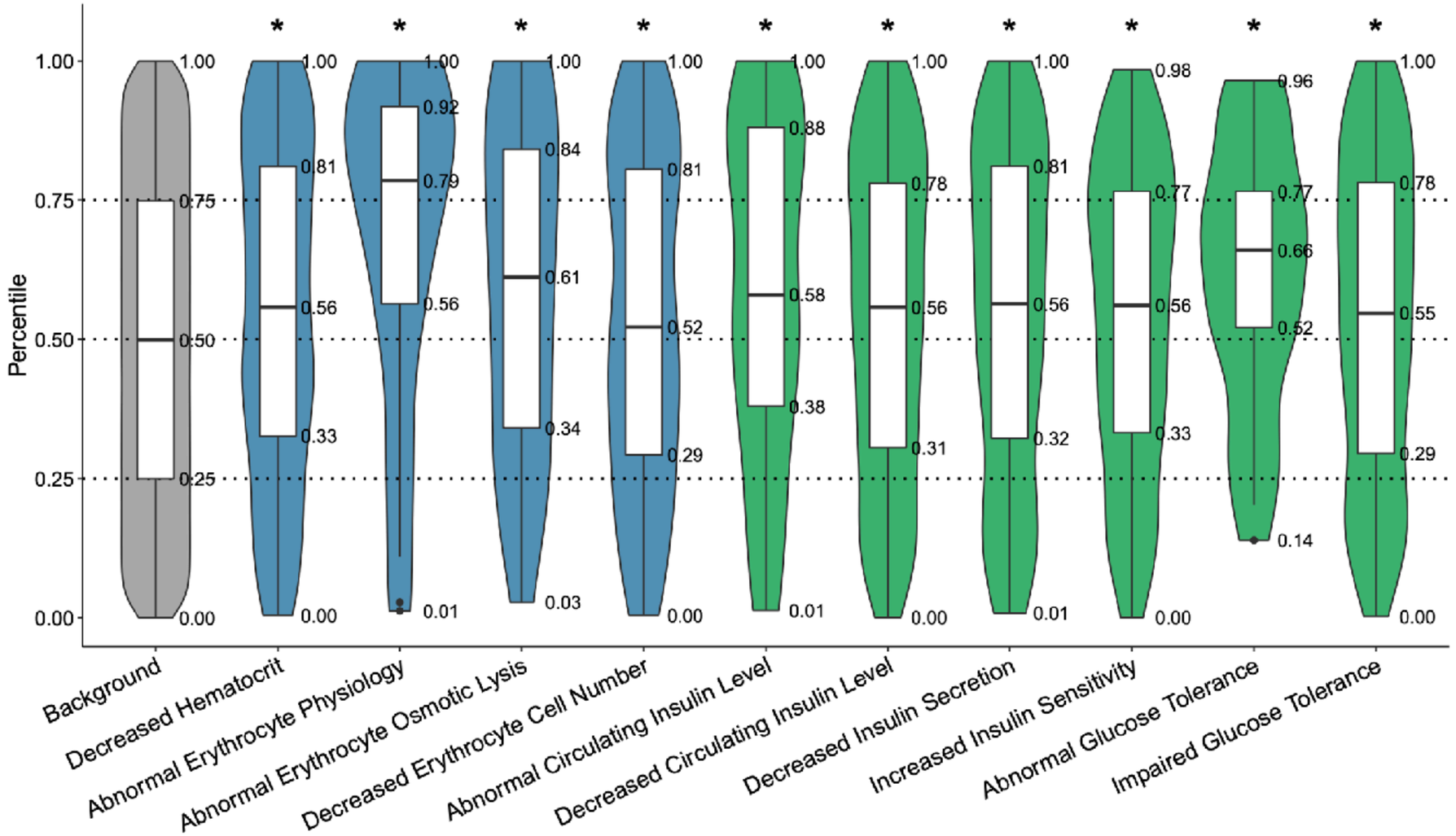
- Significant associations in many sets of genes with known function on erythrocytic lifespan in mice

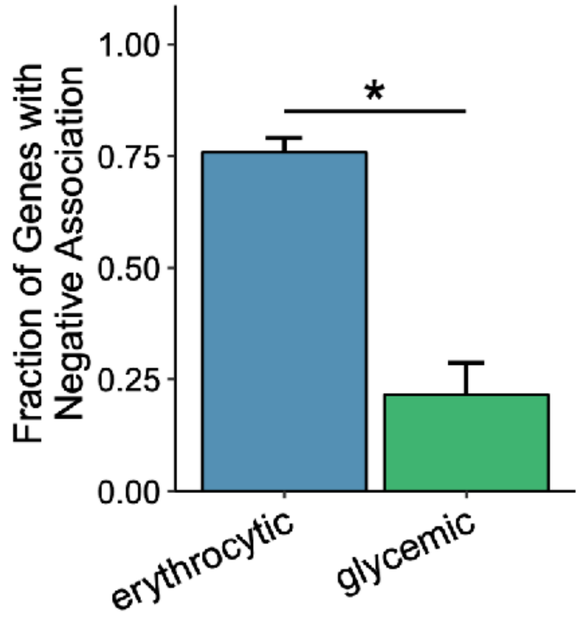


Dornbos et al, Nature Genetics 2022

Evidence for associations across many genes

- Compared to rare variants in genes involved in glycemia in mice, rare variants in erythrocytic genes are more likely to decrease HbA1c





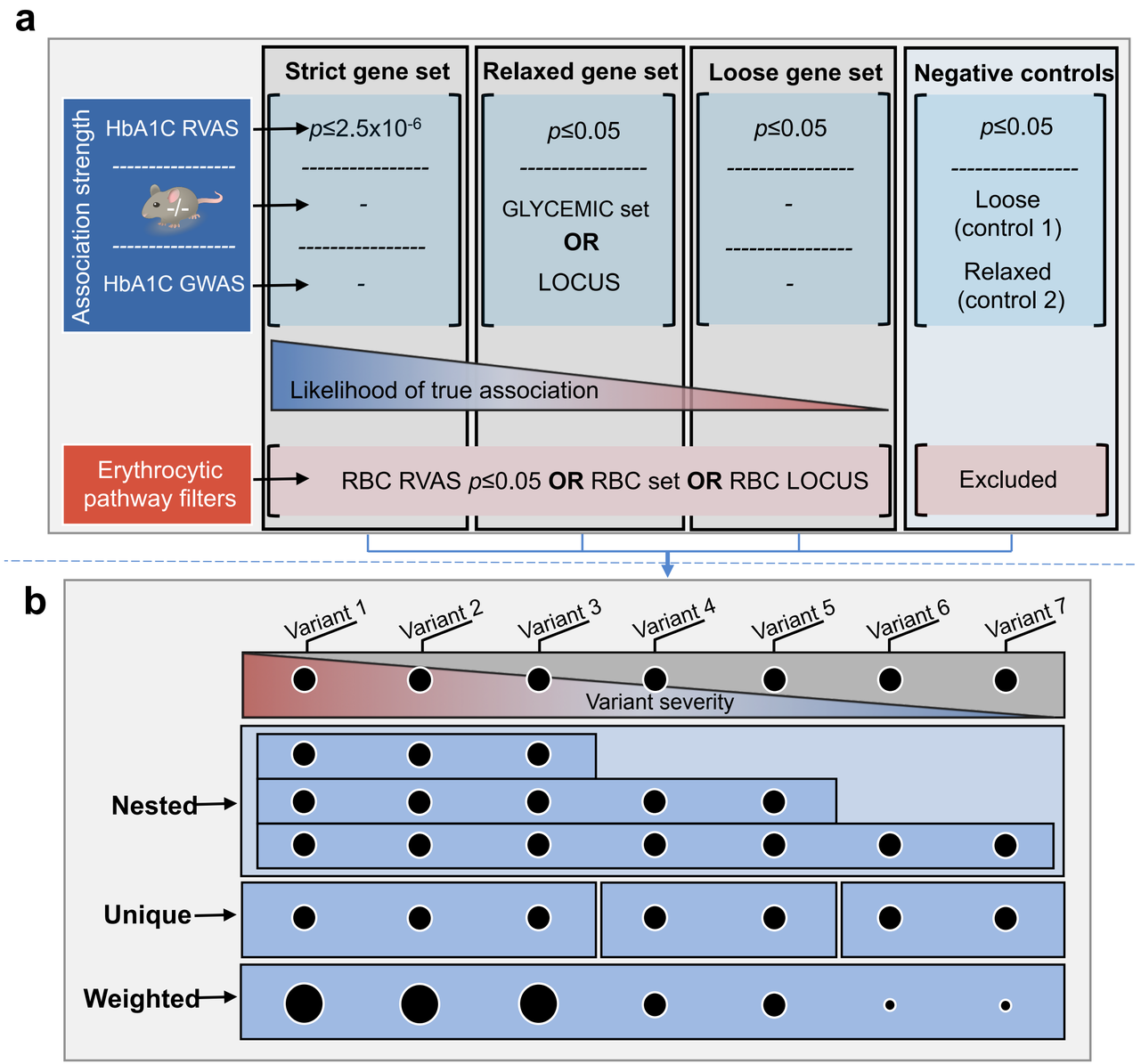
Dornbos et al, Nature Genetics 2022

Putting this together in a polygenic score

Filtering to true associations

Filtering to erythrocytic variants

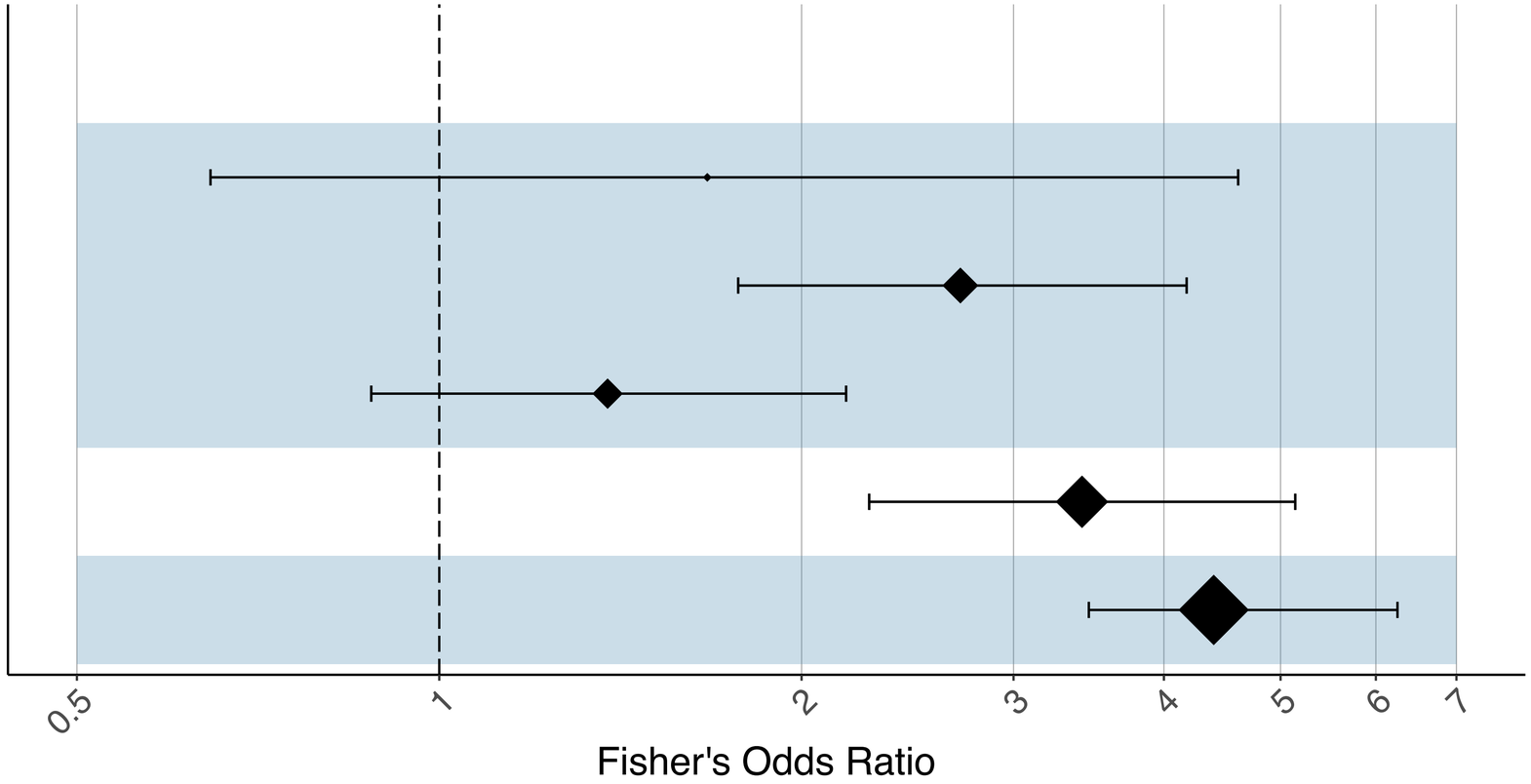
Estimate variant effect size



Dornbos et al, Nature Genetics 2022

Final model: 21,293 variants

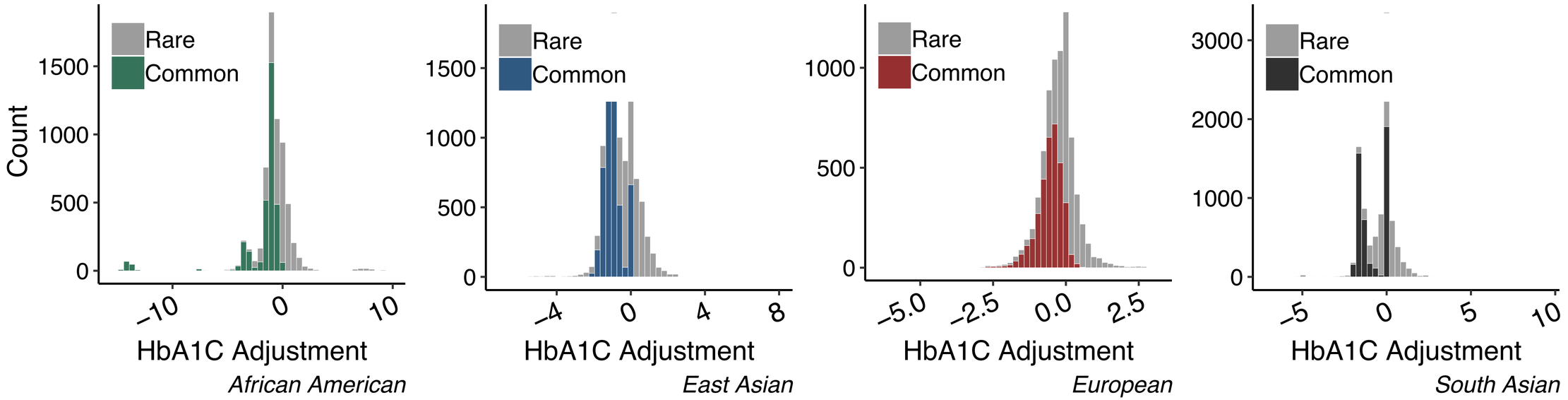
Variation	Model
Rare	PIEZO1/G6PD
	Erythrocytic Genes
	Glycemic Genes
Common	Erythrocytic Variants
Combined	Erythrocytic PGS



**Diamonds scaled to number of reclassified individuals in the test sample

Dornbos et al, Nature Genetics 2022

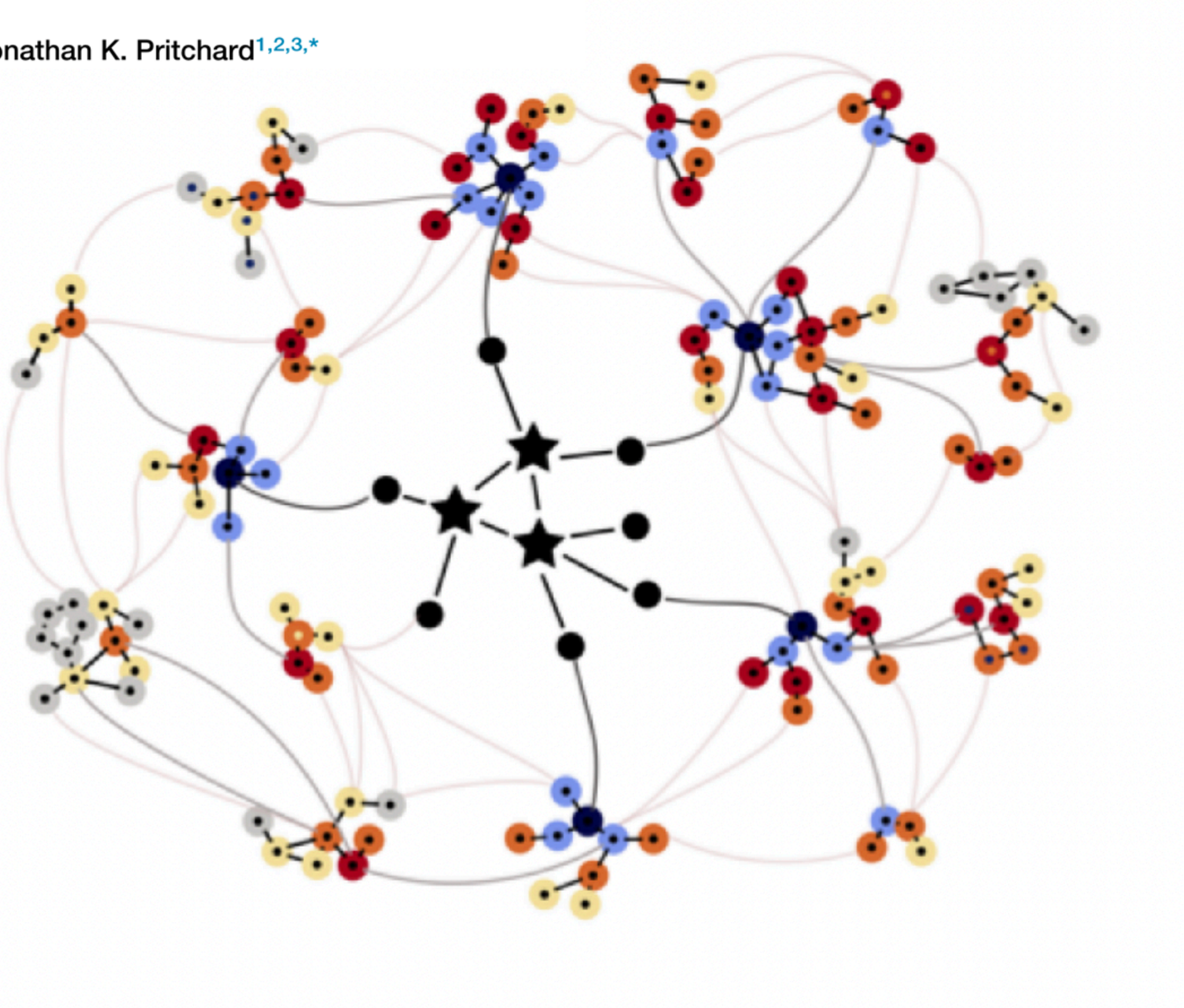
Model is highly polygenic

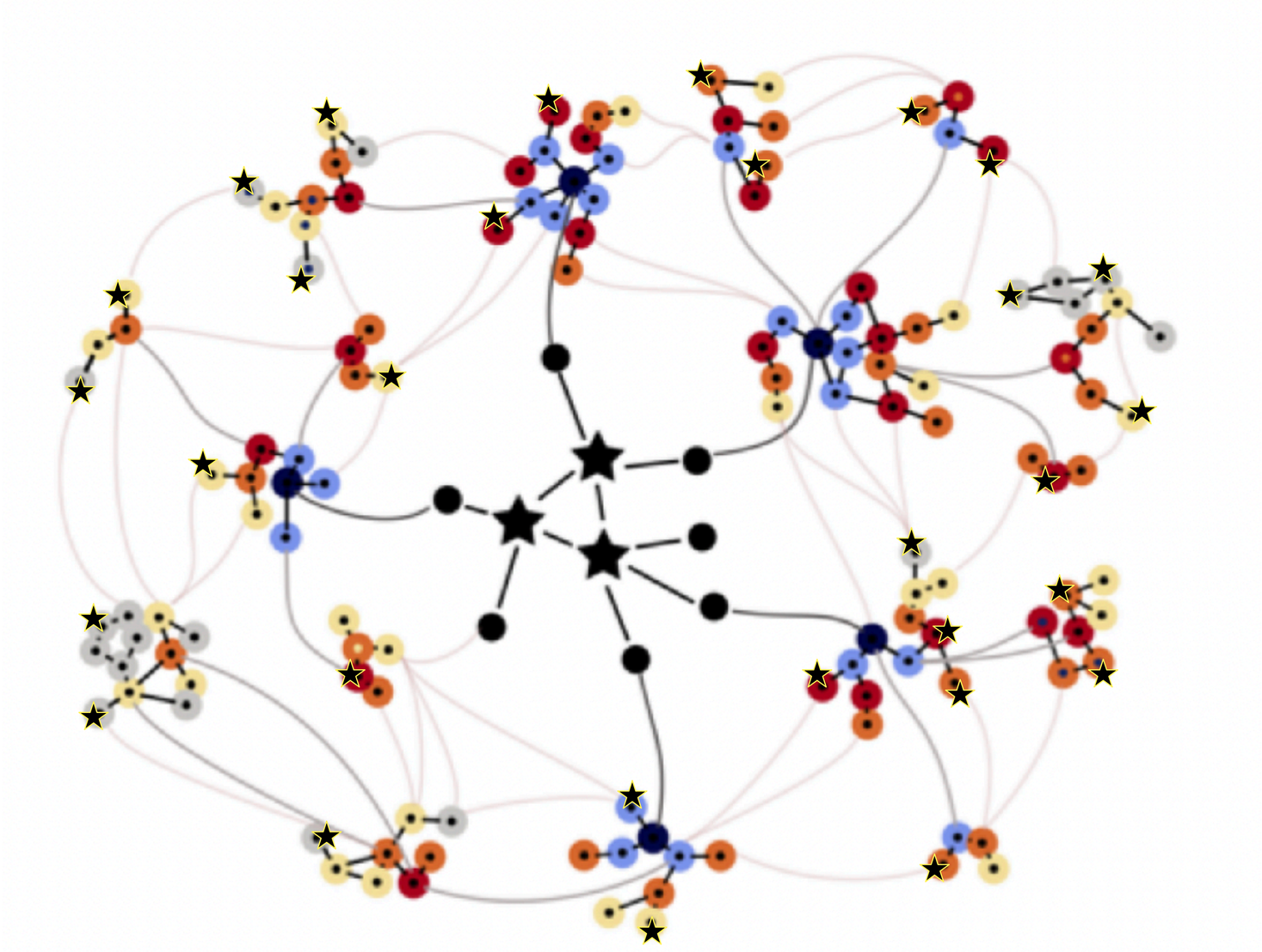


Dornbos et al, Nature Genetics 2022

An Expanded View of Complex Traits: From Polygenic to Omnigenic

Evan A. Boyle,[1,*] Yang I. Li,[1,*] and Jonathan K. Pritchard[1,2,3,*]

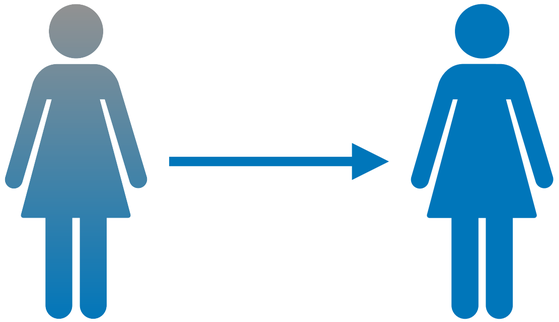


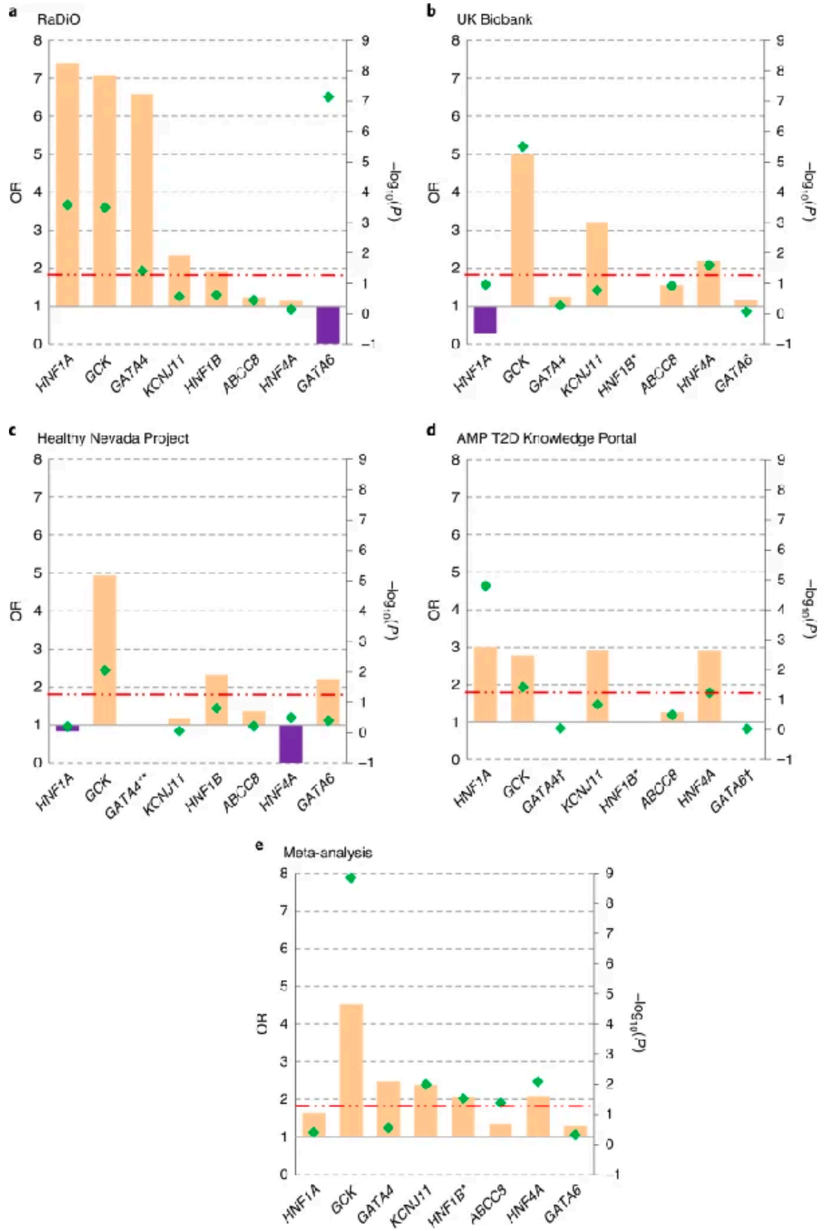


Common

Rare

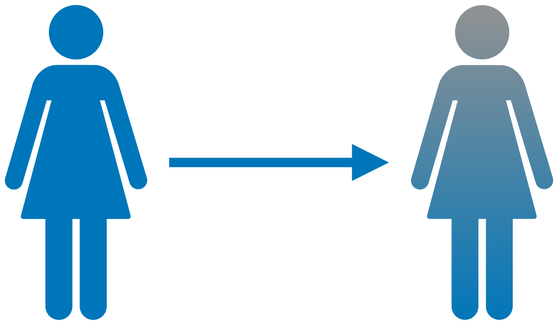
Many T2D patients carry pathogenic variants in MODY genes

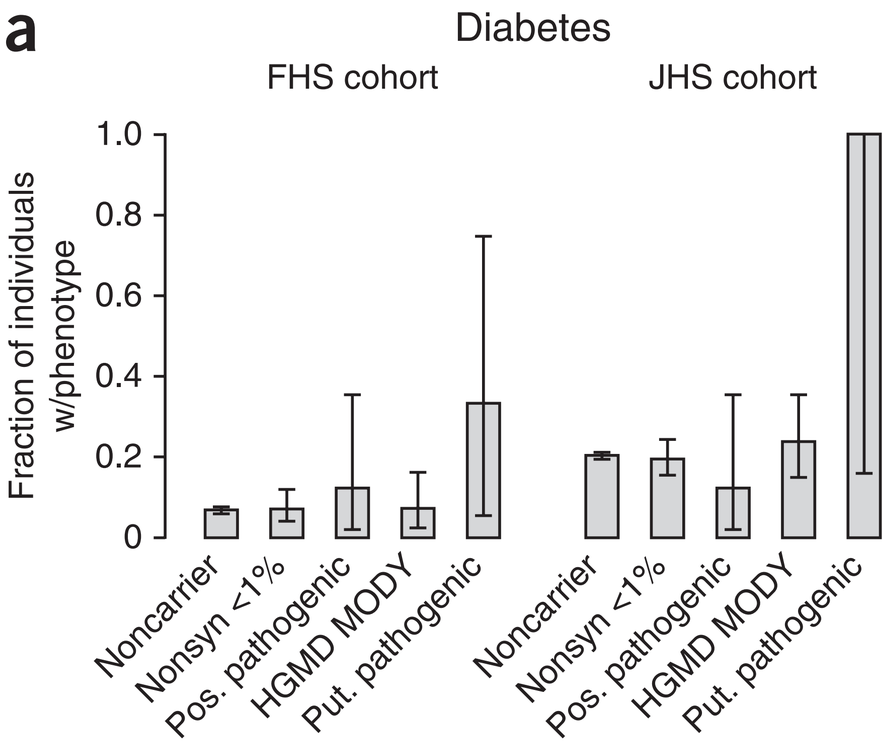


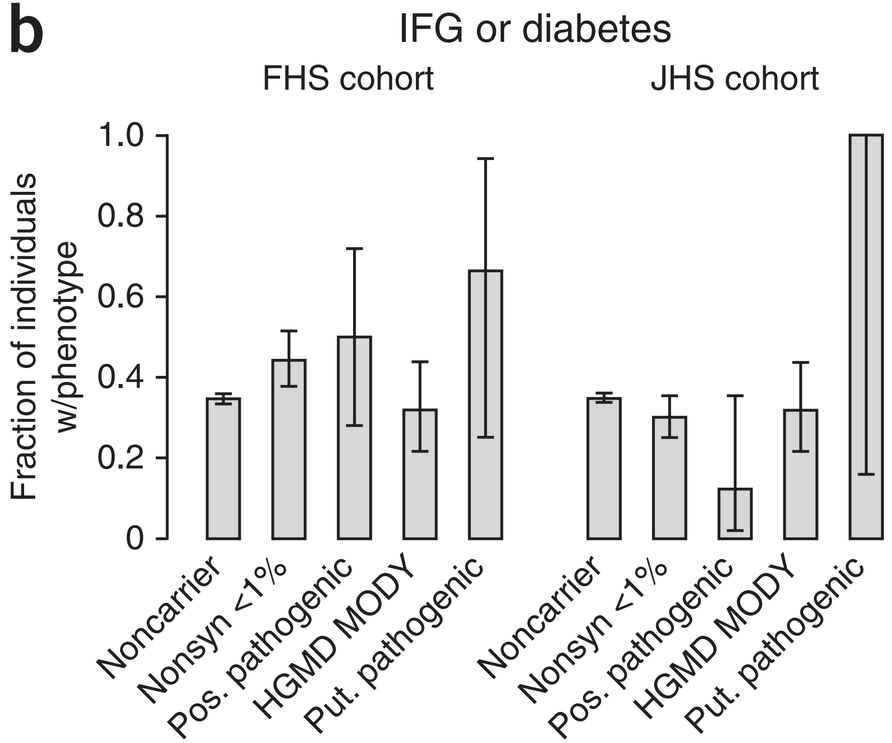


Bonnefond et al, Nature Metabolism 2020

Many damaging mutations in MODY genes are incompletely penetrant

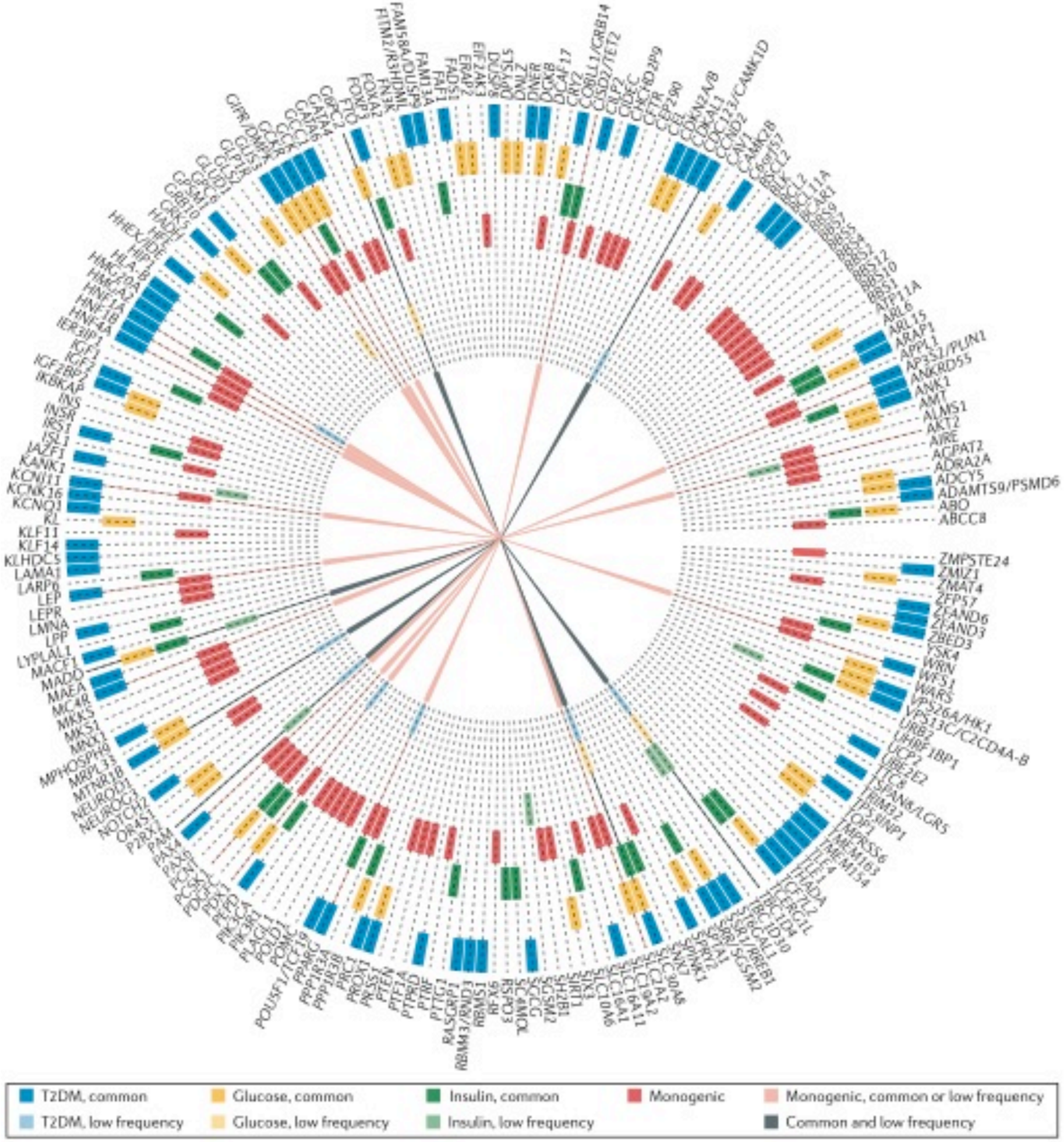






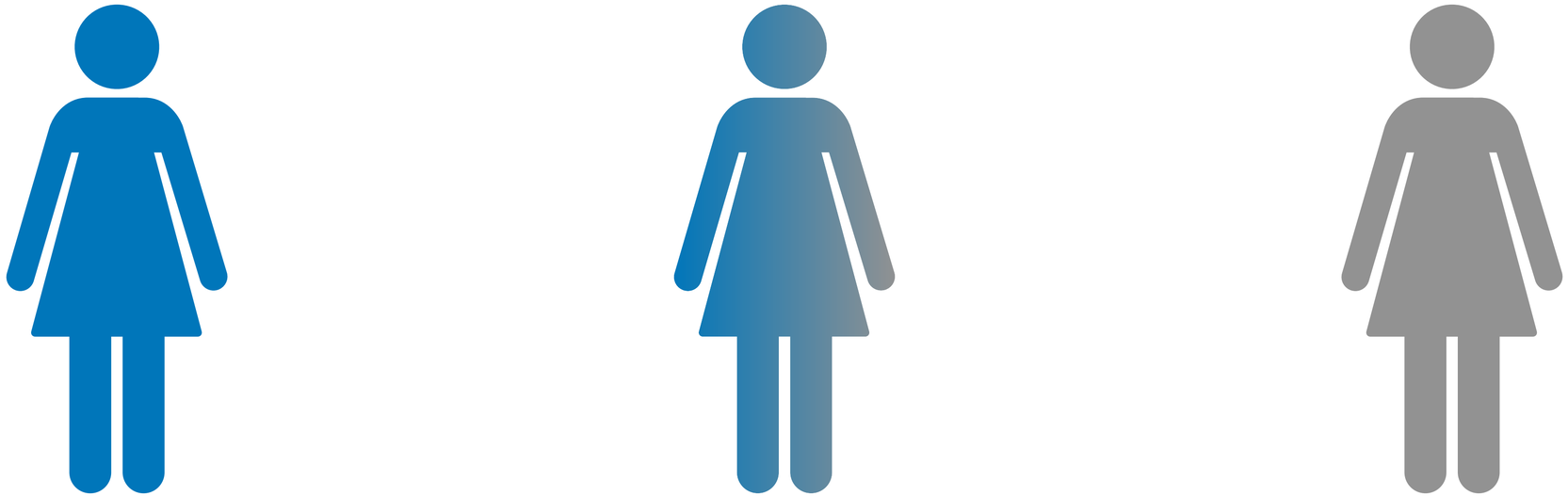
Flannick, Beer et al, Nature Genetics 2013

Rare and common forms of diabetes share genes



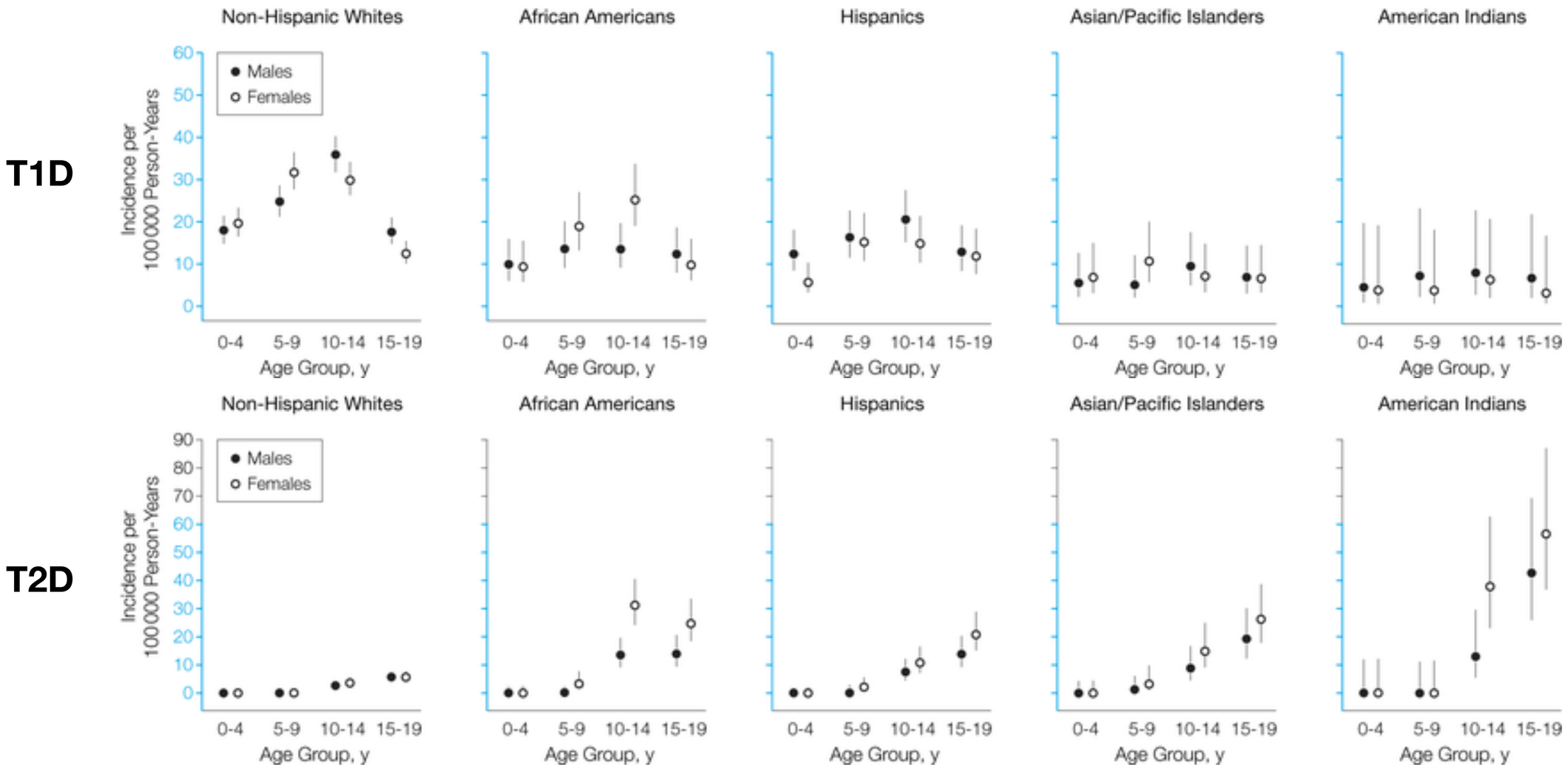
Flannick et al, 2016

?



The ProDiGY study of T2D in youth

- Incidence of ~3,700 cases/year and increasing, particularly ages 10-19
  - 15% of new diabetes cases in whites, 46-86% in minorities



Samples

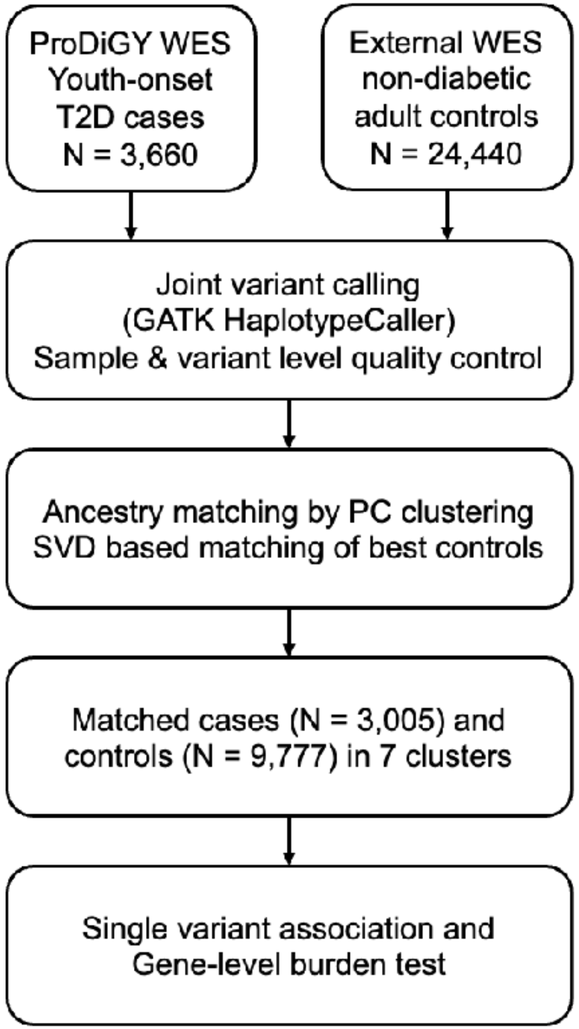
- SEARCH for Diabetes in Youth
  - Longitudinal follow up to assess natural history and complication risk factors
  - Active registry of youth diagnosed with diabetes at age < 20
- TODAY
  - Clinical trial of ages 10-17 to compare treatment efficacy of *Metformin vs Metformin+Lifestyle Intervention vs Metformin+Rosiglitazone*
  - BMI above 85th percentile
- Both studies are multi-ethnic

Ancestry	Samples
African-American	1,491 (40.8%)
East-Asian	62 (1.7%)
European	757 (20.7%)
Hispanic	1,306 (35.9%)
NA	34 (0.9%)
Total	3,650

	Total	Male	Female
N	3,650	1,294 (35.4%)	2,356 (64.6%)
Current Age	15.2±3.0	15.1±3.1	15.4±2.8
Age at Onset	13.6±2.3	13.3±2.3	14.1±2.2

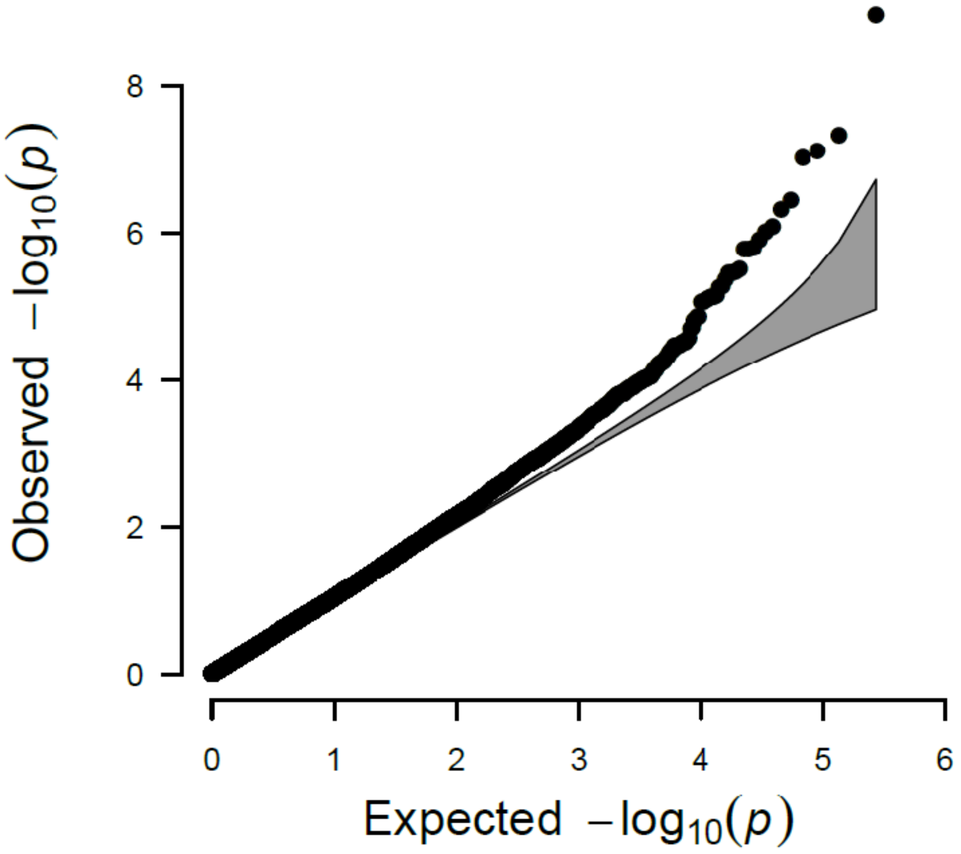
Analysis design

- Whole exome sequencing of 3,650 youth-onset T2D cases
- Match to controls from AMP-T2D exomes
  - Total analysis of 3,005 cases and 9,777 controls
- Single variant analysis (mostly for common variants)
- Gene-level analysis (for rare variants)

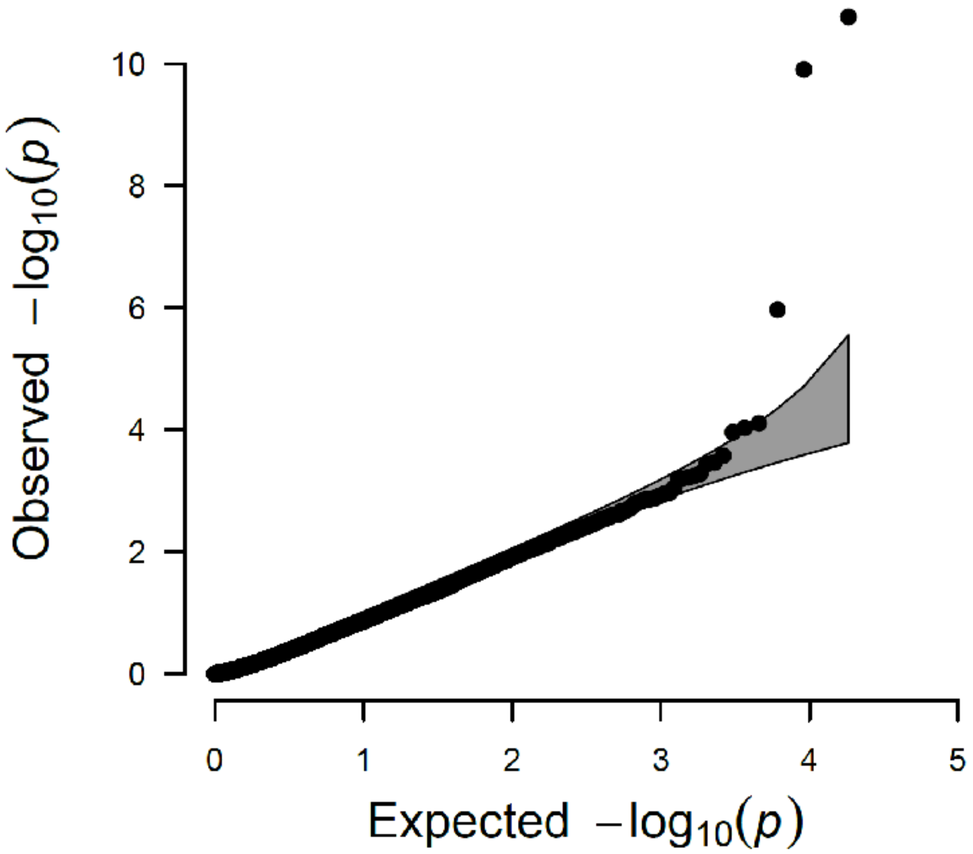


Statistics are well-calibrated

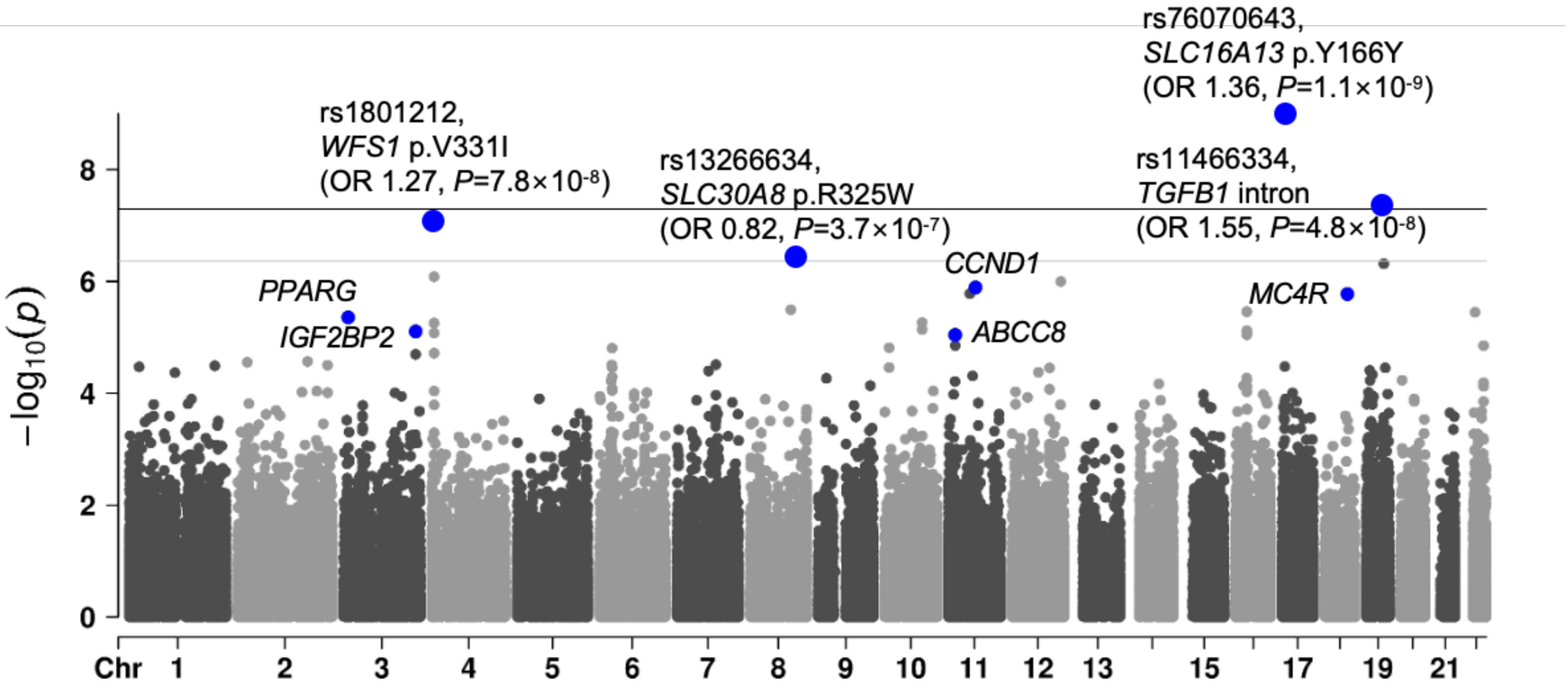
Single variant results



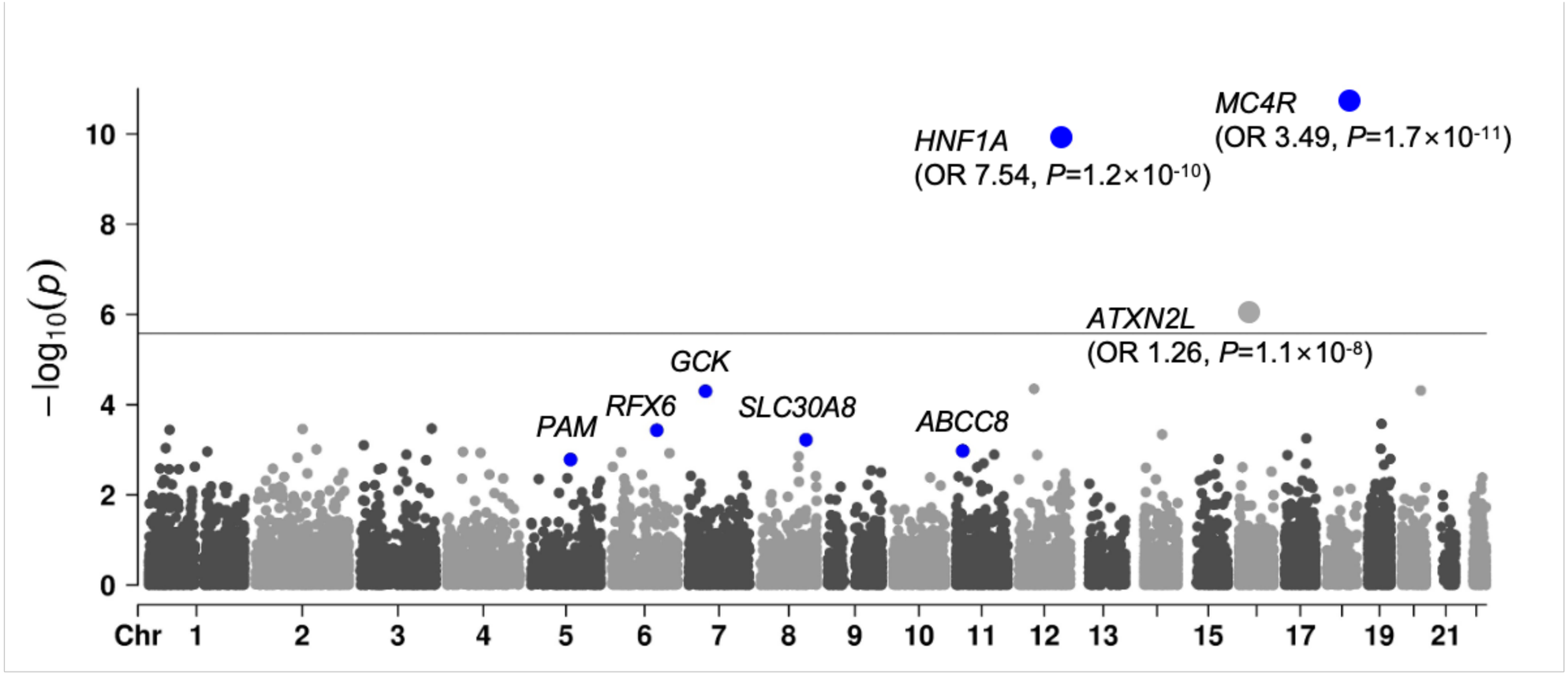
Gene-level results



Four exome-wide significant associations



Three exome-wide significant gene-level associations

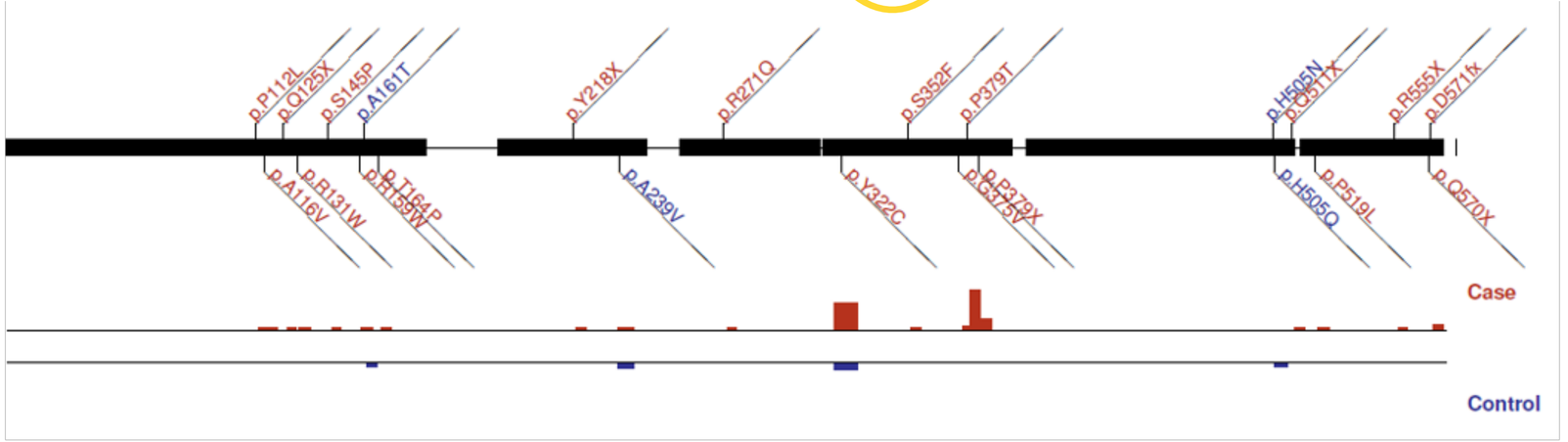


- **Additionally:** 2.1% of cases carry a monogenic diabetes causing variant

Example association: HNF1A

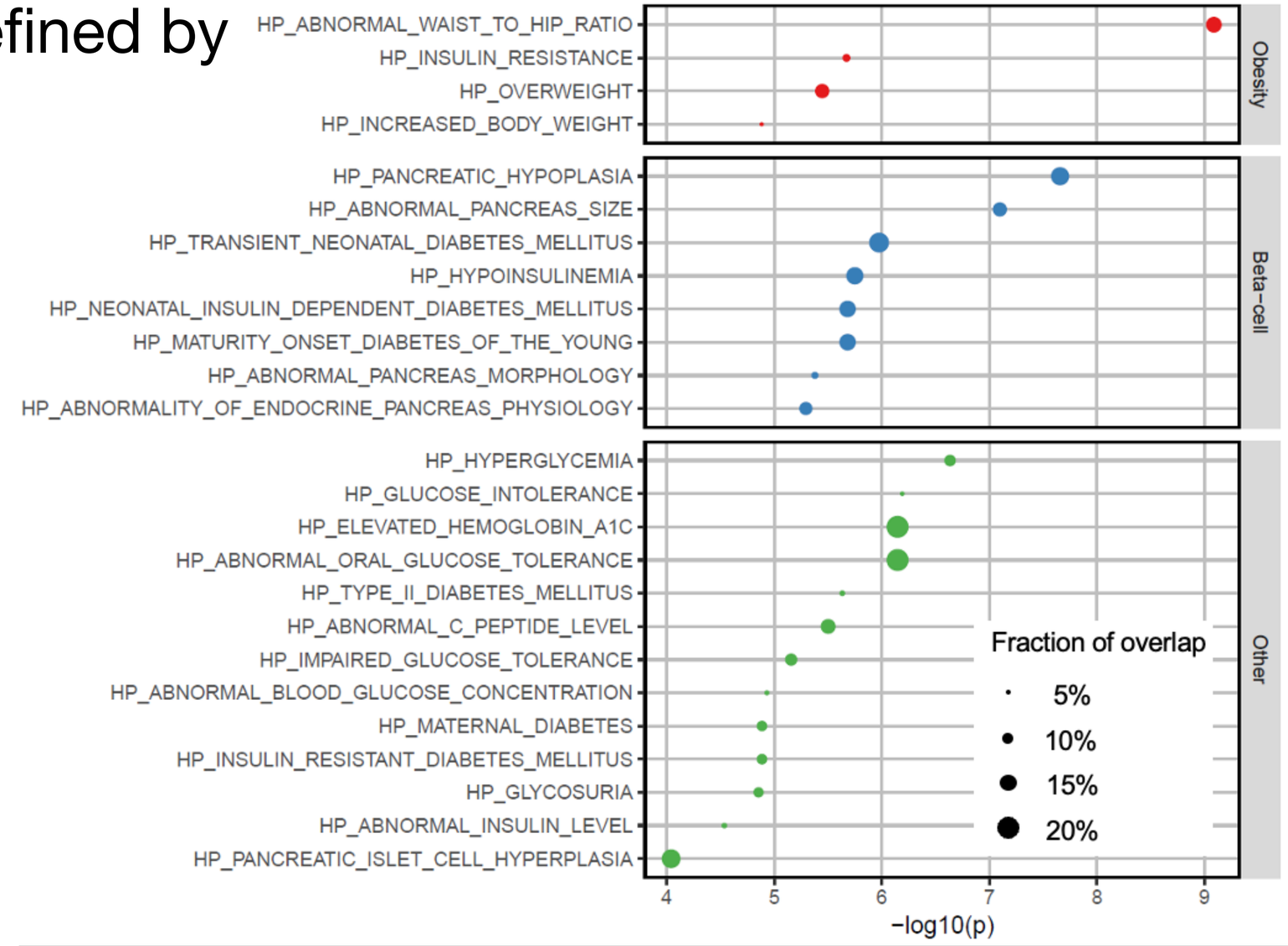
More damaging

	Total			
MASK	**# Var**	**CAF**	**OR (95% CI)**	**P**
LofTee	6	0.00047	46.0 (1.76-1198)	9.55×10^{-5}
16/16	6	0.00047	46.0 (1.76-1198)	9.55×10^{-5}
11/11	21	0.0038	7.51 (3.92-14.4)	2.24×10^{-11}
5/5	22	0.0039	6.77 (3.63-12.6)	7.59×10^{-11}
5/5 + LofTee LC 1%	23	0.0040	7.01 (3.77-13.0)	2.41×10^{-11}
5/5 + 1/5 1%	59	0.018	1.80 (1.41-2.31)	4.21×10^{-6}
5/5 + 0/5 1%	63	0.019	1.82 (1.42-2.32)	2.70×10^{-6}



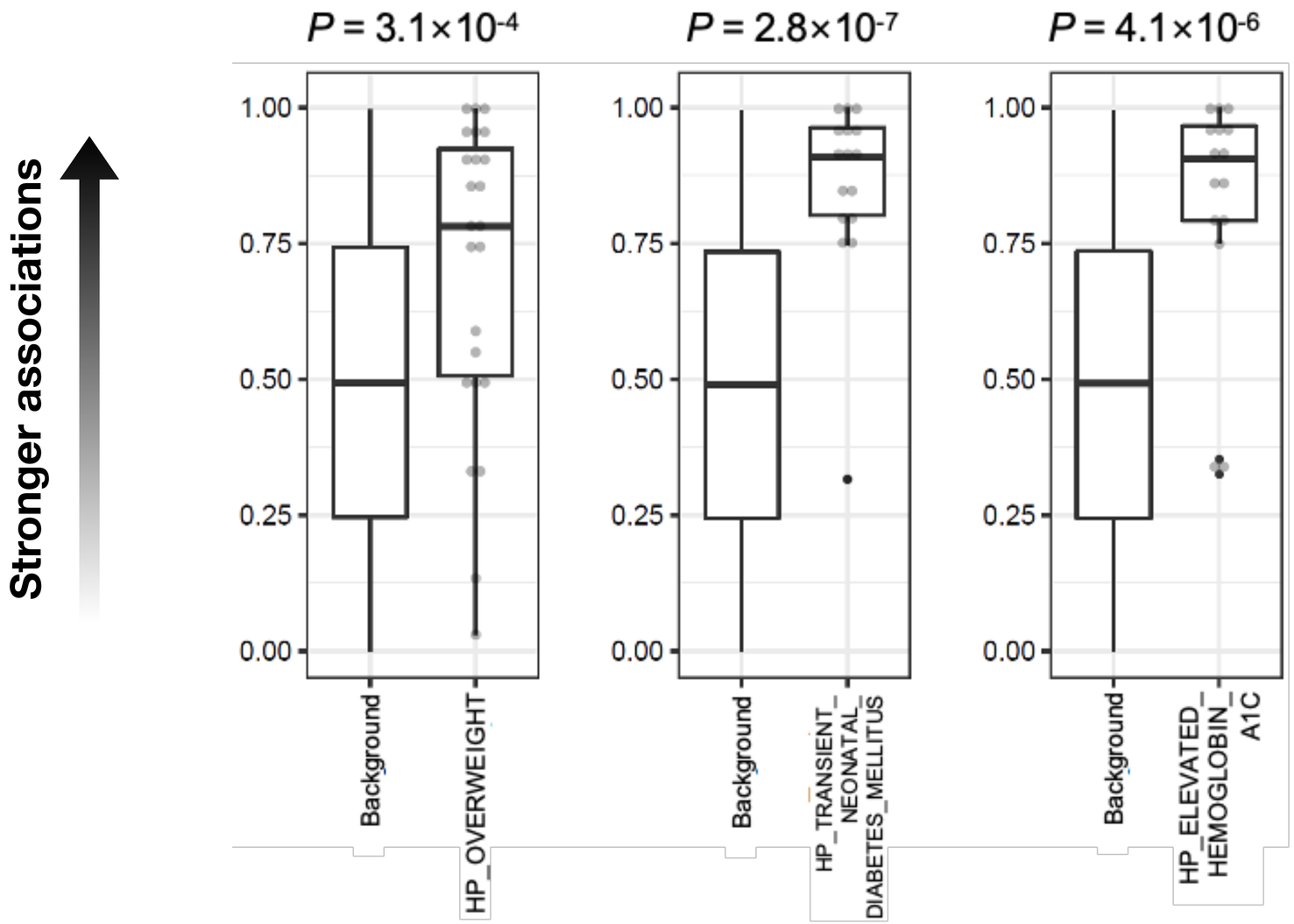
Substantial enrichment in diabetes-relevant gene sets

- Gene sets defined by HPO terms



Three categories

Enrichments are due to many genes

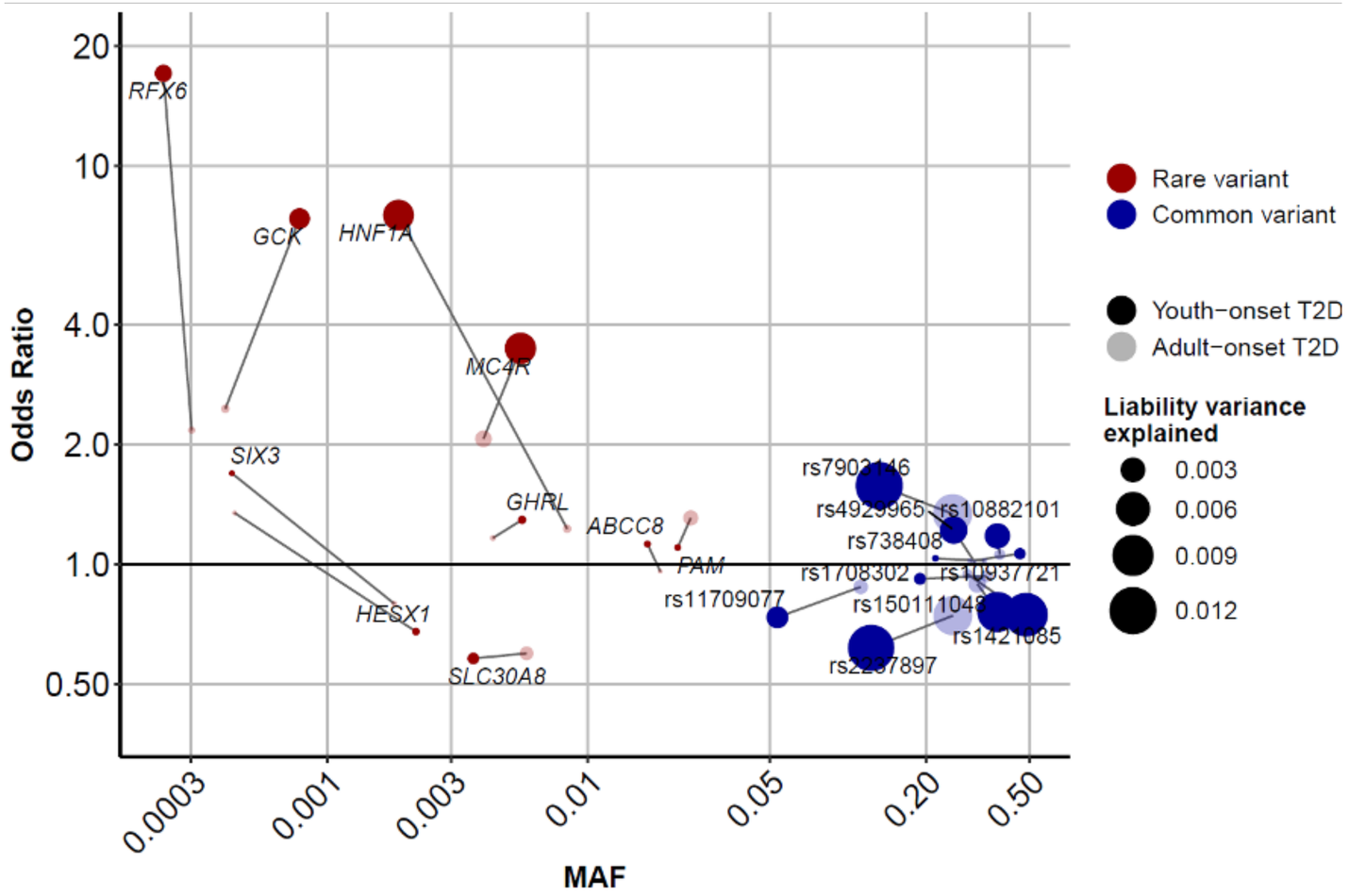


Tiers of candidate genes

- **Tier 1:** Exome-wide significant genes (*MC4R*, *HNF1A*, *ATXN2L*)
- **Tier 2:** among top 50 and causal for monogenic diabetes or T2D (*GCK*, *SLC30A8*, *ABCC8*, *PAM*)
- **Tier 3:** among the top 50 and in an enriched HPO gene set (*RFX6*, *GHRL*, *HESX1*, *SIX3*)
- **Tier 4:** $p<0.05$ and in a diabetes-relevant gene set (38 additional genes)

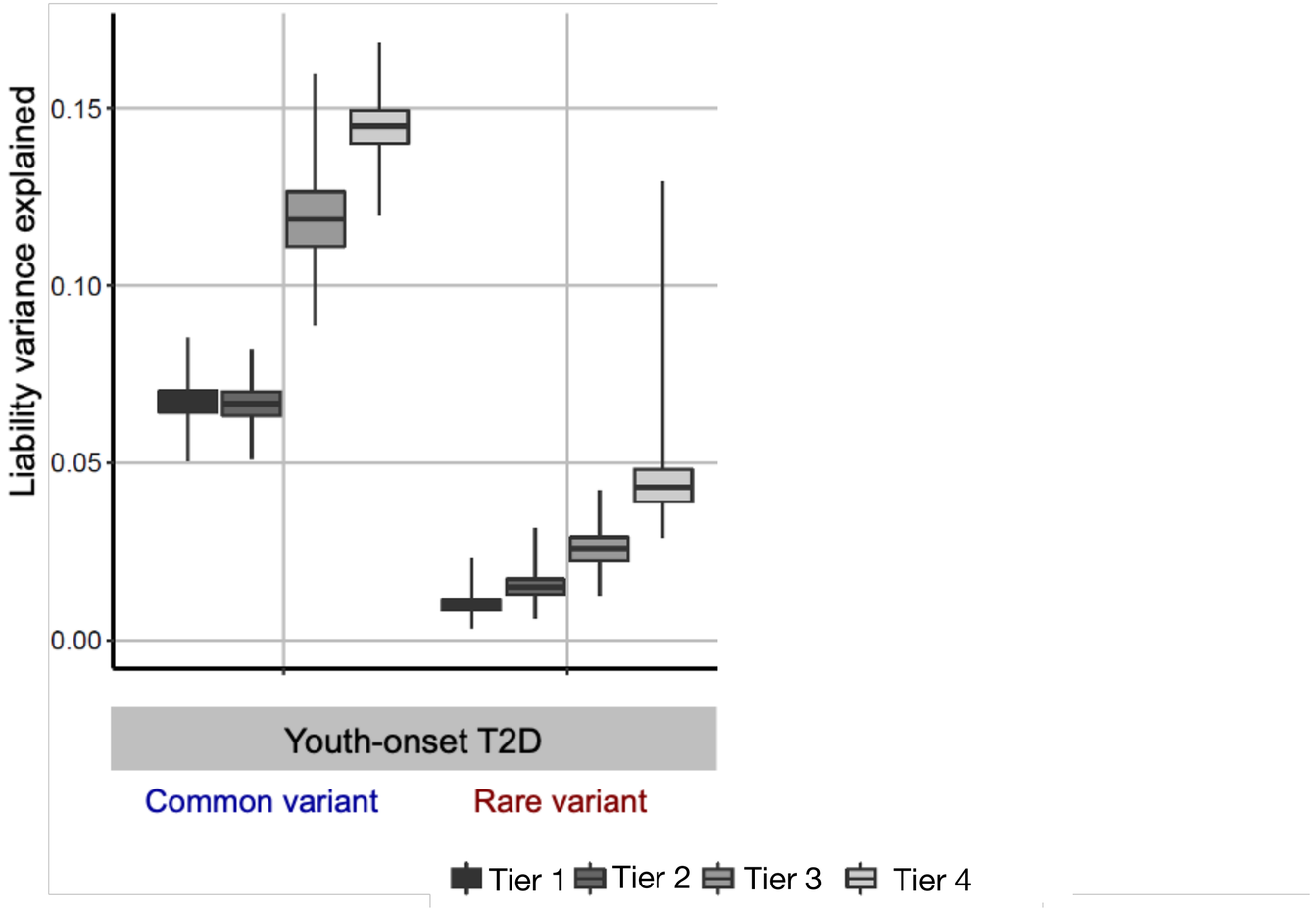
Both common and rare variants are enriched in ProDiGY

(relative to adult-onset T2D cases)



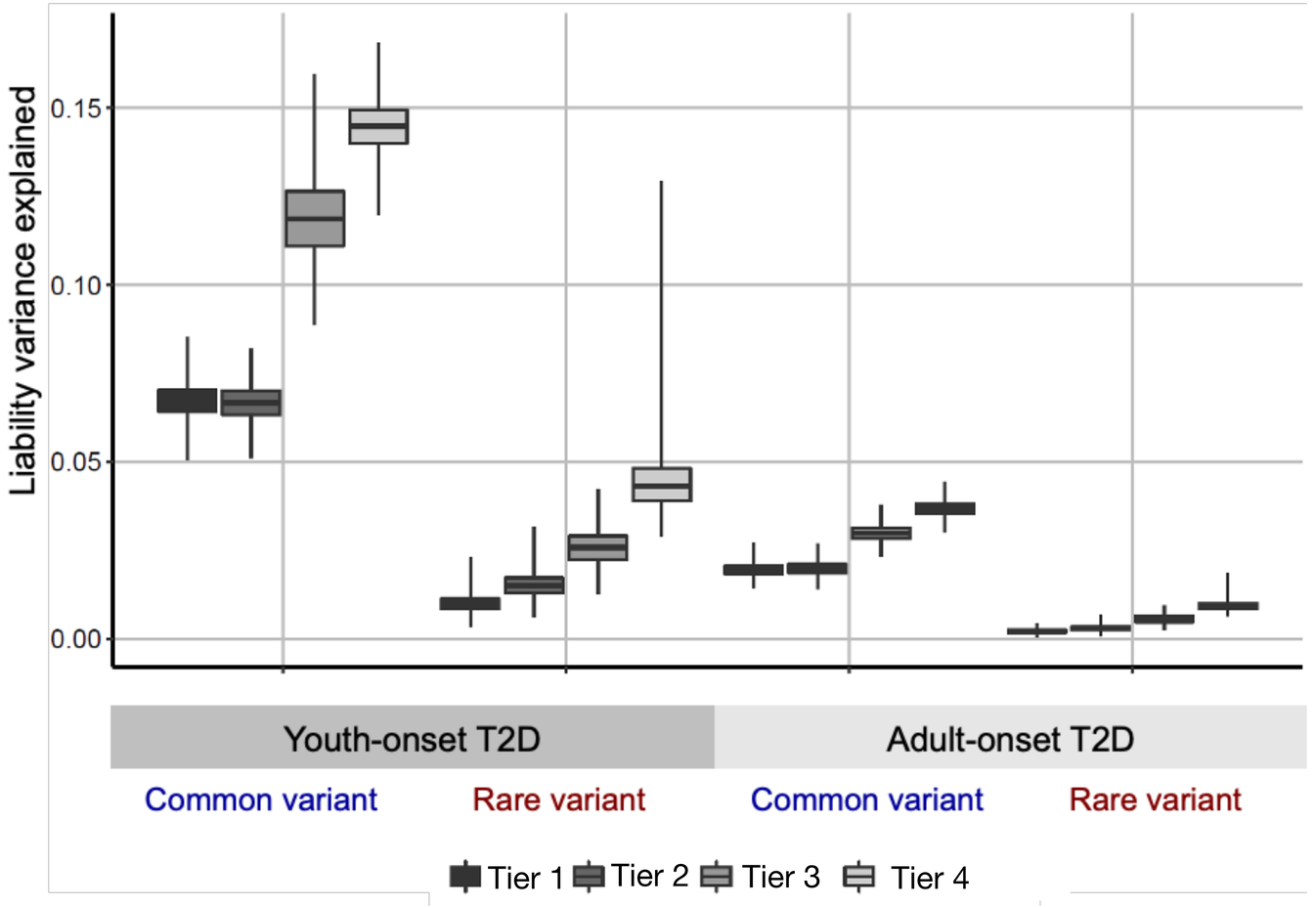
Both common and rare variants explain more heritability

(relative to adult-onset T2D cases)



Both common and rare variants explain more heritability

(relative to adult-onset T2D cases)

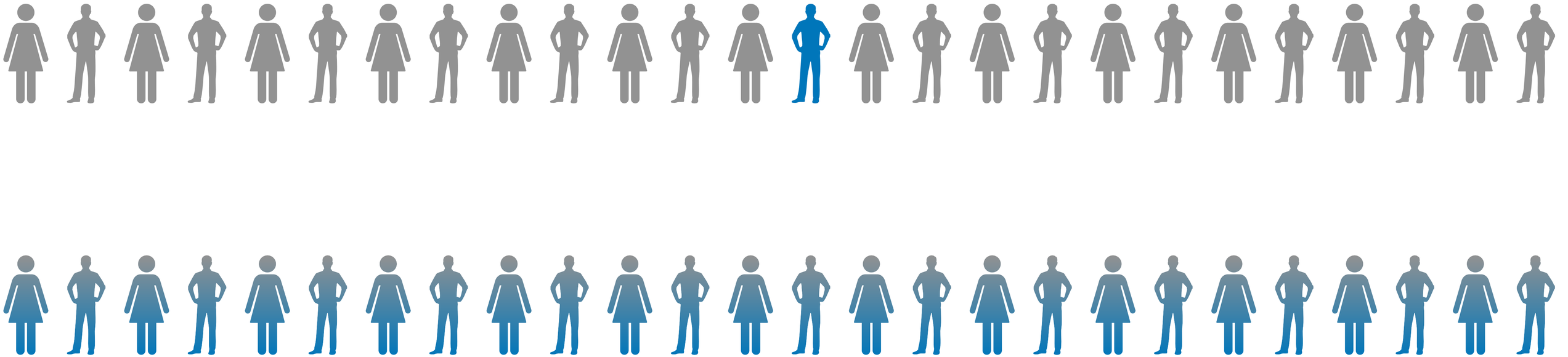


As a **population**, youth-onset T2D cases are enriched for all types of genetic risk factors

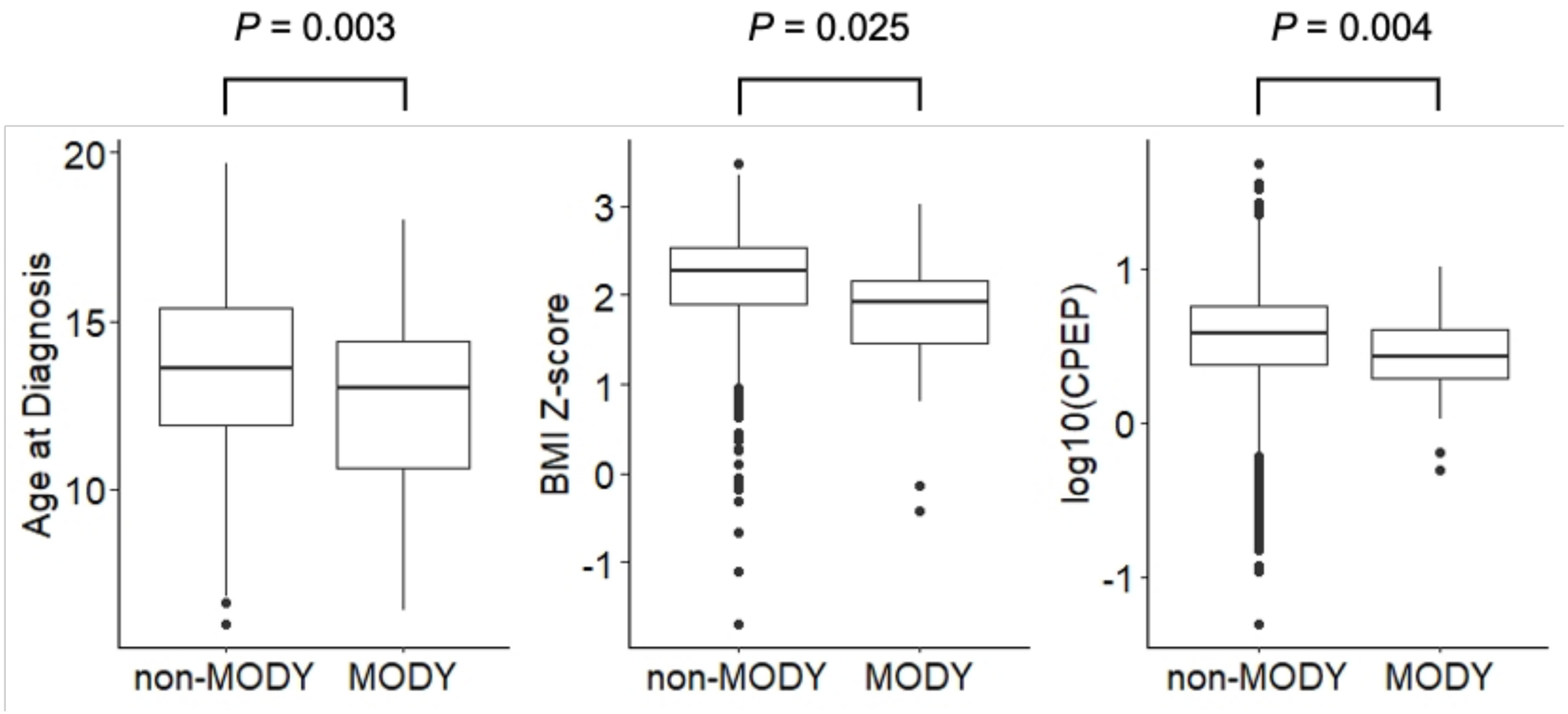
2.1% carry monogenic variants (MODY cases)
5.0-fold more rare variants than adult-onset cases
3.4-fold more common variants than adult-onset cases

skew towards **common variants in absolute** terms
skew towards **rare variants relative** to adult-onset T2D

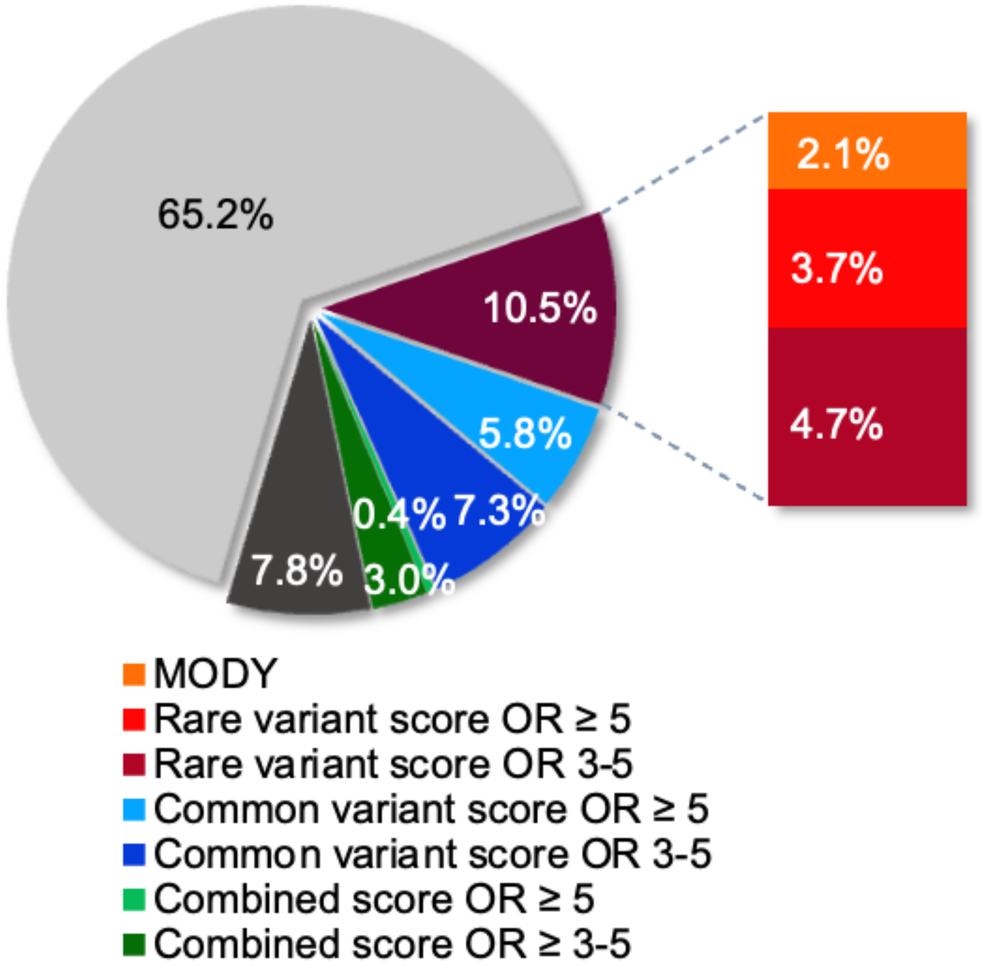
What about individually?



Cases due to MODY mutations are phenotypically different

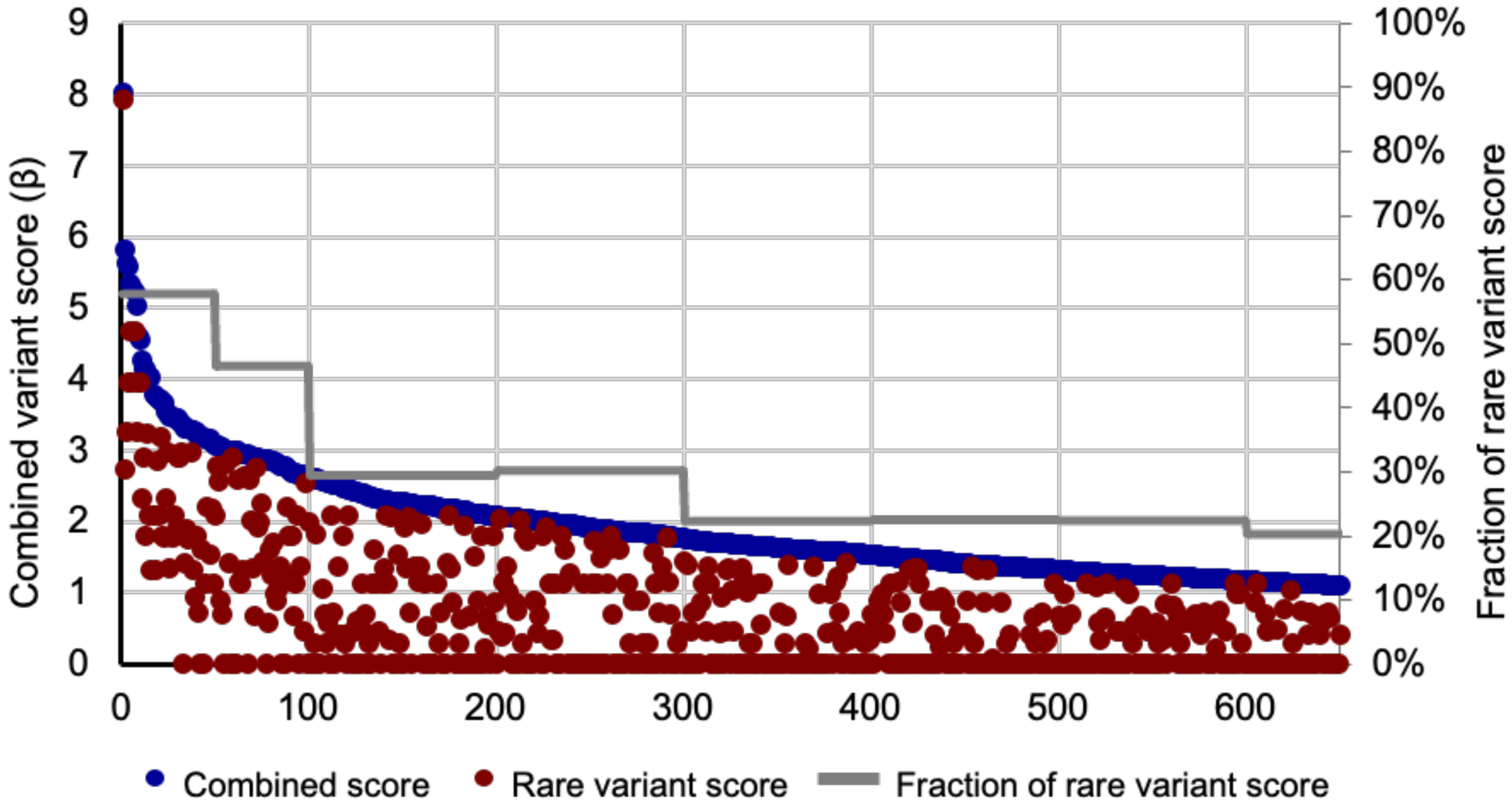


Start with cases “explained” by rare or common variants

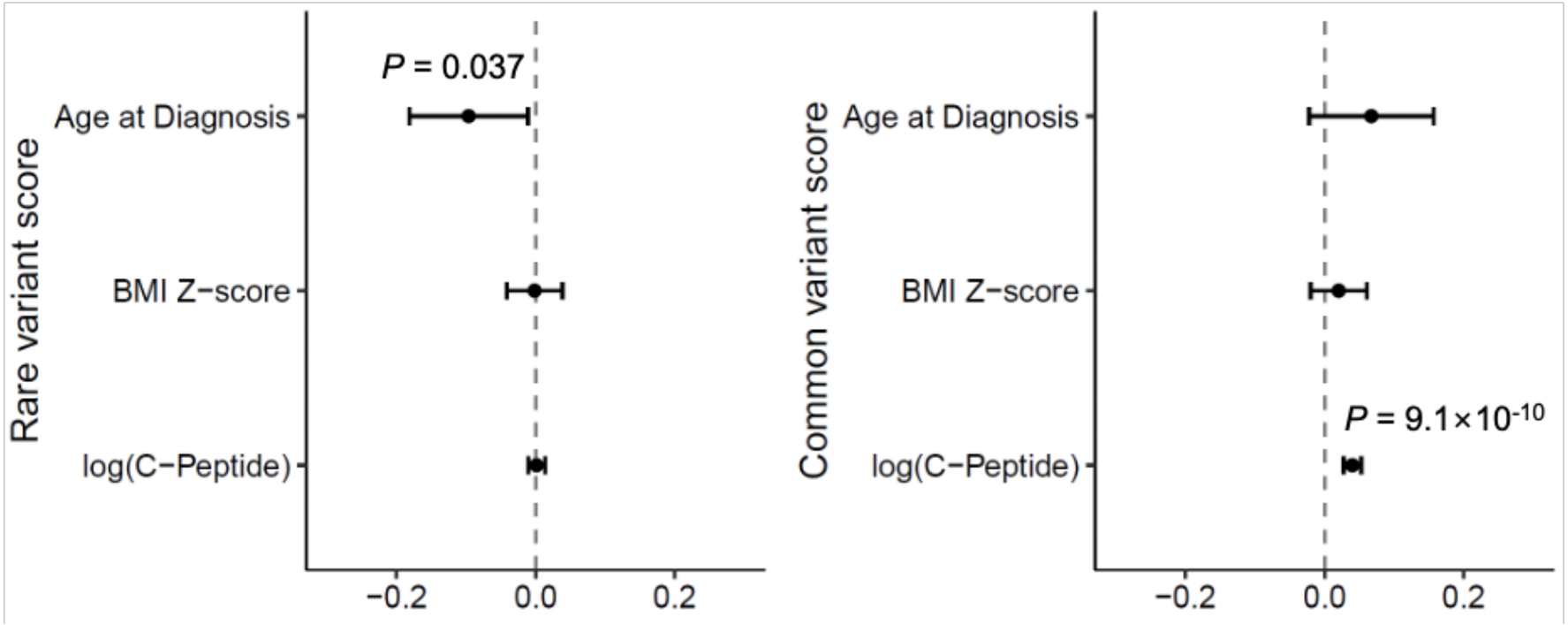


No clear dividing line between cases due to rare vs. common variants

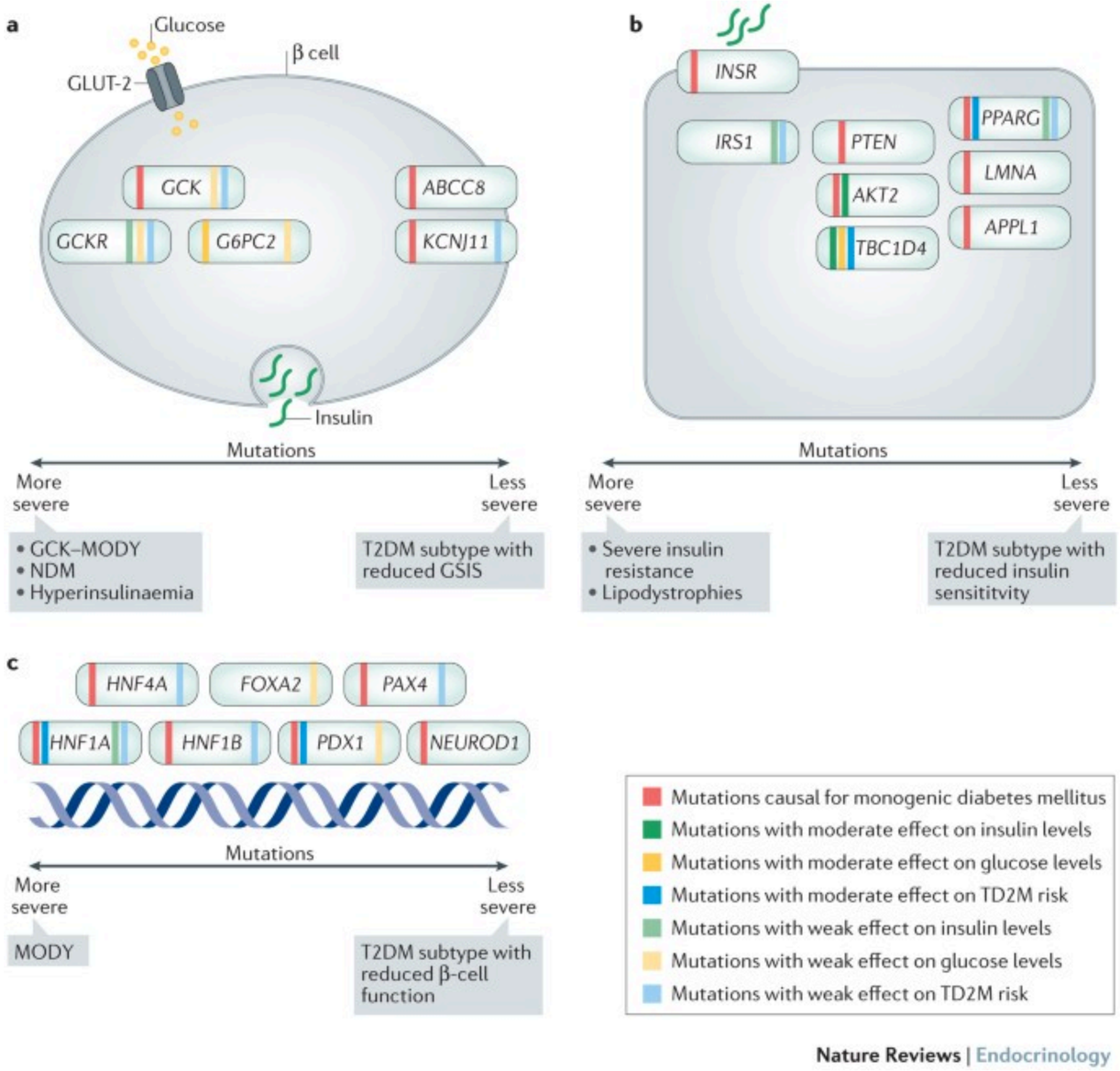
- But, a substantial amount of heterogeneity across cases



Cases due to rare vs. common variants are phenotypically different

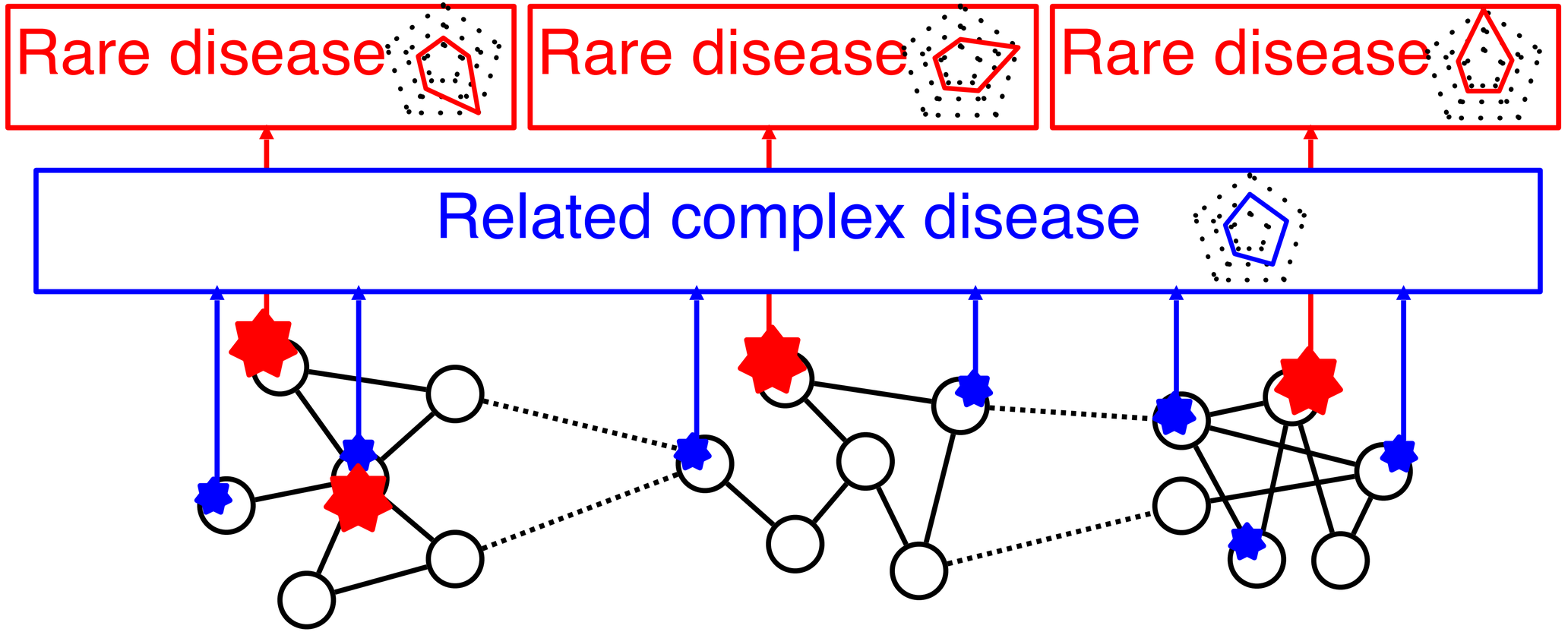


Model: allelic series are pervasive across genes and pathways

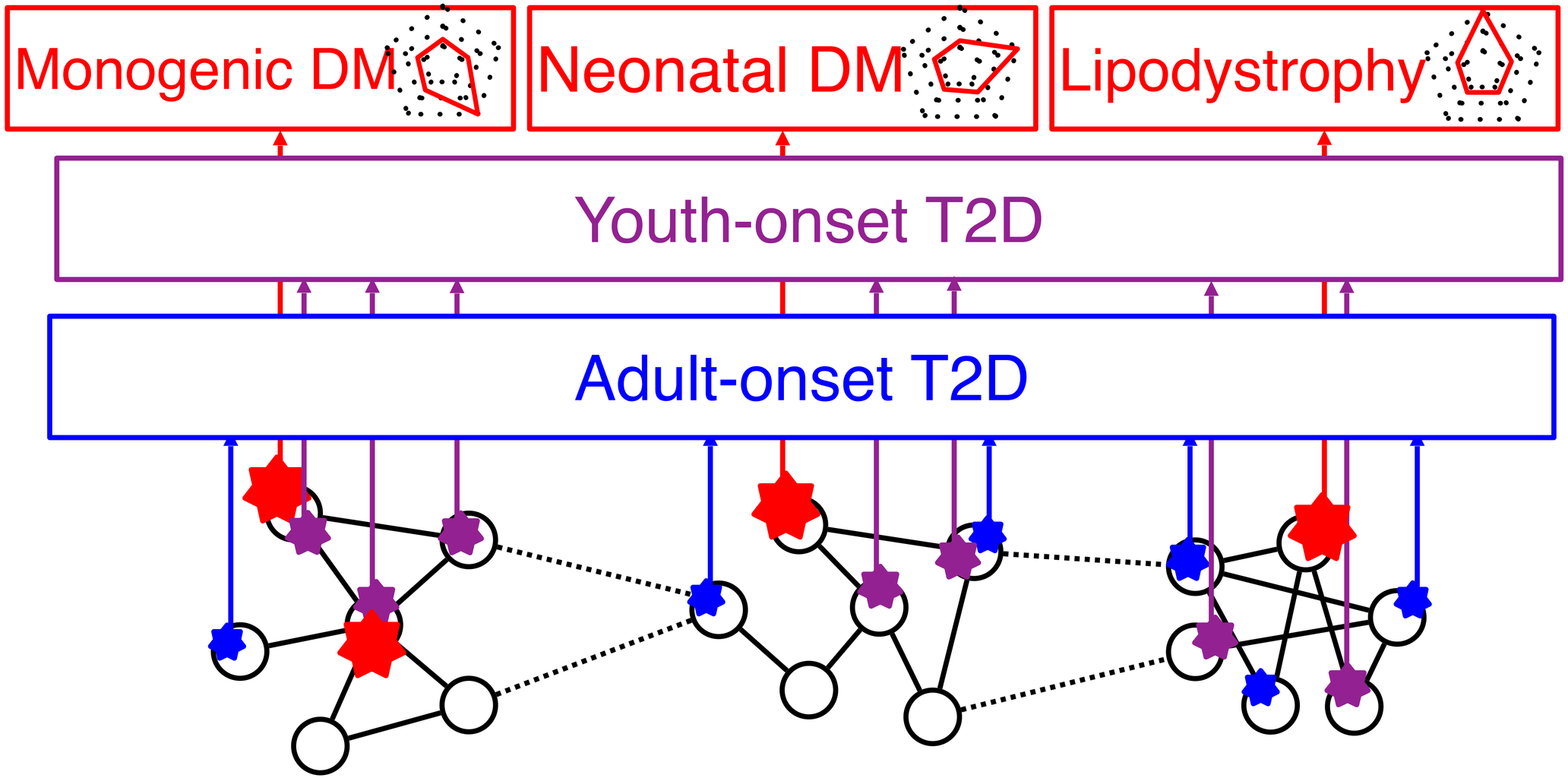


Flannick et al, 2016

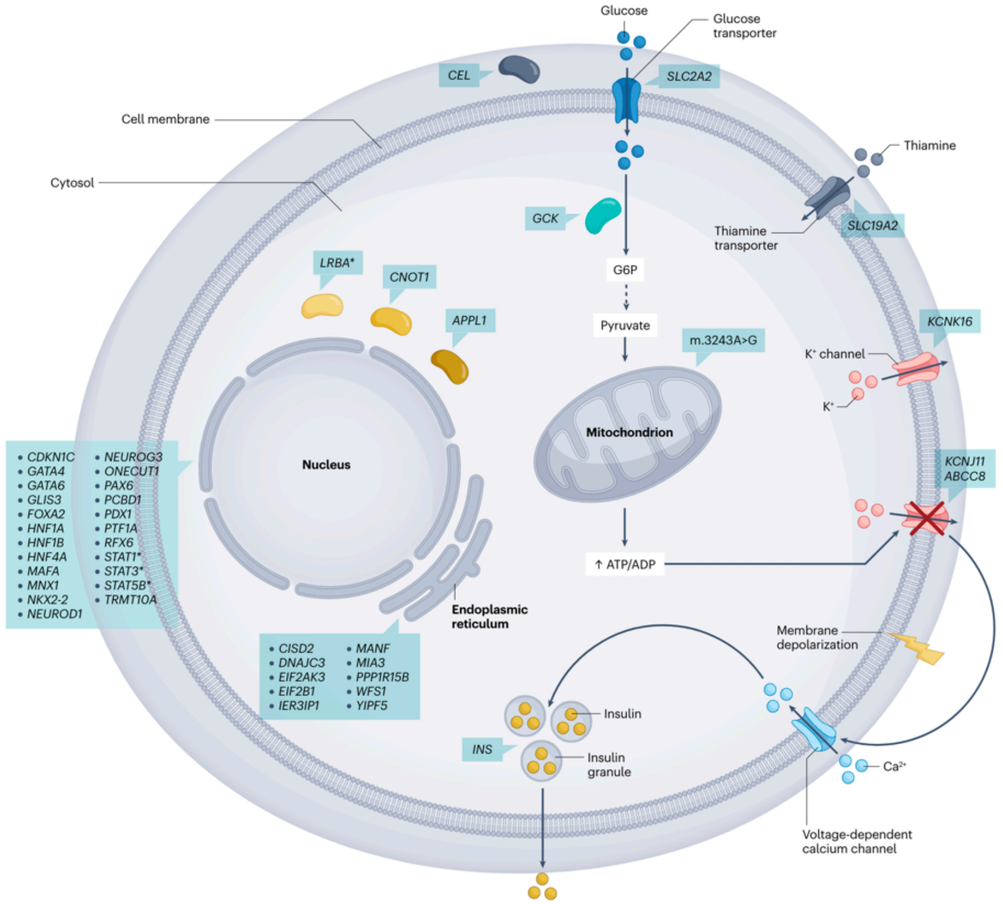
Model: allelic series are pervasive across genes and pathways



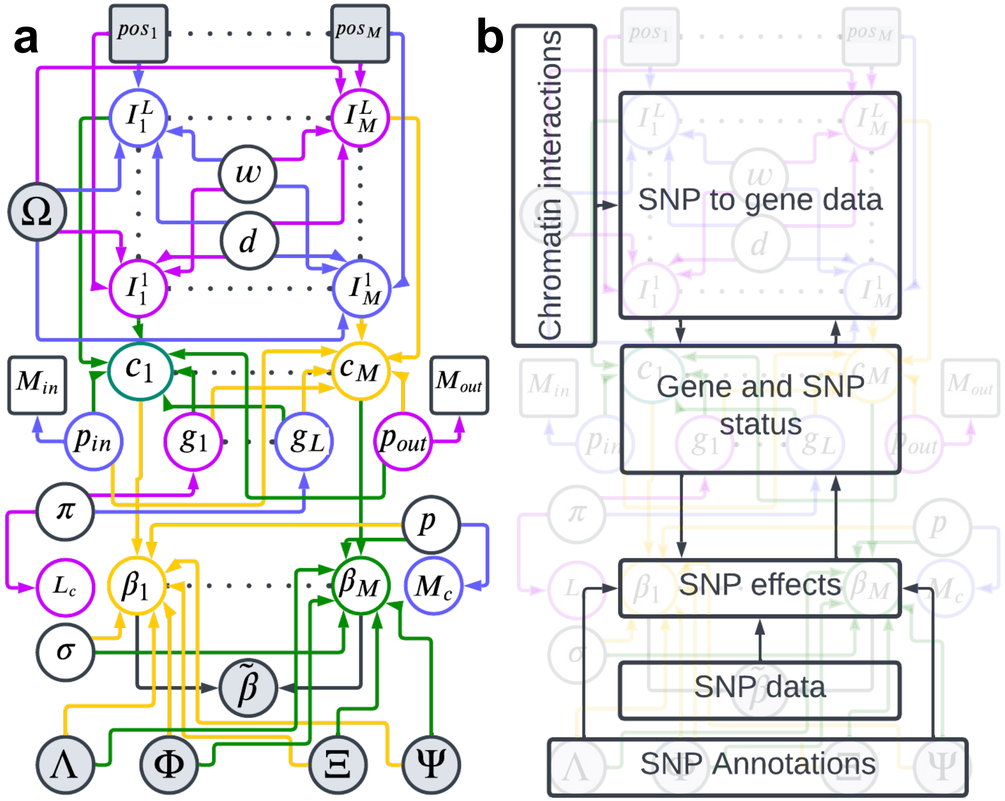
Model: allelic series are pervasive across genes and pathways



What's next?


Glucose
Glucose transporter
SLC2A2
CEL
Cell membrane
Cytosol
Thiamine
SLC19A2
Thiamine transporter
GCK
G6P
Pyruvate
m.3243A>G
LRBA*
CNOT1
APPL1
KCNK16
K+ channel
K+
Mitochondrion
Nucleus
KCNJ11
ABCC8
↑ ATP/ADP
Endoplasmic reticulum
Membrane depolarization
Insulin
INS
Insulin granule
Ca2+
Voltage-dependent calcium channel
• CDKN1C • GATA4 • GATA6 • GLIS3 • FOXA2 • HNF1A • HNF1B • HNF4A • MAFA • MNX1 • NKX2-2 • NEUROD1
• NEUROG3 • ONECUT1 • PAX6 • PCBD1 • PDX1 • PTF1A • RFX6 • STAT1* • STAT3* • STAT5B* • TRMT10A
• CISD2 • DNAJC3 • EIF2AK3 • EIF2B1 • IER3IP1
• MANF • MIA3 • PPP1R15B • WFS1 • YIPF5

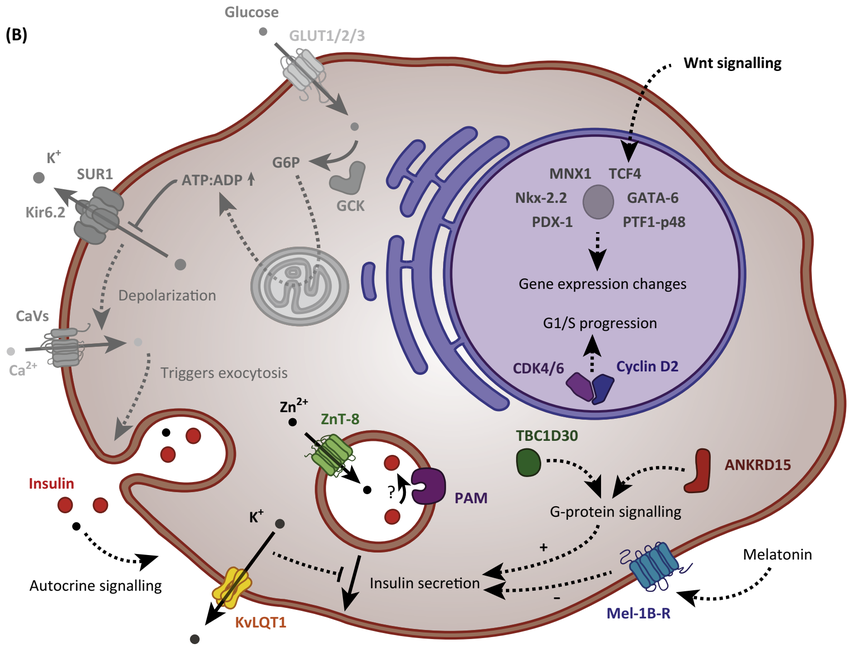
=


a
b
Chromatin interactions
SNP to gene data
Gene and SNP status
SNP effects
SNP data
SNP Annotations

Biological insights

vs.

Genetic architecture



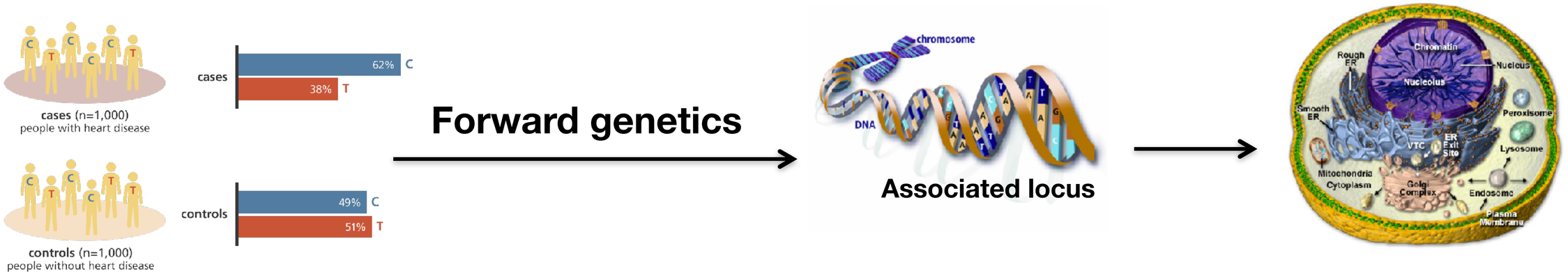
Genetic architecture

Biology

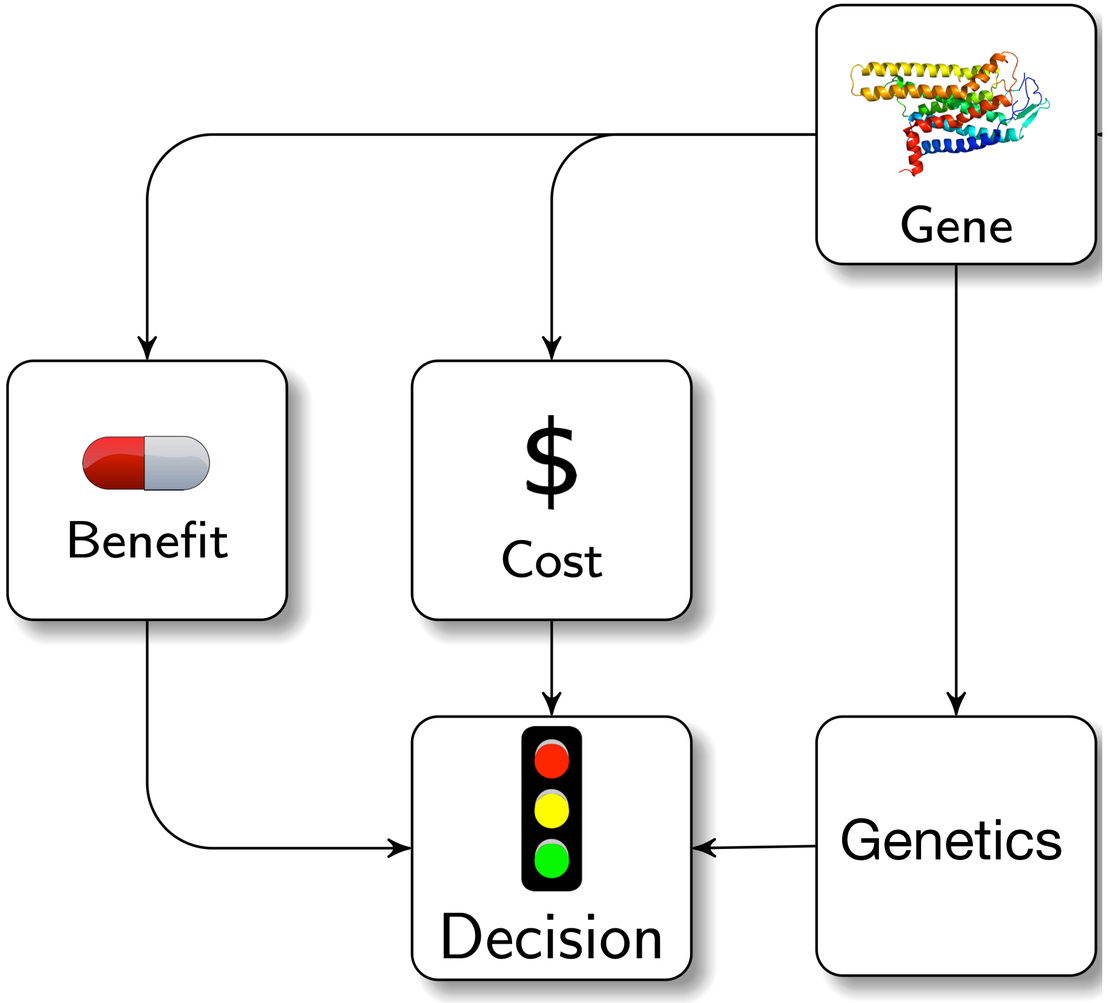
Process Name	% CPU	CPU Time	Threads	Idle Wake Ups	Kind	% GPU	GPU Time	PID	
ke[illegible]sk	41.6	26:23:40.49	601	2151	Apple	0.0	0.00	0	root
W[illegible]er	27.8	11:47:16.65	29	268	Apple	19.3	2:14:13.01	386	_wir
mds_stores	19.9	7:28:41.39	5	1	Apple	0.0	0.00	542	root
Google Chrome Helper (GPU)	12.0	3:29:07.62	33	103	Apple	8.2	31:53.17	757	flann
Google Chrome	9.4	4:53:08.90	48	68	Apple	0.0	0.00	558	flann
screencapture	6.0	0.48	2	0	Apple	0.0	0.00	46086	flann
Zotero	5.4	3:37:59.23	71	2	Intel	0.0	12.95	31740	flann
Google Chrome Helper (Renderer)	5.3	1.12	23	14	Apple	0.0	0.00	46022	flann
Slack Helper (Renderer)	3.2	3:40:56.17	18	5	Apple	0.0	0.00	2228	flann
Google Chrome Helper	2.5	2:52:45.36	31	43	Apple	0.0	0.00	758	flann
Activity Monitor	2.3	7.76	5	2	Apple	0.0	0.00	45817	flann
Slack Helper (GPU)	2.2	45:24.27	11	130	Apple	0.2	8:11.00	2221	flann
Google Chrome Helper (Renderer)	1.5	32.79	25	0	Apple	0.0	0.00	97973	flann
Cisco AnyConnect Secure Mobility Client	1.2	8:07:36.75	7	4	Intel	0.0	0.00	74765	flann
se_agent	1.1	3:41:14.86	7	1	Apple	0.0	0.00	750	root
com.cisco.anyconnect.macos.acsockext	1.0	2:35:20.73	8	0	Apple	0.0	0.00	524	root
ServiceDaemon	1.0	5:40:31.87	16	0	Intel	0.0	0.00	363	root
repmgr	0.8	16:29.45	24	10	Apple	0.0	0.00	3614	root
bluetoothd	0.7	15:27.91	10	0	Apple	0.0	0.00	375	root
JamfDaemon	0.7	1:45:11.20	6	0	Apple	0.0	0.00	323	root
httpd	0.7	5:24.92	1	0	Apple	0.0	0.00	19369	_ww
mdworker_shared	0.6	0.08	4	0	Apple	0.0	0.00	46075	flann
Screen Shot	0.6	0.10	3	0	Apple	0.0	0.00	46087	flann
sysmond	0.6	5:27.07	2	0	Apple	0.0	0.00	1084	root
sharingd	0.4	11:13.38	4	0	Apple	0.0	0.00	621	flann
Google Chrome Helper (Renderer)	0.4	1:53.20	25	2	Apple	0.0	0.00	819	flann
airportd	0.4	1:01:05.84	9	0	Apple	0.0	0.00	400	root
mds	0.3	1:27:40.12	6	7	Apple	0.0	0.00	336	root
fseventsd	0.3	34:34.76	12	11	Apple	0.0	0.00	310	root
logd	0.3	27:33.59	4	0	Apple	0.0	0.00	306	root
http://localhost	0.3	13:56.65	6	3	Apple	0.0	9.12	1852	flann
Adobe Acrobat Synchronizer	0.3	42:04.38	12	1	Apple	0.0	0.00	814	flann
Acrobat Reader Synchronizer	0.3	42:28.93	12	1	Apple	0.0	0.00	847	flann
ForeScout SecureConnector	0.3	28:36.14	4	0	Intel	0.0	0.00	344	root
mDNSResponder	0.2	21:38.17	4	8	Apple	0.0	0.00	428	_mdn
Toolkit	0.2	20:36.72	38	0	Intel	0.0	0.00	5131	flann

Stop

Where will genetic associations lead us?



Can we lead human genetics instead?

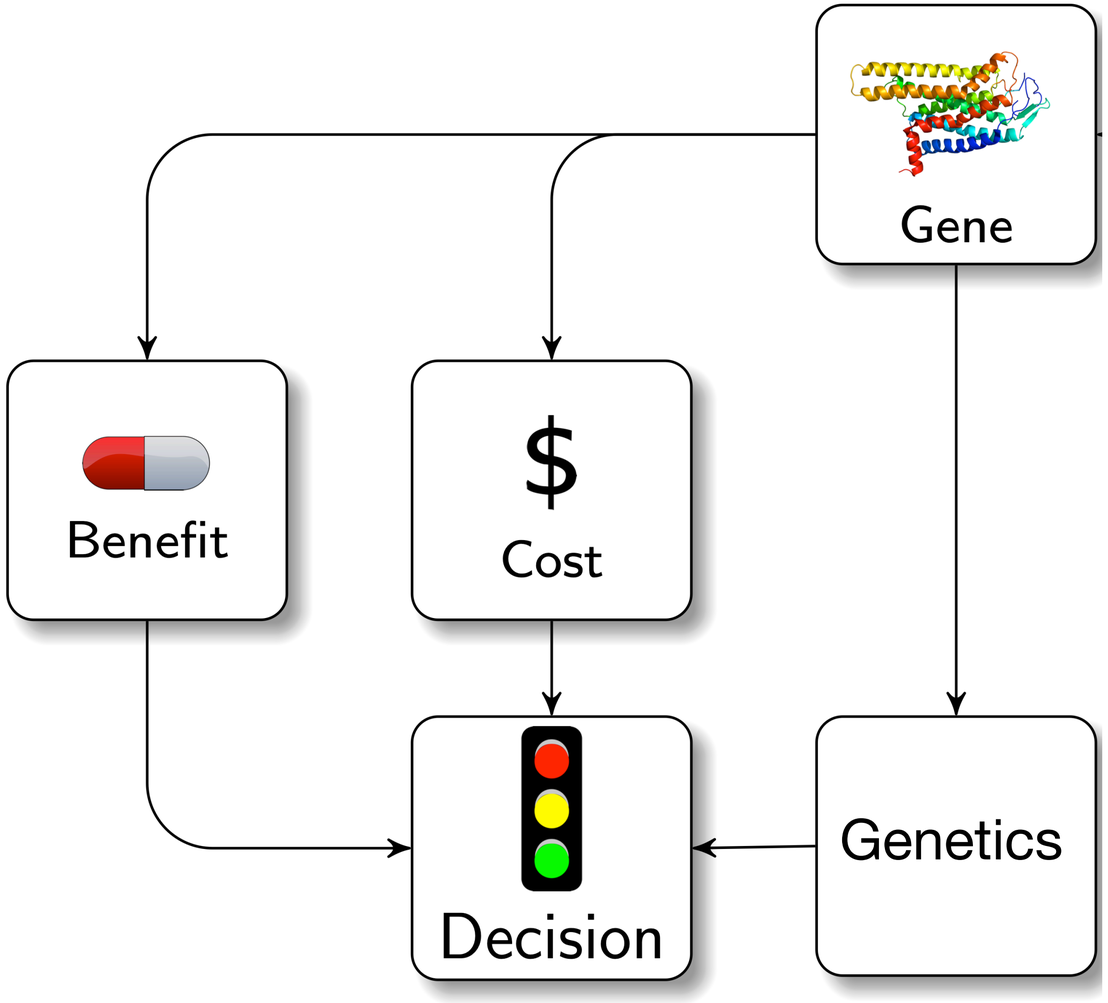


Reverse genetics

Our organizing question

What does human genetic data tell us about a gene?

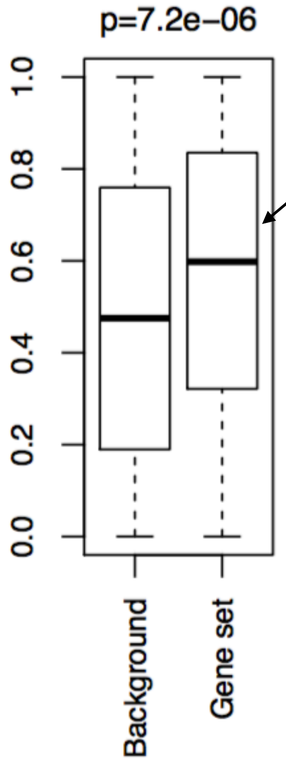
explicitly or implicitly

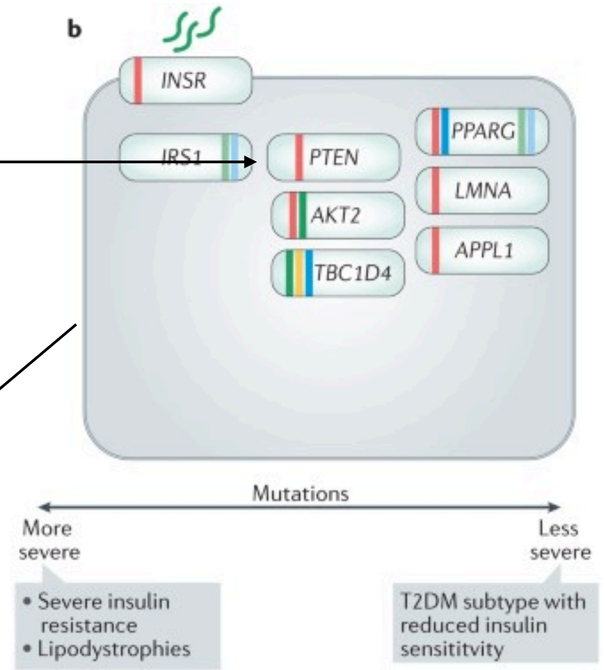



Rare variants
Common variants
JAKE-CLARK.TUMBLR
imgflip.com

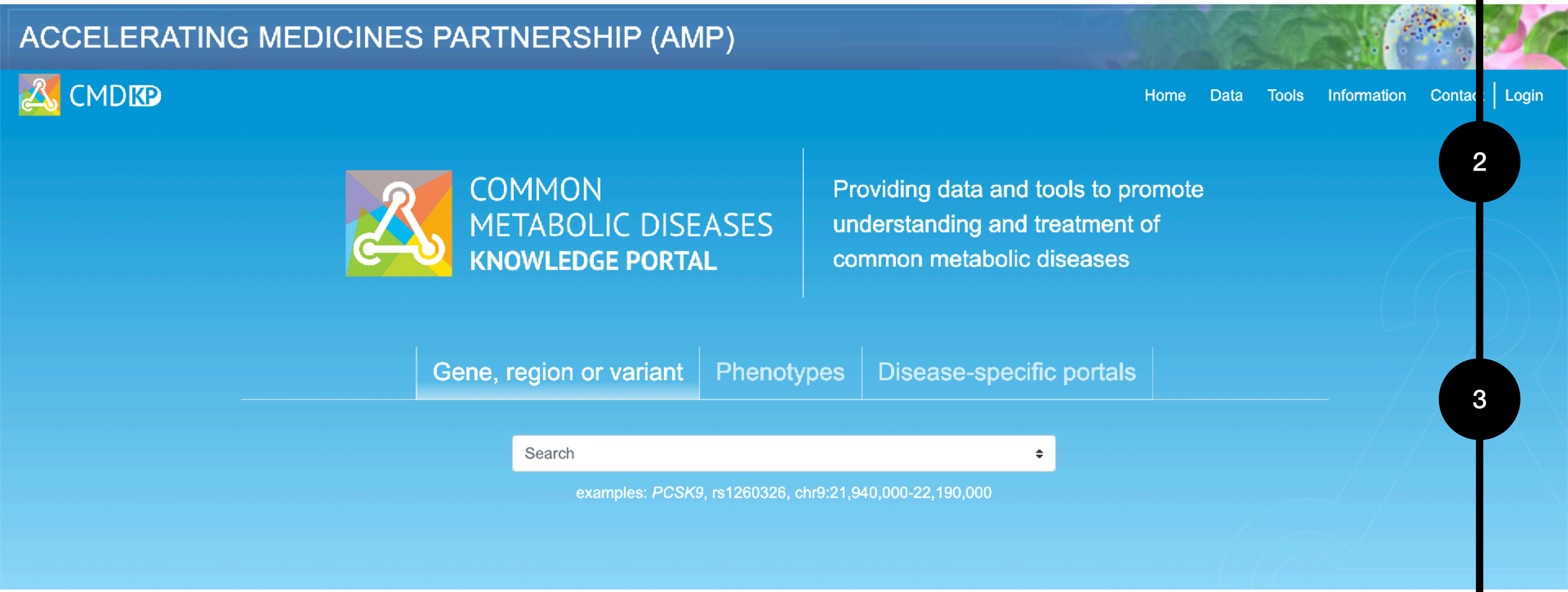

Why don't we have both?
imgflip.com

***PTEN*:**
rare p=0.048
common p=0.007





1. Make the data available



1

2

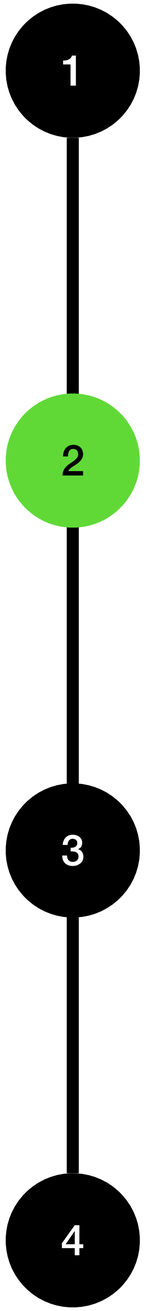
3

4

Costanzo et al, Cell Metabolism, 2023

2. Help interpret the data

Gene	GWAS	Exome
SIN3A	Minimum p=9.2e-16	p=0.59
FOXO1	Minimum p=1.91e-5	p=0.036



2. Help interpret the data

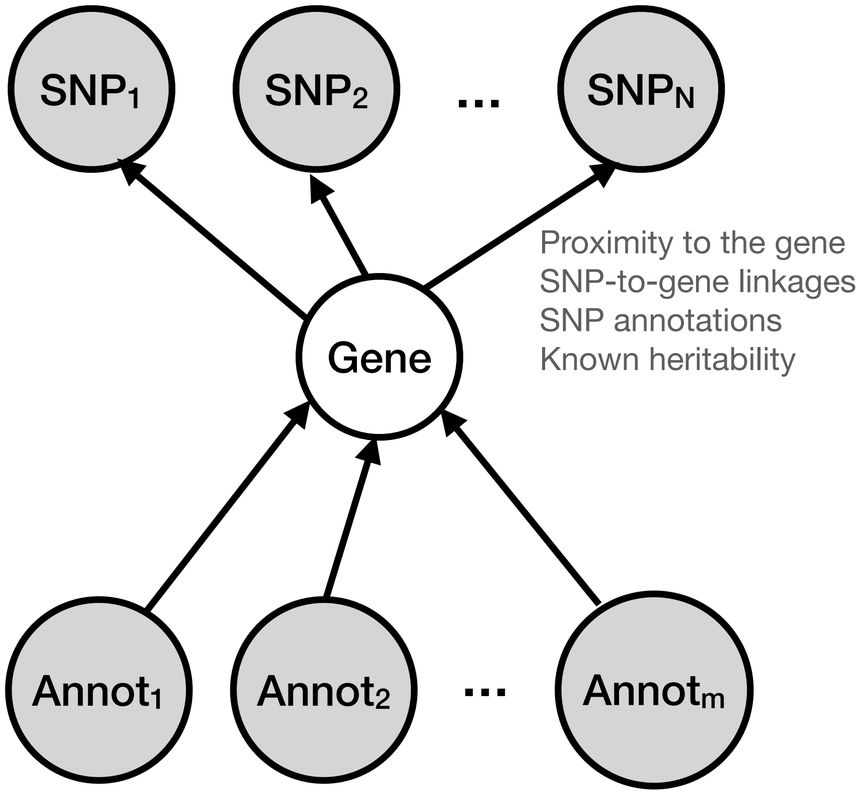
Common Variation	No evidence $p \geq 0.1$	Weak $p < 0.1$	Nominal $p < 0.05$	Strong $p < 1 \times 10^{-3}$	Exome-wide $p < 2.5 \times 10^{-6}$
Causal coding variant	**Compelling** 95% | 99%	**Compelling** 95% | 99%	**Compelling** 99% | 9%	**Compelling** 99% | 99%	**Compelling** 99% | 99%
Nearest gene	**Very Strong** 70% | 90%	**Very Strong** 80% | 95%	**Extreme** 90% | 95%	**Compelling** 99% | 99%	**Compelling** 99% | 99%
Coding variant	**Strong** 50% | 85%	**Very Strong** 60% | 90%	**Very Strong** 75% | 95%	**Compelling** 95% | 99%	**Compelling** 99% | 99%
GWAS locus	**Moderate** 15% | 40%	**Moderate** 20% | 55%	**Moderate** 30% | 70%	**Very Strong** 75% | 95%	**Compelling** 99% | 99%
No evidence	**No evidence** 5% | 20%	**Anecdotal** 5% | 25%	**Moderate** 15% | 45%	**Strong** 50% | 85%	**Compelling** 95% | 99%

Rare Variation

1
2
3
4

Dornbos et al, Cell Metabolism 2022

3. Build ever more sophisticated models

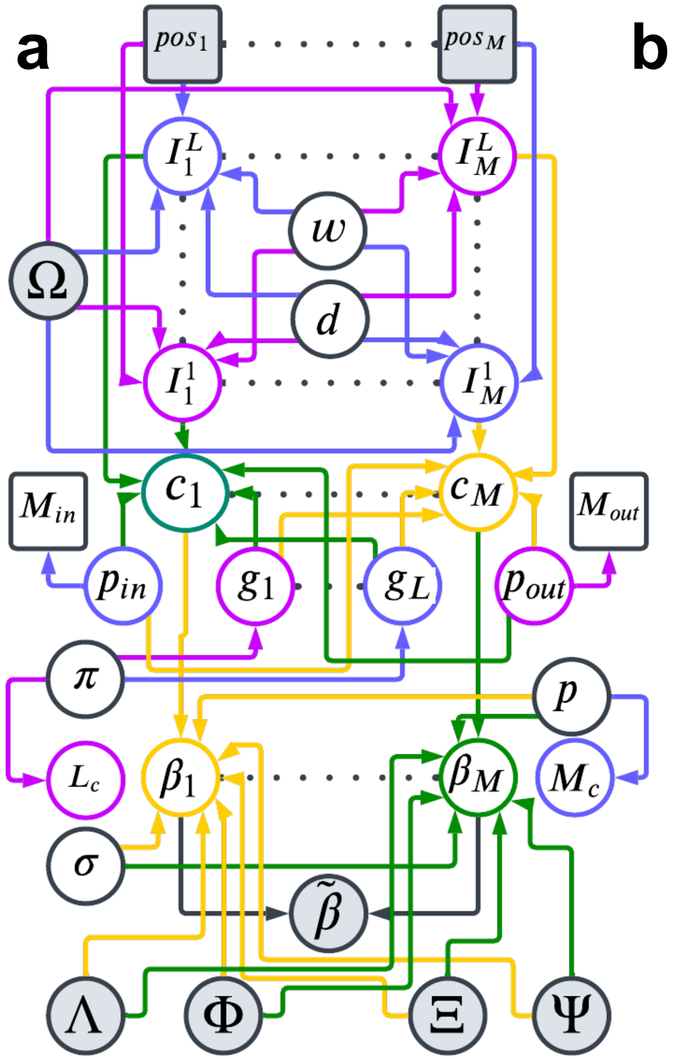


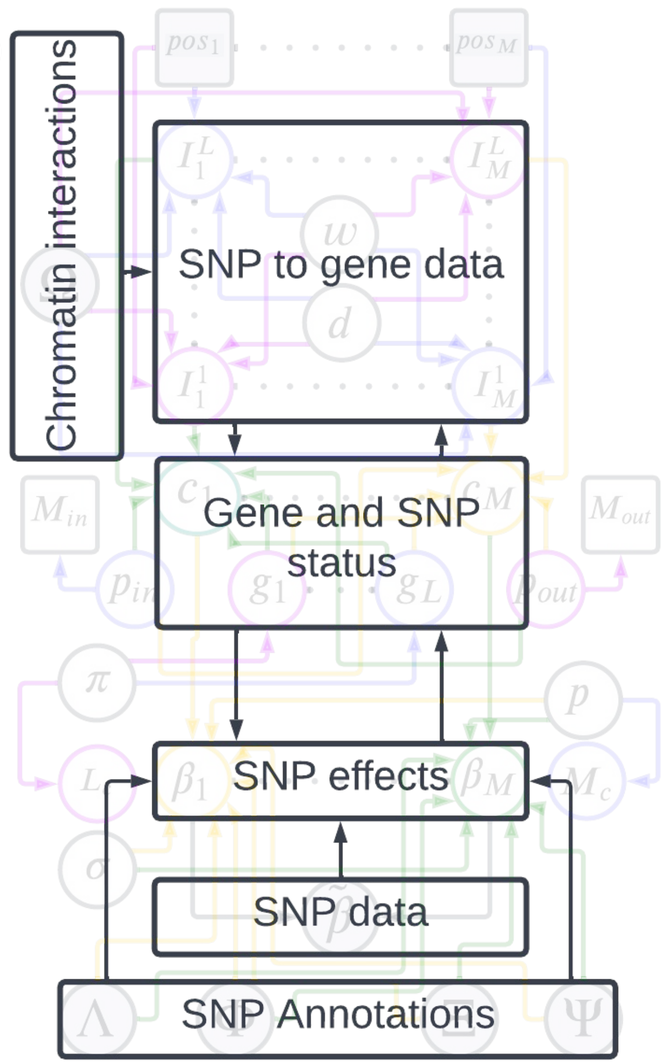
Proximity to the gene
SNP-to-gene linkages
SNP annotations
Known heritability

Expression pattern
Mouse phenotypes
Literature terms

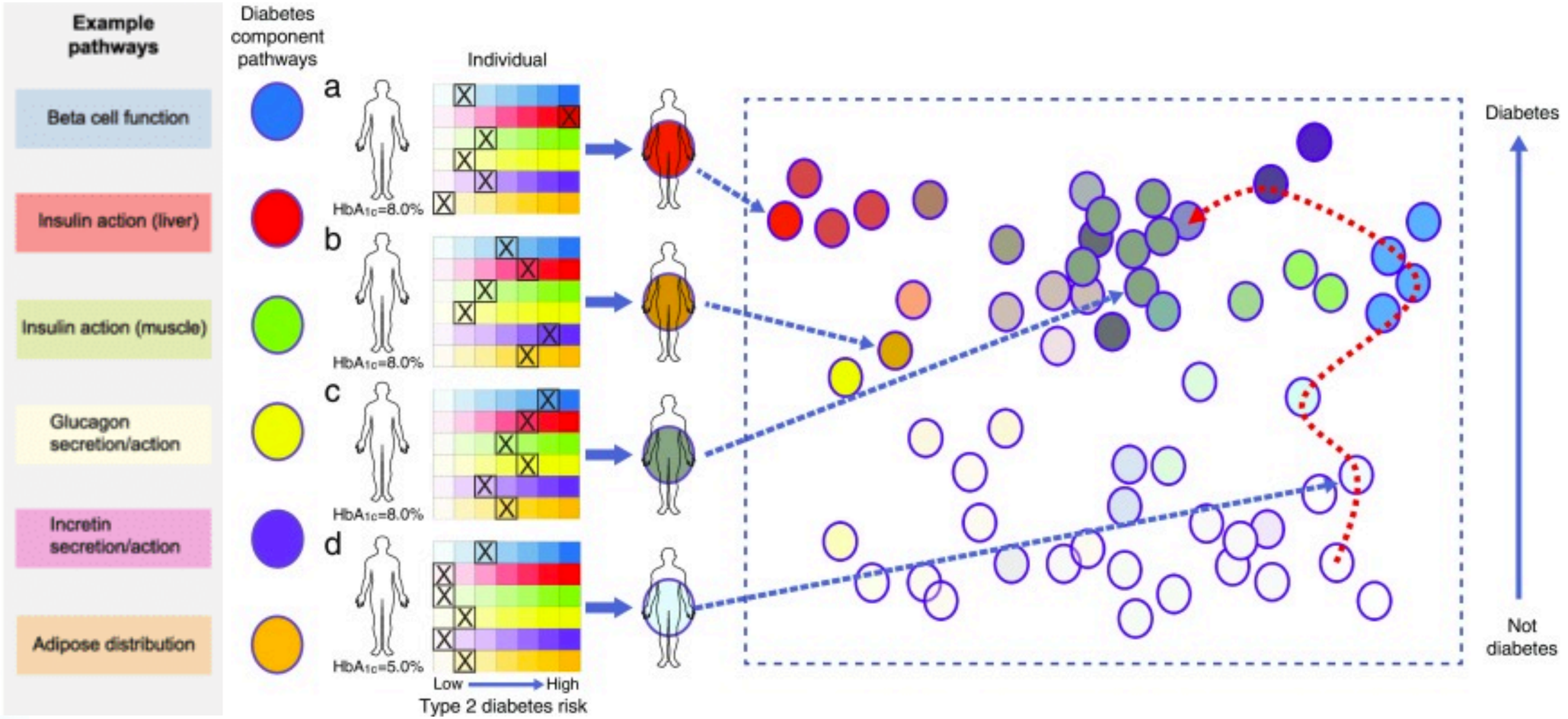
1
2
3
4

3. Build ever more sophisticated models





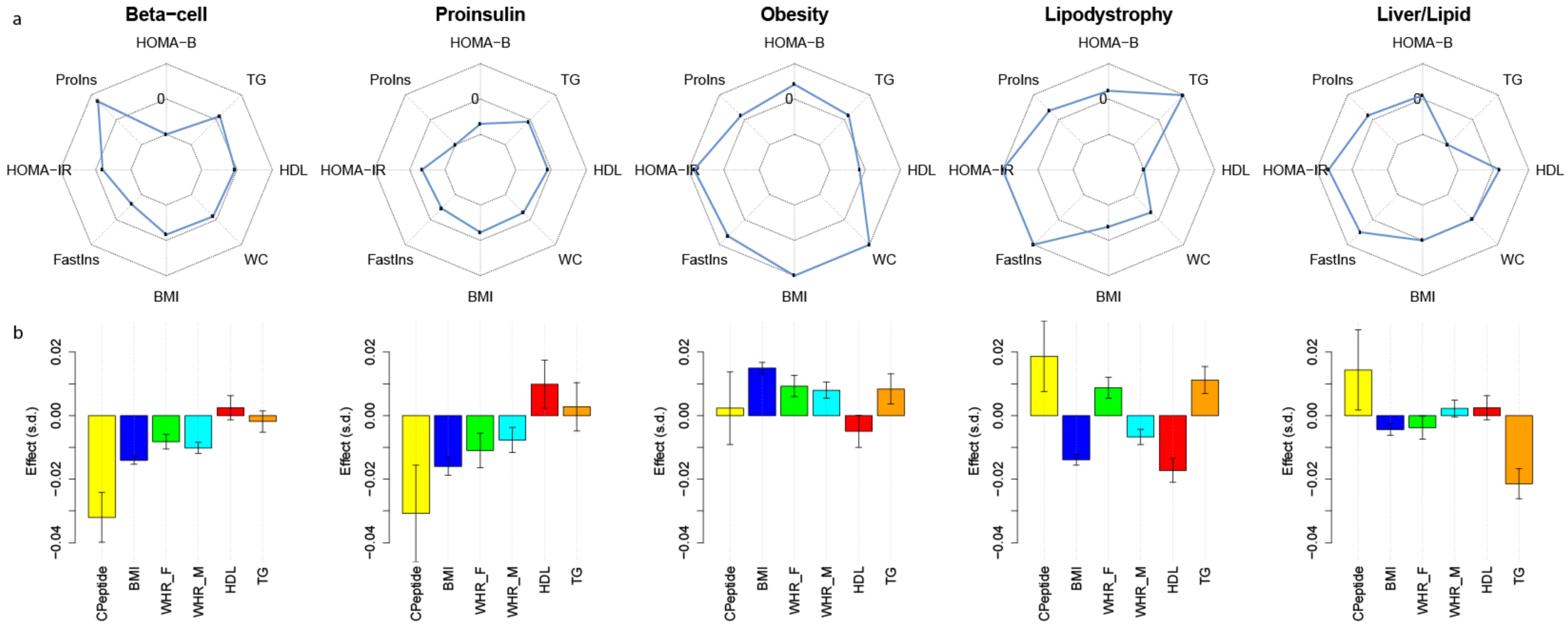
4. Extend beyond genes to pathways



McCarthy, Diabetologia, 2017

1
2
3
4

4. Extend beyond genes to pathways



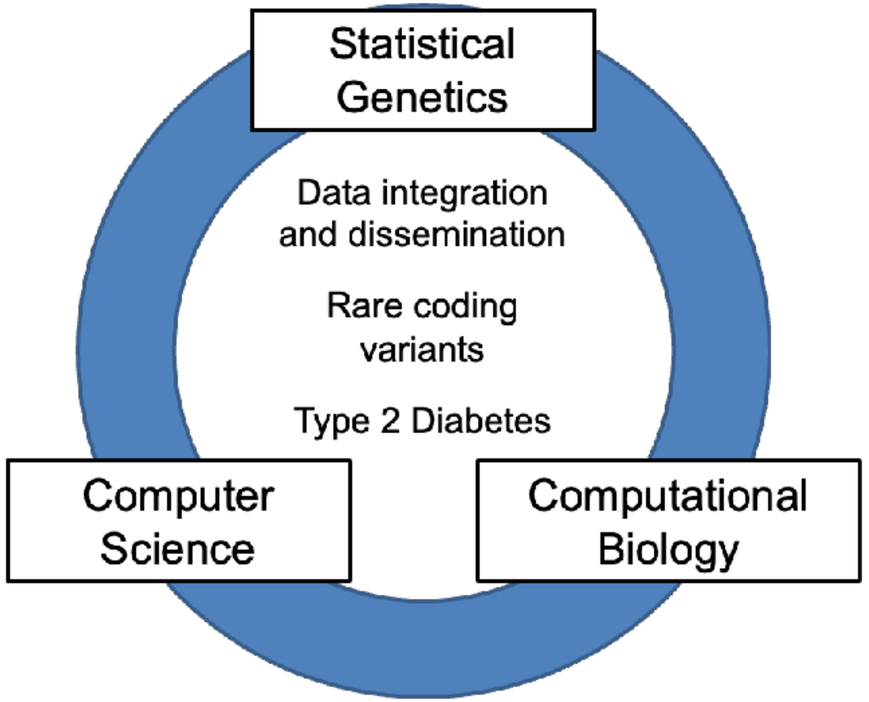
1

2

3

4

Udler et al, 2018



FlannickLab

We are always seeking collaborators and motivated new members!

Contact flannick@broadinstitute.org or http://flannicklab.org