

ARTICLE

Genetic underpinning of the comorbidity between type 2 diabetes and osteoarthritis

Ana Luiza Arruda,^{1,2,4} April Hartley,⁵ Georgia Katsoula,^{1,4} George Davey Smith,⁵ Andrew P. Morris,^{1,6} and Eleftheria Zeggini^{1,3,*}



Arruda et al., 2023, *The American Journal of Human Genetics* 110, 1304–1318
August 3, 2023 © 2023 The Author(s).
<https://doi.org/10.1016/j.ajhg.2023.06.010>

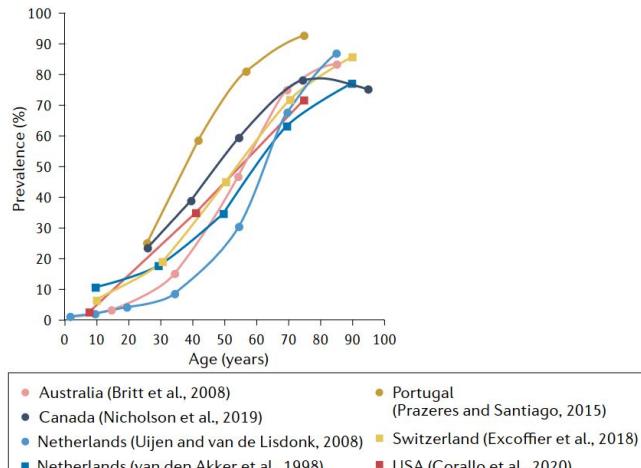


Presentation by Dylan Maher
Masters Comprehensive Exam
May 21st, 2024

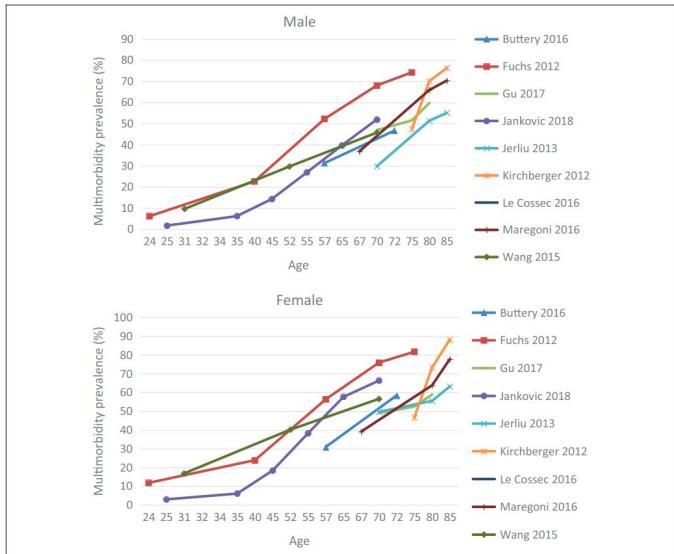
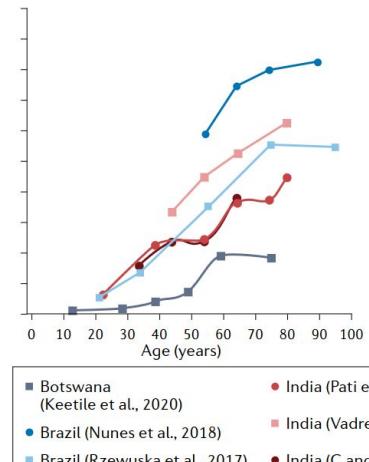


Multimorbidity

a Multimorbidity prevalence high-income countries



b Multimorbidity prevalence low-income and middle-income countries



Multimorbidity is defined as the coexistence of multiple chronic diseases in a single individual.¹ Worldwide, over 50% of the population older than 65 years is affected by more than one long-term medical condition simultaneously.² Commensurate with the rise in life expectancy

Figure 2. Age- and sex-specific prevalence of multimorbidity.

[Left] Skou et al. (2022)
[Right] Nguyen et al. (2019)



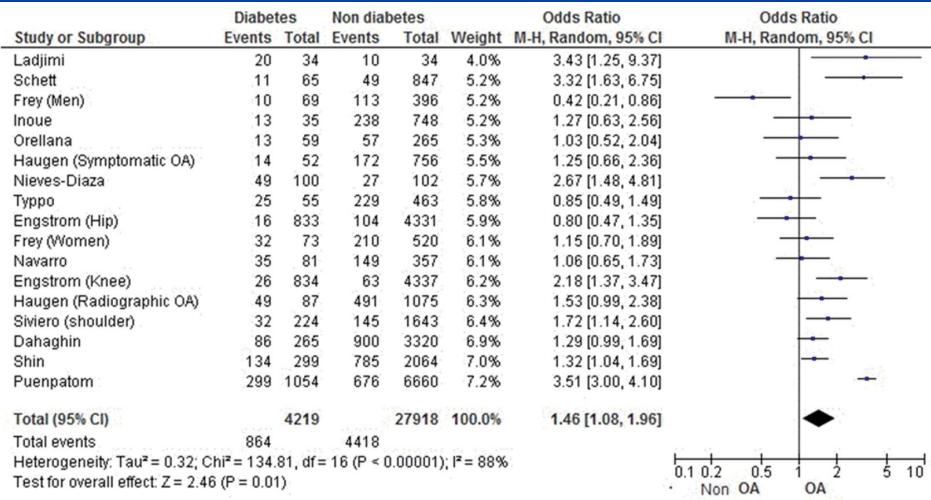
T2D & Osteoarthritis

Table 4. Prevalence of number of diseases and disease combinations among community-dwelling people aged 50 and over with multimorbidity.

Number of diseases	% (95%CI)	Main combination of chronic diseases	
Males			
2	61.4 (60.0–62.8)	1 st	Hypertension and diabetes
		2 nd	Hypertension and heart attack
		3 rd	Hypertension and osteoarthritis
		4 th	Hypertension and COPD
		5 th	Hypertension and rheumatism
3	24.8 (23.6–26.0)	1 st	Hypertension, diabetes and heart attack
		2 nd	Hypertension, diabetes and osteoarthritis
		3 rd	Hypertension, osteoarthritis and heart attack
		4 th	Hypertension, rheumatism and osteoarthritis
		5 th	Hypertension, heart attack and COPD
4	9.4 (8.5–10.4)	1 st	Hypertension, heart attack, diabetes and osteoarthritis
		2 nd	Hypertension, heart attack, stroke and diabetes
		3 rd	Hypertension, heart attack, diabetes and COPD
		4 th	Hypertension, diabetes, rheumatism and osteoarthritis
		5 th	Hypertension, heart attack, rheumatism and osteoarthritis
5 or more	4.4 (3.9–5.0)	1 st	Hypertension, diabetes, heart attack, cancer and osteoarthritis
		2 nd	Hypertension, diabetes, heart attack, osteoarthritis and COPD
		3 rd	Hypertension, heart attack, diabetes, rheumatism and osteoarthritis
		4 th	Hypertension, heart attack, rheumatism, osteoarthritis and COPD
		5 th	Hypertension, diabetes, heart attack, stroke and osteoarthritis
Females			
2	54.6 (53.5–55.7)	1 st	Hypertension and osteoarthritis
		2 nd	Hypertension and diabetes
		3 rd	Hypertension and rheumatism
		4 th	Hypertension and heart attack
		5 th	Hypertension and emotional disorder
3	27.5 (26.6–28.5)	1 st	Hypertension, diabetes and osteoarthritis
		2 nd	Hypertension, rheumatism and osteoarthritis
		3 rd	Hypertension, heart attack and osteoarthritis
		4 th	Hypertension, osteoarthritis and emotional disorder
		5 th	Hypertension, rheumatism and diabetes
4	11.7 (11.0–12.5)	1 st	Hypertension, diabetes, rheumatism and osteoarthritis
		2 nd	Hypertension, rheumatism, osteoarthritis and heart attack
		3 rd	Hypertension, rheumatism, osteoarthritis and emotional disorder
		4 th	Hypertension, diabetes, osteoarthritis and heart attack
		5 th	Hypertension, osteoarthritis, heart attack and emotional disorder
5 or more	6.2 (5.6–6.8)	1 st	Hypertension, diabetes, heart attack, rheumatism and osteoarthritis
		2 nd	Hypertension, heart attack, rheumatism, arthritis and COPD
		3 rd	Hypertension, diabetes, osteoarthritis, rheumatism and COPD
		4 th	Hypertension, diabetes, rheumatism, emotional disorder and COPD
		5 th	Hypertension, diabetes, osteoarthritis, rheumatism and emotional disorder



Osteoarthritis with/without T2D

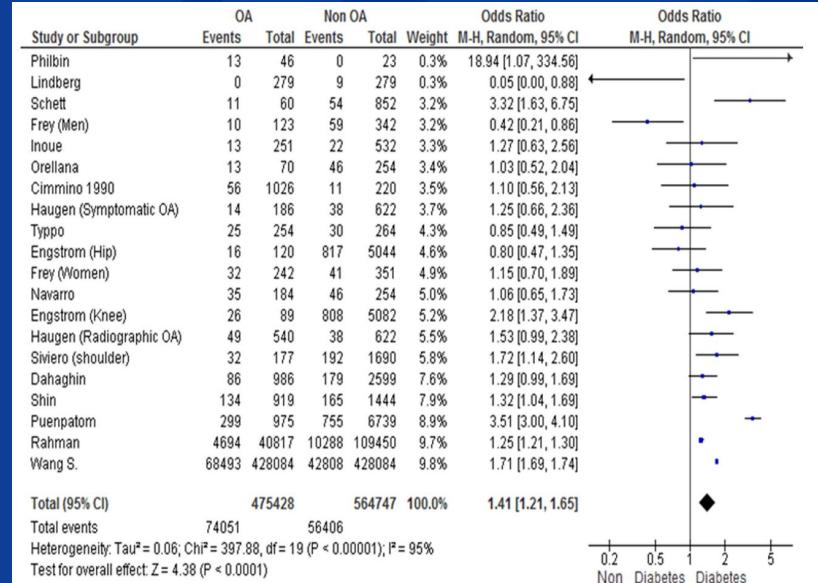


$OR_{RE} = 1.46(1.06, 1.56)$
 $I^2 = 88\%$

Louati et al. (2015)

T2D & Osteoarthritis

T2D with/without Osteoarthritis



$OR_{RE} = 1.41(1.21, 1.65)$
 $I^2 = 95\%$



T2D & Osteoarthritis

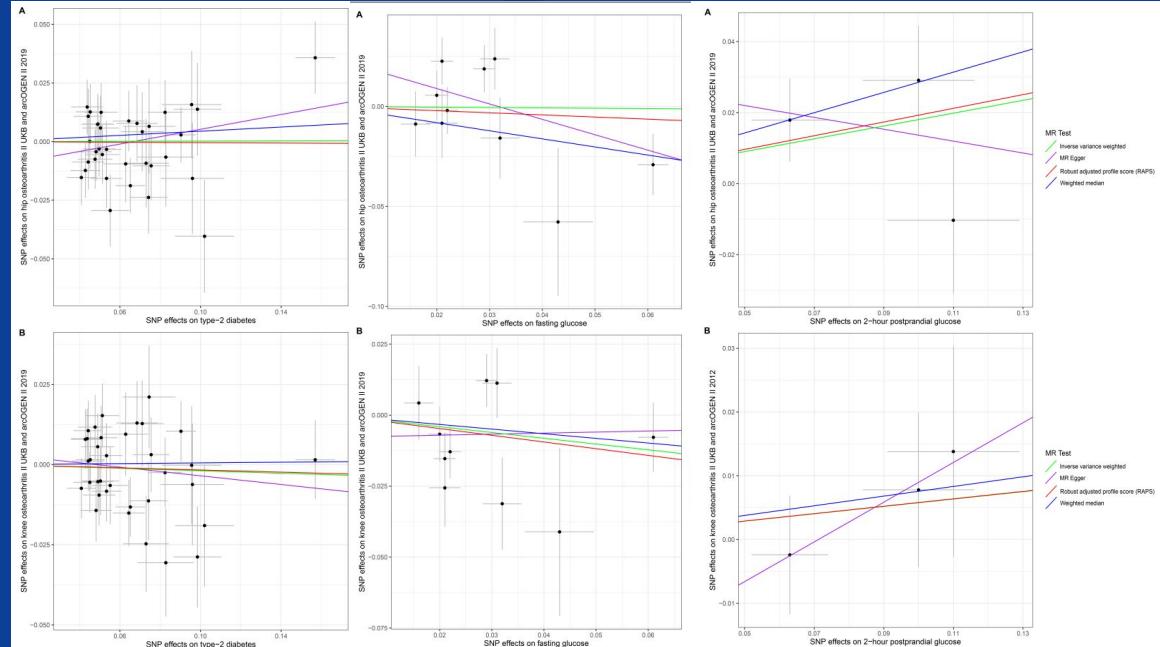
(Previous Mendelian Randomization studies)

Type 2 Diabetes and Glycemic Traits
Are Not Causal Factors of
Osteoarthritis: A Two-Sample
Mendelian Randomization Analysis



(No arrow means no causal relation)

Top panel: Hip OA
Bottom panel: Knee OA



OA ~ T2D

OA ~ FG

OA ~ 2HPPG

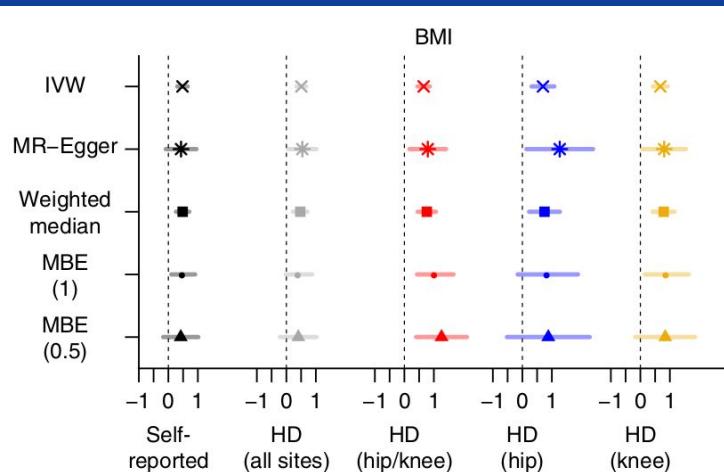


University of
Pittsburgh®

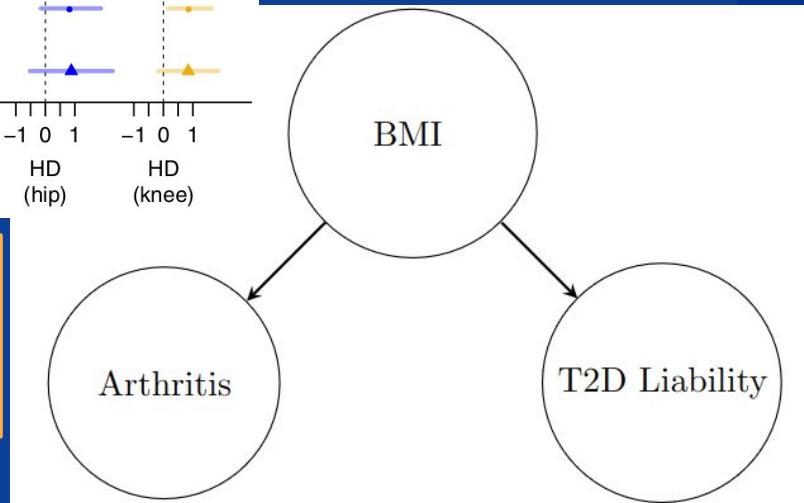
School of
Public Health

T2D & Osteoarthritis

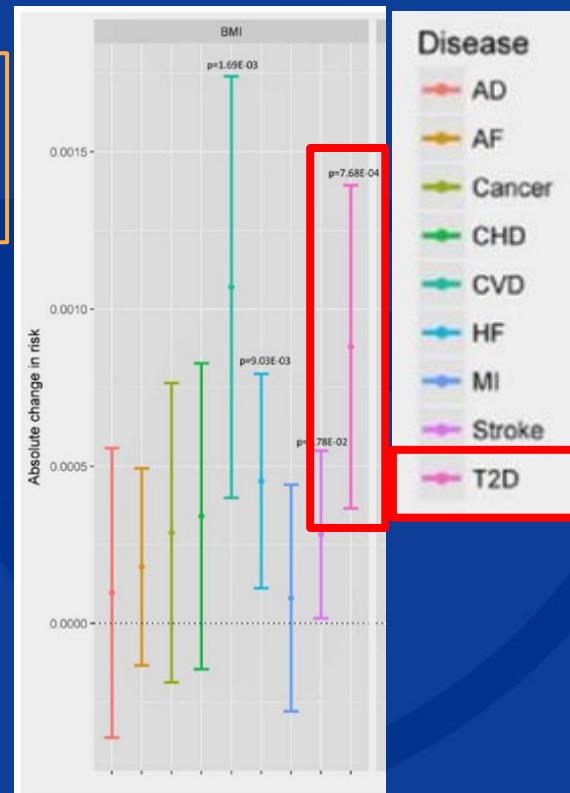
(Previous Mendelian Randomization studies)



Zengini et al.
(2018)



He et al.
(2017)

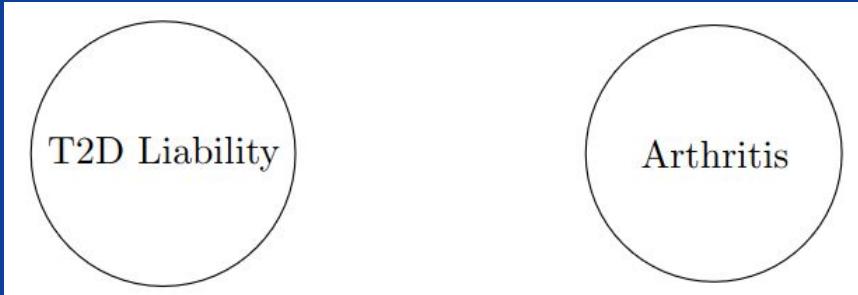




T2D & Osteoarthritis

(Previous Mendelian Randomization studies)

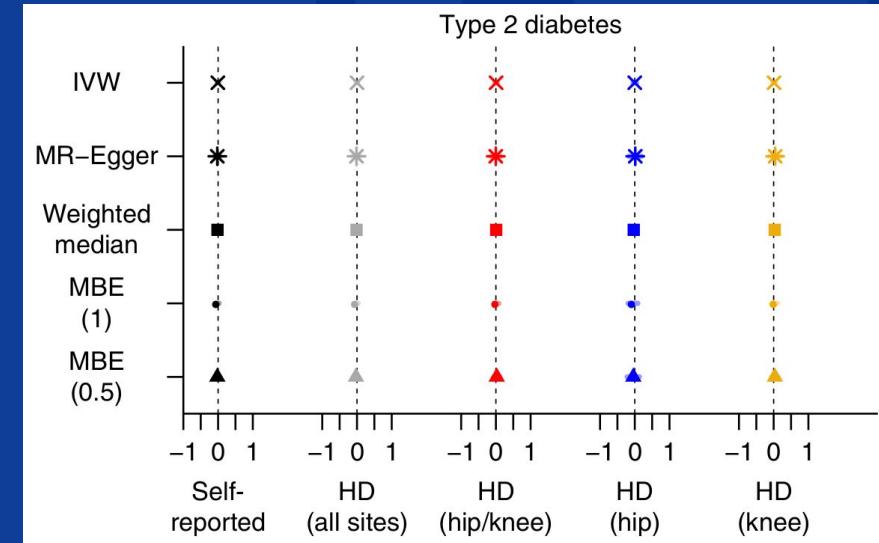
(No arrow means no causal relation)



Genome-wide analyses using UK Biobank data provide insights into the genetic architecture of osteoarthritis

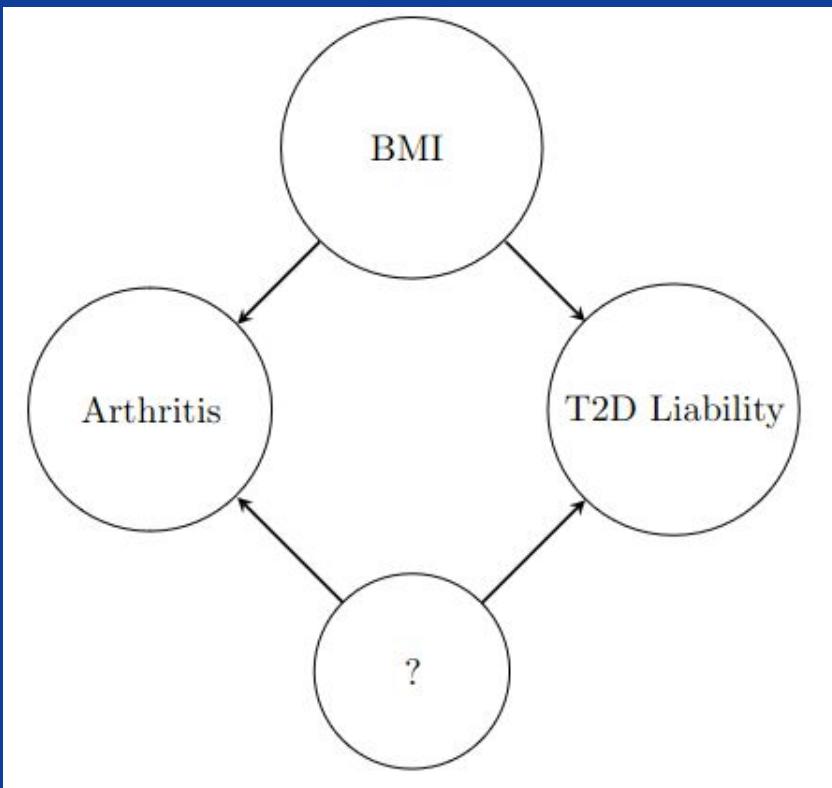
Zengini et al. (2021)

endophenotypes, and in five signals we identified genes that were differentially expressed in degraded compared with intact articular cartilage from patients with osteoarthritis. We established causal effects on osteoarthritis for higher body mass index but not for triglyceride levels or genetic predisposition to type 2 diabetes.





T2D & OA



Open Questions & Implications

global health challenge. However, the majority of health and drug development research is focused on treating and/or preventing individual diseases, leading to interventions that are currently not optimally designed to assist individuals suffering from multiple health conditions.

have yielded conflicting results.^{4,9,14} Considering that obesity is a major risk factor for both diseases studied here, genetic variants associated with different physiological characteristics of increased adiposity are expected to be shared risk variants for the comorbidity. However, those variants could exert their effects on the comorbidity through alternative biological pathways to obesity through horizontal pleiotropy.¹⁰

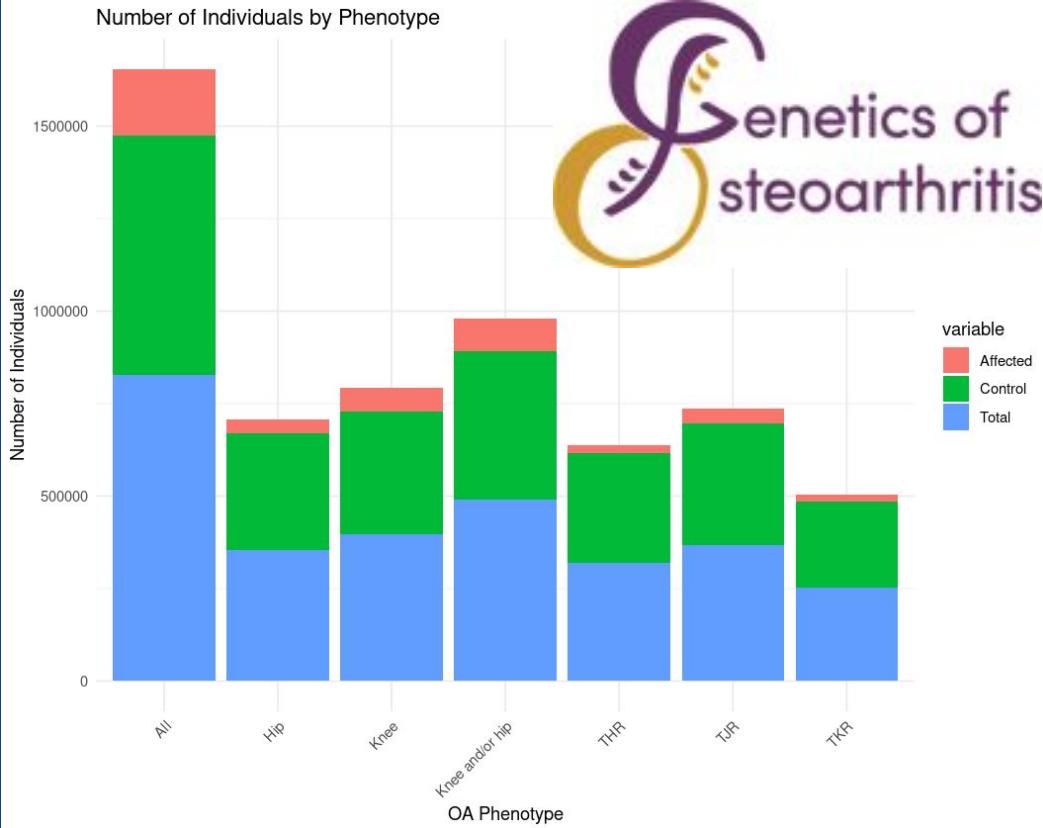


University of
Pittsburgh®

School of
Public Health

Main Study

(Osteoarthritis Sample)



Phenotype	Affected	Control	Total	Affected/Control Ratio
All	177517	649173	826690	0.27
Knee	62497	333557	396054	0.19
Knee and/or hip	89741	400604	490345	0.22
Hip	36445	316943	353388	0.11
TKR	18200	233841	252041	0.08
TJR	40887	327689	368576	0.12
THR	23021	296016	319037	0.08

Genetics of Osteoarthritis Consortium



University of
Pittsburgh®

School of
Public Health



DIAGRAM

DIAMANTE

DIAMANTE Consortium

- 1,159,055 controls
- 180,834 cases

Main Study

(T2D Sample)

European Subgroup

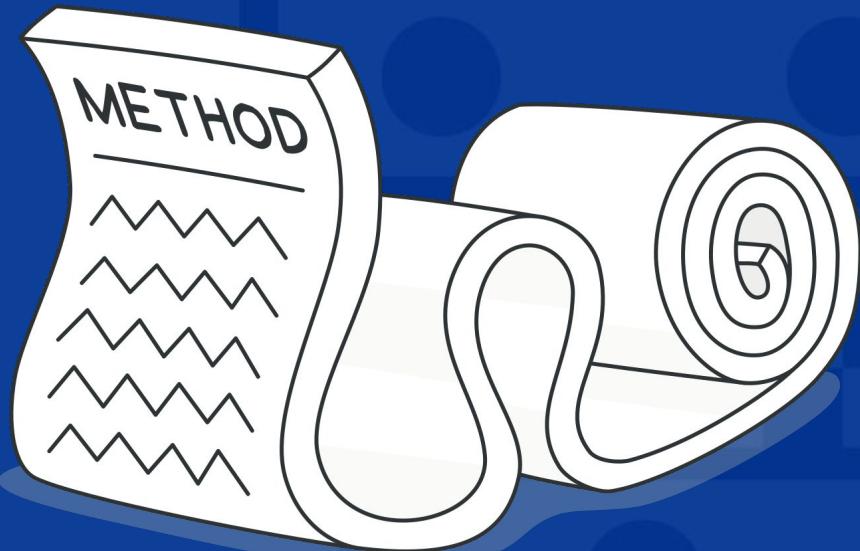
- 824,006 controls
- 74,124 cases

tal individuals for each study can be found in Table 1. For type 2 diabetes, the GWAS meta-analysis unadjusted for BMI from the DIAMANTE consortium was used.¹⁶ It includes data from 898,130 individuals (74,124 affected individuals) of European ancestry.



1. Genetic overlap of T2D and OA
2. Colocalization analysis
3. KO mouse phenotypes
4. Rare & syndromic phenotypes
5. Differential gnxp
6. Mutli-trait colocalization (eQTL & pQTL)
7. Scoring potential effector genes
8. Mutli-trait colocalization (adiposity measures)
9. Pathway analysis
10. Classification of high-confidence genes
11. “Druggable genome”
12. Causal inference analysis
13. Two-step MR
14. Tissue-specific approaches

Methods





1. Genetic overlap of T2D and OA

2. Colocalization analysis
3. KO mouse phenotypes
4. Rare & syndromic phenotypes
5. Differential gnxp
6. Mutli-trait colocalization (eQTL & pQTL)
7. Scoring potential effector genes
8. Mutli-trait colocalization (adiposity measures)
9. Pathway analysis
10. Classification of high-confidence genes
11. “Druggable genome”
12. Causal inference analysis
13. Two-step MR
14. Tissue-specific approaches

Methods

- LDSC to assess r_g
- 1KG LD Scores (EUR)
- Hypothesis testing
 - Permutation testing (10,000x)
 - Hold T2D fixed, permute OA
 - Run LDSC on each permutation



University of
Pittsburgh®

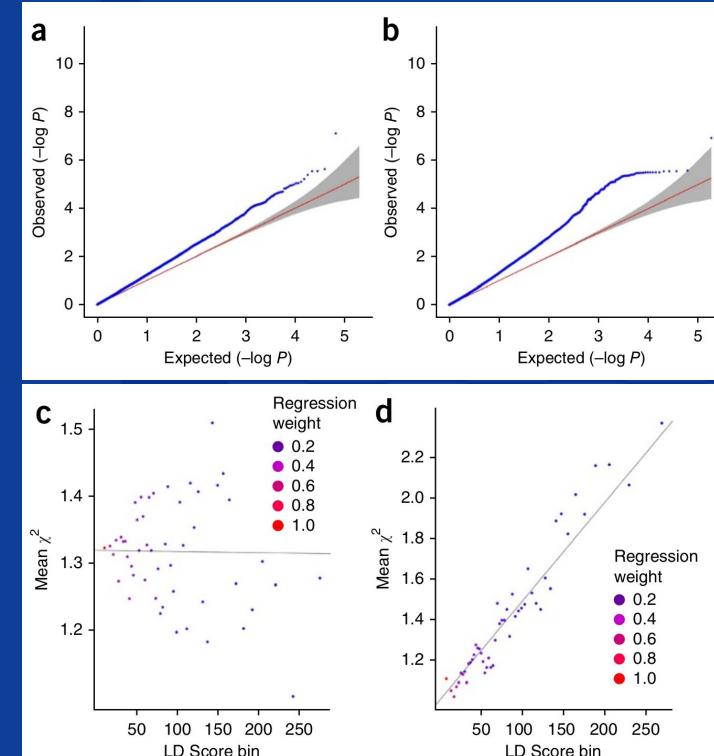
School of
Public Health

LD Score regression distinguishes confounding from polygenicity in genome-wide association studies

Brendan K Bulik-Sullivan¹⁻³, Po-Ru Loh^{1,4}, Hilary K Finucane^{4,5}, Stephan Ripke^{2,3}, Jian Yang⁶, Schizophrenia Working Group of the Psychiatric Genomics Consortium⁷, Nick Patterson¹, Mark J Daly¹⁻³, Alkes L Price^{1,4,8} & Benjamin M Neale¹⁻³

- First proposed by Bulik-Sullivan et al. (2015)
- Several uses
 - Estimate genome-wide inflation
 - Estimate heritability
 - Estimate partitioned heritability
 - **Estimate genetic correlation**
- Only requires GWAS summary statistics

LDSC



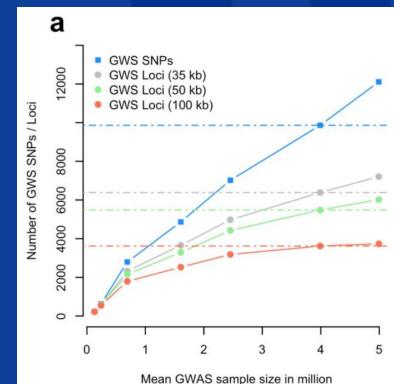
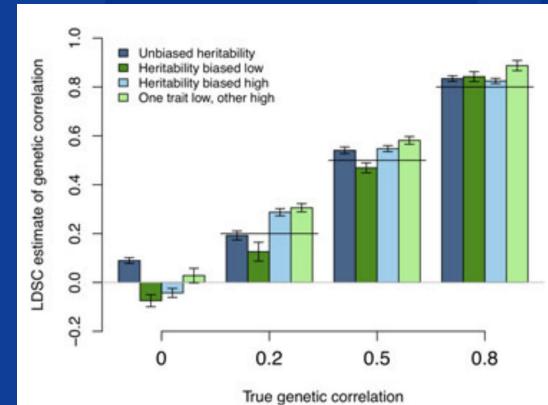


The accuracy of LD Score regression as an estimator of confounding and genetic correlations in genome-wide association studies

James J. Lee¹ | Matt McGue¹ | William G. Iacono¹ | Carson C. Chow²

In 2018, Lee et al. showed LDSC was actually more robust than had previously been assumed. They showed:

1. If the genetic effects & LD scores in blocks are independent, h^2_{LDSC} is unbiased (usually not true)
2. LDSC intercept reflects useful measure of inflation even in presence of env. confounding
3. LD Score regression provides an accurate estimate of the genetic correlation between two traits, even if neither trait's heritability is well estimated.



[Top] Lee et al. (2018)
Estimates of r_g with varying degrees of h^2 accuracy

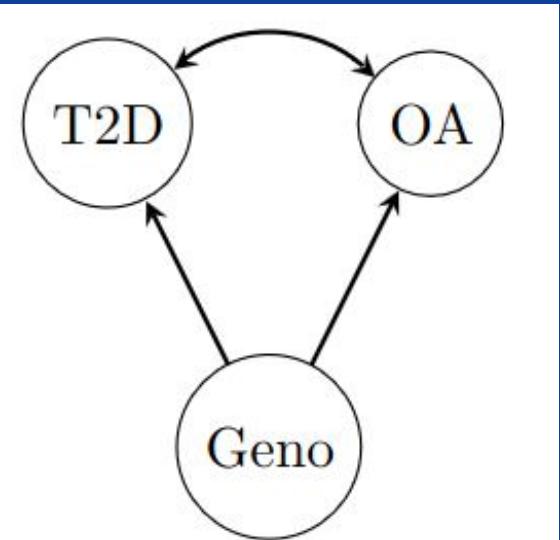
[Bottom] Yengo et al. (2022)-Supplements
Empirical estimates of SNP discovery from downsampled height GWAS



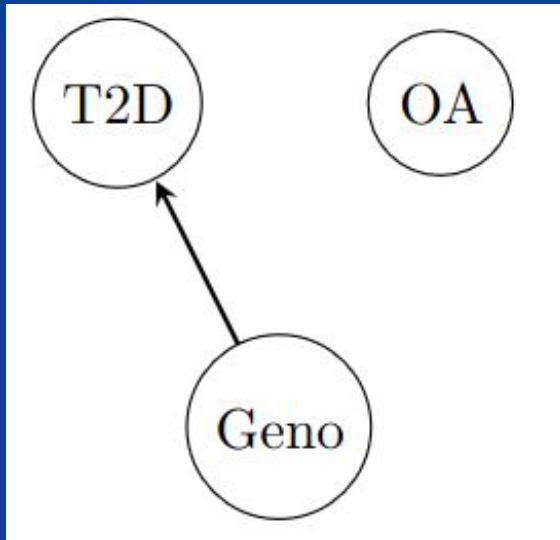
Methods

Permuting
(randomization)
breaks causal link
between geno &
OA

Provides empirical
distribution under
null (no correlation)



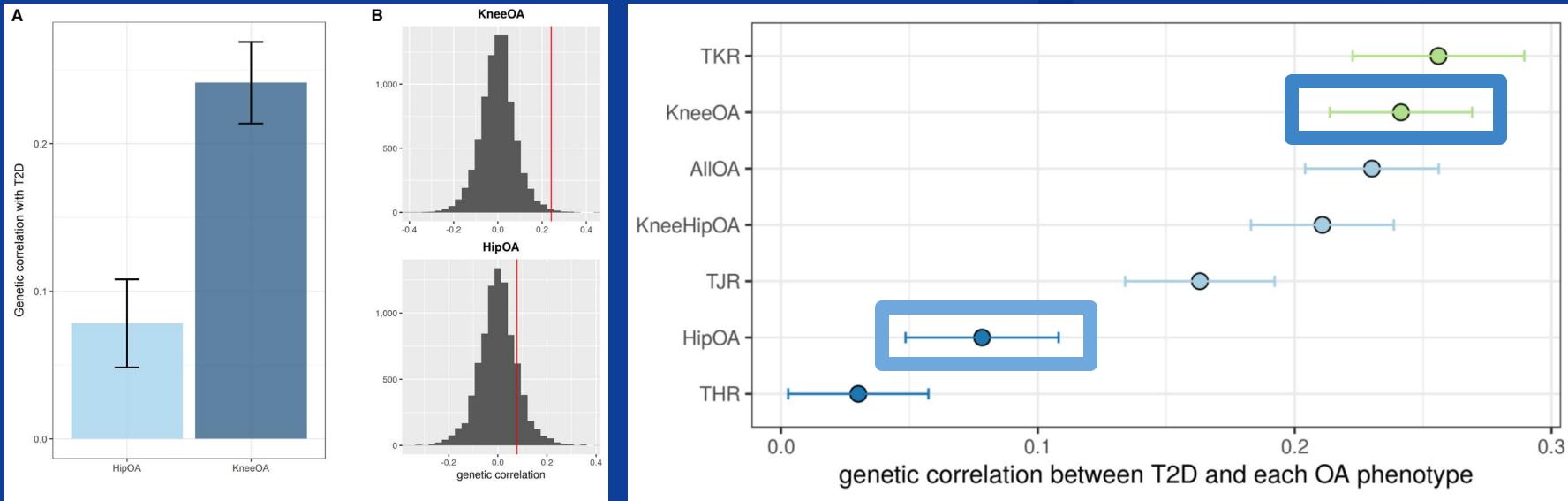
$$r_g \geq 0$$



$$r_g = 0$$

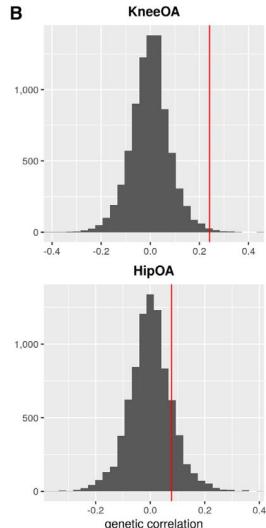
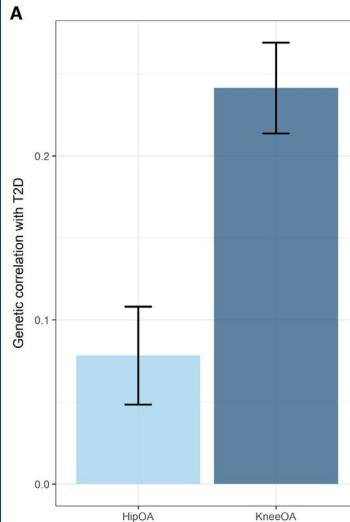


Results

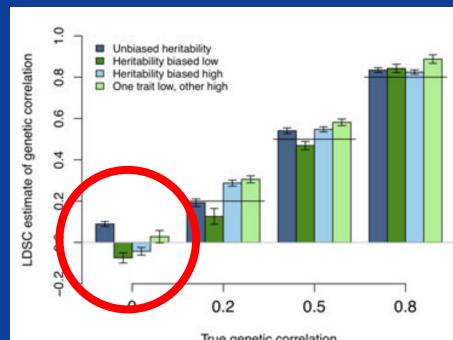

$$r_g(\text{T2D}, \text{knee}) > r_g(\text{T2D}, \text{hip})$$



Interpretation



	Knee OA	Hip OA
Sample Sizes	$N_{\text{cases}} = 62,497$ $N_{\text{controls}} = 333,557$	$N_{\text{cases}} = 36,445$ $N_{\text{controls}} = 316,943$
LDSC Estimate	$r_g = 0.241$ $SE = 0.028$ $p = 2.65e-18$	$r_g = 0.078$ $SE = 0.029$ $p = 8e-3$
Permutation testing	$p = 5e-3$	$p = 0.142$



the hip ($r_g = 0.078$, $SE = 0.029$, $p = 0.008$) (Figure 1A). To assess the potential for bias due to overlapping samples and different sample sizes, we also performed a permutation-based analysis (empirical p value for knee = 0.005, empirical p value for hip = 0.142) (Figure 1B and Table S2).

Sample Overlap: deCODE, Rotterdam, NHS, UKBB, one other (?)

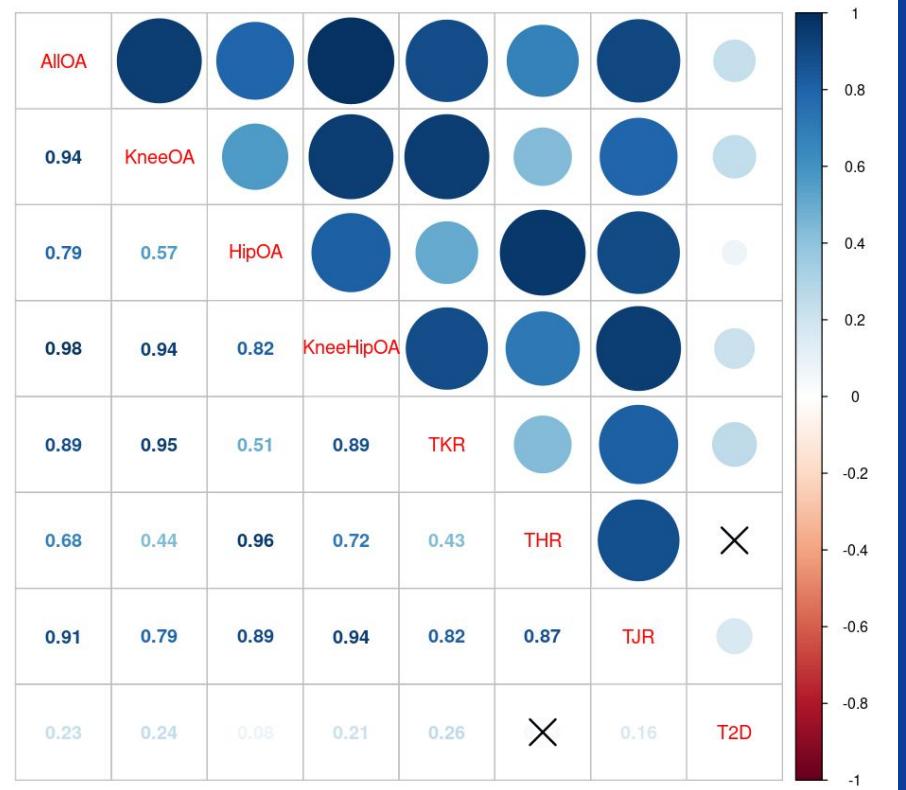
reasonable accuracy. Our mathematical analysis and simulation results suggest that the estimate should be treated with caution if it is statistically significant but nevertheless small. Our derivation in the Supporting



University of
Pittsburgh®

School of
Public Health

Interpretation



OA_phenotype	p_LDSC	p_permutation	permutation_significant
HipOA	0.008469	0.1423	No
THR	3.592e-14	0.3238	No
TKR	2.607e-18	0.0035	Yes
KneeOA	8.073e-19	0.0049	Yes
KneeHipOA	0.2689	0.0096	Yes
AllOA	1.899e-14	0.0125	Yes
TJR	1.915e-08	0.0168	Yes



1. Genetic overlap of T2D and OA

2. Colocalization analysis

3. KO mouse phenotypes

4. Rare & syndromic phenotypes

5. Differential gnxp

6. Mutli-trait colocalization (eQTL & pQTL)

7. Scoring potential effector genes

8. Mutli-trait colocalization (adiposity measures)

9. Pathway analysis

10. Classification of high-confidence genes

11. "Druggable genome"

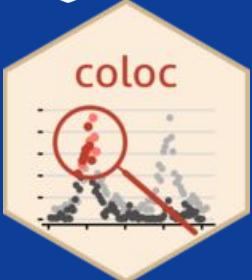
12. Causal inference analysis

13. Two-step MR

14. Tissue-specific approaches

Methods

- Define significant SNPs
 - T2D ($p < 5e-8$)
 - OA ($p < 1.3e-8$)
- Consider regions around these
 - 2 Mb (+/- 1Mb centered on SNP)
- Run colocalization analysis (coloc R package)
- Evidence of colocalization: $H4$ (PP4) > 0.8



Prior justification:

$p_1 = p_2 = 1e-4$ broadly correspond to 99% “belief” that $p < 5e-8$ in a GWAS corresponds to a “true” association

p_{12} is trickier, some intuition provided by ShinyApp
(<https://chr1swallace.shinyapps.io/coloc-priors/>)

(Approximate) Bayes Factor colocalisation analyses

Introduction

The idea behind the ABF analysis is that the association of each trait with SNPs in a region may be summarised by a vector of 0s and at most a single 1, with the 1 indicating the causal SNP (so, assuming a single causal SNP for each trait). The posterior probability of each possible configuration can be calculated and so, crucially, can the posterior probabilities that the traits share their configurations. This allows us to estimate the support for the following cases:

- H_0 : neither trait has a genetic association in the region
- H_1 : only trait 1 has a genetic association in the region
- H_2 : only trait 2 has a genetic association in the region
- H_3 : both traits are associated, but with different causal variants
- H_4 : both traits are associated and share a single causal variant

https://chr1swallace.github.io/coloc/articles/a03_enumeration.html

Methods

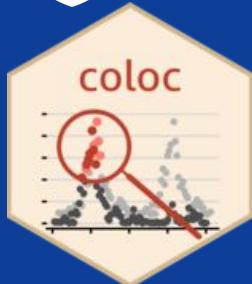
Wallace (2020)

hyp	configuration	num	prior
H_0	○—○—○—○—○	$\times 1$	
H_1	{ ○—○—○—○—○ } ○—○—○—○—○ } ... } $\times n$		p_1 p_1
H_2	{ ○—○—○—○—○ } ○—○—○—○—○ } ... } $\times n$		p_2 p_2
H_3	{ ○—○—○—○—○ } ○—○—○—○—○ } ... } $\times n(n - 1)$		$p_1 p_2$ $p_1 p_2$
H_4	{ ○—○—○—○—○ } ○—○—○—○—○ } ... } $\times n$		p_{12} p_{12}



University of
Pittsburgh®

School of
Public Health



Shiny

Prior justification:

$p_1 = p_2 = 1e-4$ broadly correspond to 99% “belief” that $p < 5e-8$ in a GWAS corresponds to a “true” association

p_{12} is trickier, some intuition provided by ShinyApp

Prior explorer for coloc

Input: Type of Per SNP priors

select one

Raw

Marginal/Conditional

Input: Parameter values

Number of SNPs in region

p_{12}

p_1

p_2

nsnps > 0
 $p_1 * p_2 < p_{12} < \min(p_1, p_2)$
 $0 < p_1 < 1/nsnps$
 $0 < p_2 < 1/nsnps$

Per SNP priors - raw

- p_{12} - Prior probability a random SNP in the region is jointly causal for both traits
- p_1 - Prior probability a random SNP in the region is causally associated with trait 1 and not trait 2
- p_2 - Prior probability a random SNP in the region is causally associated with trait 2 and not trait 1

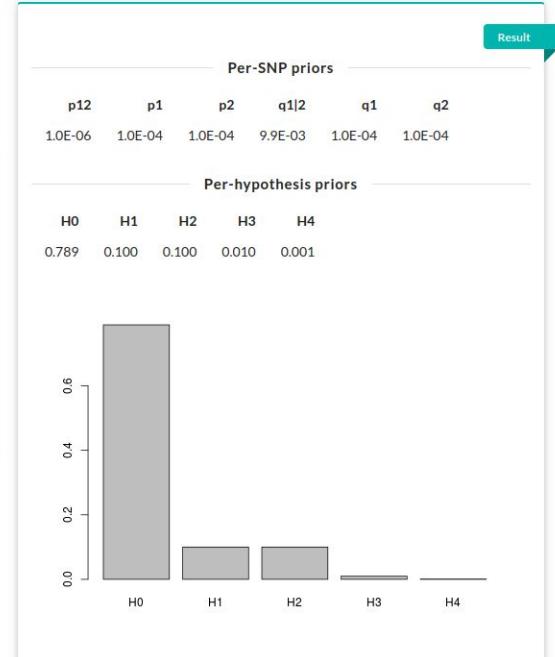
Per SNP priors - marginal and conditional

- p_{12} - Prior probability a random SNP in the region is jointly causal for both traits given it is causal for trait 2
- p_1 - Prior probability a random SNP in the region is causally associated with trait 1 (whether or not causal for trait 2)
- p_2 - Prior probability a random SNP in the region is causally associated with trait 2 (whether or not causal for trait 1)

Per-hypothesis priors

- H0 = no association with either trait in region
- H1 = association with trait 1 in region, but not trait 2
- H2 = association with trait 2 in region, but not trait 1
- H3 = association with both traits in region, but separate causal variants
- H4 = association with both traits in region, same causal variants

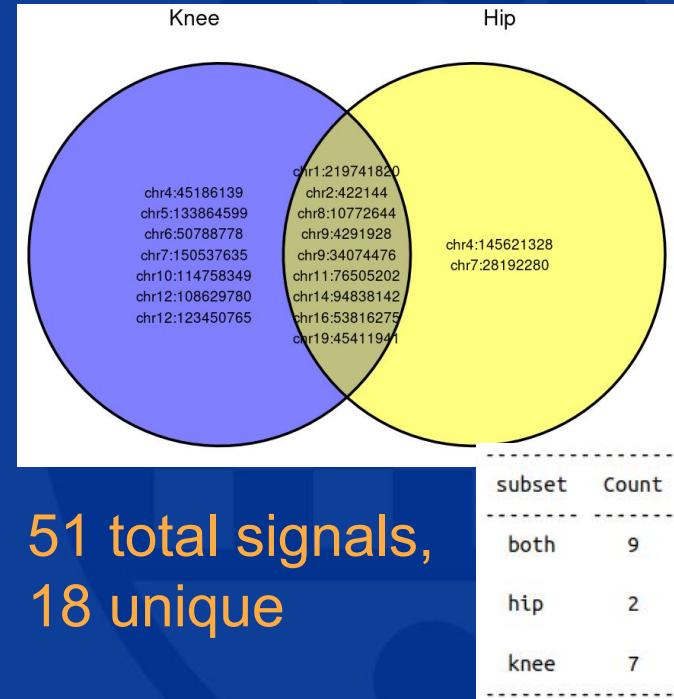
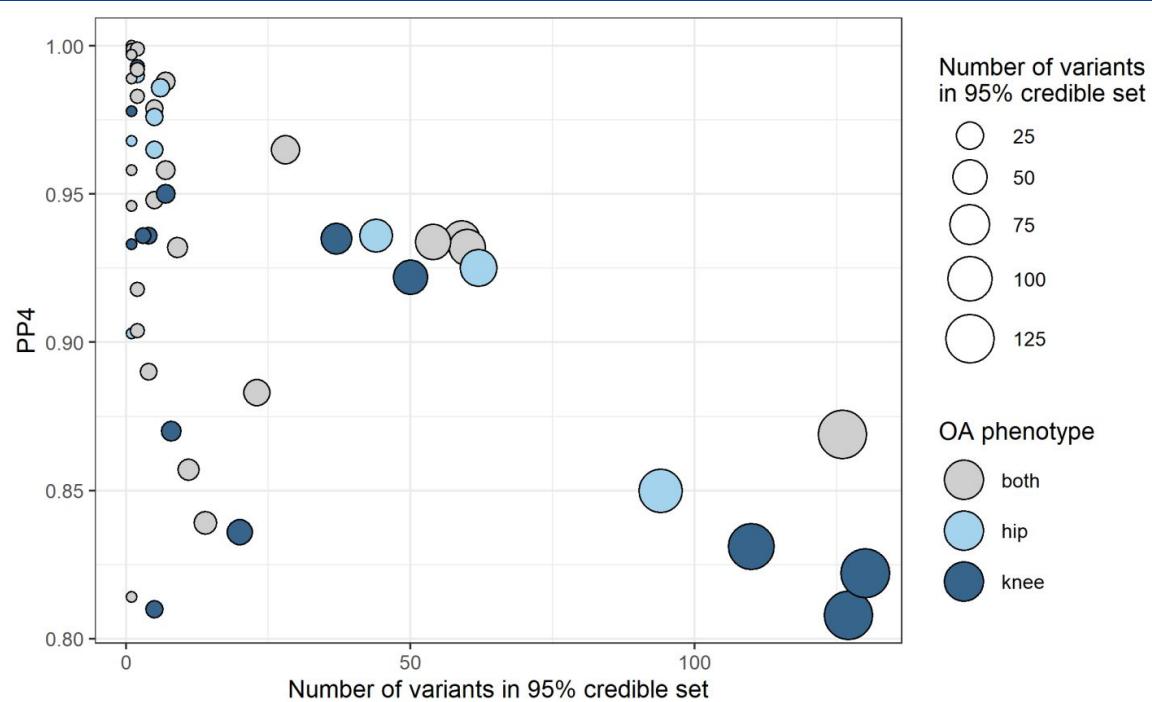
Methods



(<https://chr1swallace.shinyapps.io/coloc-priors/>)



Results



(Table S2 and Figures S5–S22). Ten of those loci colocalize with type 2 diabetes for both hip and knee osteoarthritis, two colocalize for hip osteoarthritis only, and six colocalize only for knee osteoarthritis. In three genomic loci, the 95% credible set for the causal variant from the colocalization



Recap

- ✓ Genetic overlap of T2D and OA
- ✓ Colocalization analysis

- ❑ KO mouse phenotypes
- ❑ Rare & syndromic phenotypes
- ❑ Differential gnxp
- ❑ Multi-trait colocalization (eQTL & pQTL)
- ❑ Scoring potential effector genes
- ❑ Multi-trait colocalization (adiposity measures)
- ❑ Pathway analysis
- ❑ Classification of high-confidence genes
- ❑ “Druggable genome”
- ❑ Causal inference analysis
- ❑ Two-step MR
- ❑ Tissue-specific approaches

What have we done so far?

- Confirmed genetic correlation of T2D & OA
 - Knee (robust)
 - Hip (less robust)
- Used GWAS from OA (various sub-types) and T2D as input for colocalization
 - 51 total signals
 - 18 unique signals

**Q: What does that leave us with?
A: About 36 Mb**

(or, about 1.125% of the genome)



36 Mb contain 906 candidate genes

How to prioritize these?



Moving Forward

Author's Propose Six Lines of Evidence

1. KO mice
 2. OMIM
 3. DGE
 4. molQTL-colocalization
 5. molQTL-missense variants
 6. “Established” effector genes
 - a. T2D Knowledge portal
 - b. GO consortium
- } mt-coloc



1. *Genetic overlap of T2D and OA*
2. *Colocalization analysis*

3. KO mouse phenotypes

4. *Rare & syndromic phenotypes*
5. *Differential gnxp*
6. *Mutli-trait colocalization (eQTL & pQTL)*
7. *Scoring potential effector genes*
8. *Mutli-trait colocalization (adiposity measures)*
9. *Pathway analysis*
10. *Classification of high-confidence genes*
11. *“Druggable genome”*
12. *Causal inference analysis*
13. *Two-step MR*
14. *Tissue-specific approaches*

Methods

For each gene showing evidence of colocalization, schematic search was performed:

1. International Mouse Phenotyping Consortium (IMPC)
2. Mouse Genome Informatics (MGI)
3. Rat Genome Database (RGD)

Word search conducted to find KO mice with phenotypes similar to T2D & OA



University of
Pittsburgh®

School of
Public Health

Methods



IMPC

International Mouse Phenotyping Consortium



Word Search

T2D

Insulin, glucose, diabetes,
hyperglycemia, pancreas, pancreatic,
obesity, BMI, body weight, body
mass, body fat, beta cell, glucosuria

OA

Skeletal, muscle, bone, osteo,
arthritis, muscular, joint, body size,
growth, stature, height



1. *Genetic overlap of T2D and OA*
2. *Colocalization analysis*
3. *KO mouse phenotypes*

4. Rare & syndromic phenotypes

5. *Differential gnxp*
6. *Mutli-trait colocalization (eQTL & pQTL)*
7. *Scoring potential effector genes*
8. *Mutli-trait colocalization (adiposity measures)*
9. *Pathway analysis*
10. *Classification of high-confidence genes*
11. *“Druggable genome”*
12. *Causal inference analysis*
13. *Two-step MR*
14. *Tissue-specific approaches*

Methods

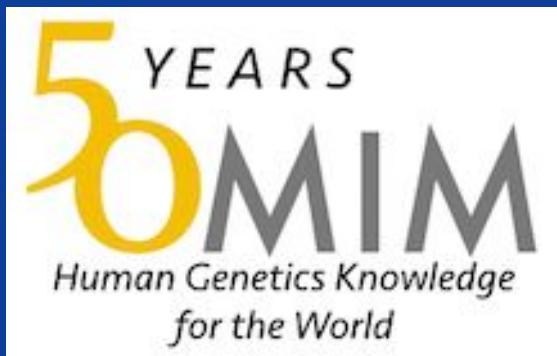
For each gene showing evidence of colocalization, schematic search was performed in OMIM (Online Mendelian Inherited Disorders in Man)

Word search conducted to find genes associated with monogenic disorders



University of
Pittsburgh®

School of
Public Health



Methods

T2D

Insulin, glycemia, glucose, diabetes,
pancreas, pancreatic, obesity, BMI,
body weight, body mass, body fat, beta
cell, glucosuria, Martsolf, aciduria,
Aicardia-Goutières, FINCA*

OA

Bone, muscle, growth skeletal,
stature, height, hand-foot-uterus,
syntosis, Martsolf, Warburg,
leukodystrophy, squalene, FINCA*

***FINCA:**
Fibrosis
Neurodegeneration
Cerebral
Angiomatosis



1. *Genetic overlap of T2D and OA*
2. *Colocalization analysis*
3. *KO mouse phenotypes*
4. *Rare & syndromic phenotypes*

5. Differential gnxp

6. *Mutli-trait colocalization (eQTL & pQTL)*
7. *Scoring potential effector genes*
8. *Mutli-trait colocalization (adiposity measures)*
9. *Pathway analysis*
10. *Classification of high-confidence genes*
11. *“Druggable genome”*
12. *Causal inference analysis*
13. *Two-step MR*
14. *Tissue-specific approaches*

Methods

OA Data

Paired vs intact OA cartilage (n=124)

T2D Data*

Pancreatic tissue from metabolically phenotyped surgery patients (n=57)

DGE: Linear model

*Covariates: age, sex, BMI
Criteria: >(+-)1.5-fold & p < 0.05

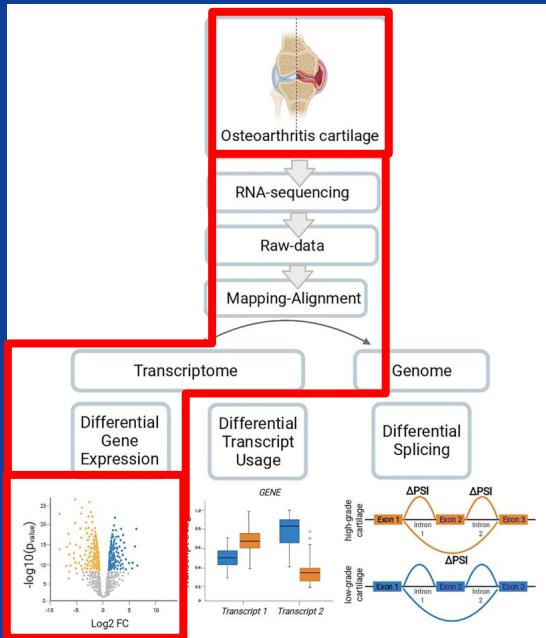


University of
Pittsburgh®

School of
Public Health

OA Data

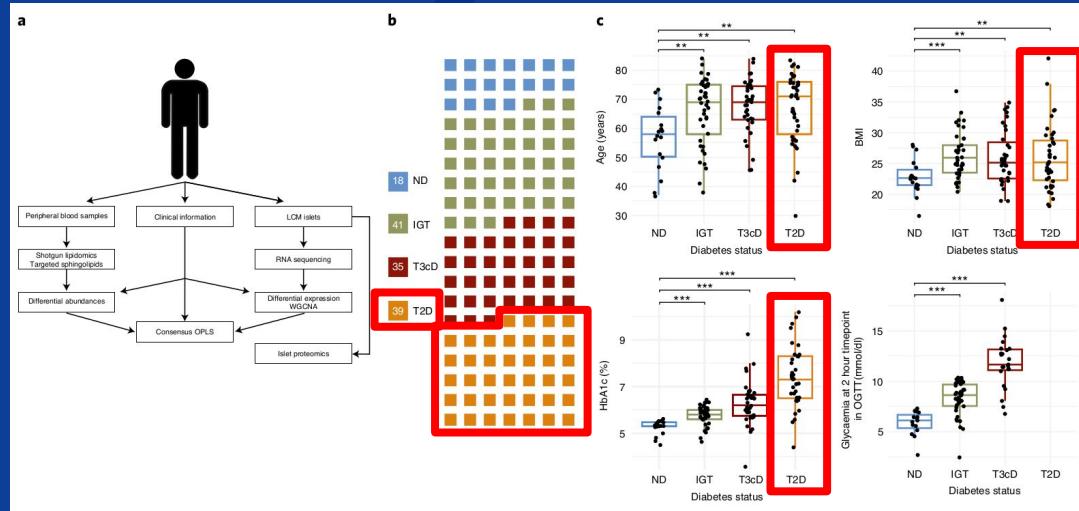
Paired vs intact OA cartilage (n=124)



Methods

T2D Data

Pancreatic tissue from metabolically phenotyped surgery patients (n=57)



OA Data: Katsoula et al. (2022)
T2D Data: Wigger et al. (2021)



OA Sample

eQTL Data

- Intact cartilage (n=95)
- Degenerated cartilage (n=87)
- Synovium (n=77)

pQTL Data

- Degenerated cartilage (n=99)

T2D Sample

eQTL Data

- Pancreatic islets
- 383 controls
- 37 cases

INspire Consortium



1. *Genetic overlap of T2D and OA*
2. *Colocalization analysis*
3. *KO mouse phenotypes*
4. *Rare & syndromic phenotypes*
5. *Differential gnxp*

6. Mutli-trait colocalization (eQTL & pQTL)

7. *Scoring potential effector genes*
8. *Mutli-trait colocalization (adiposity measures)*
9. *Pathway analysis*
10. *Classification of high-confidence genes*
11. *“Druggable genome”*
12. *Causal inference analysis*
13. *Two-step MR*
14. *Tissue-specific approaches*

Methods

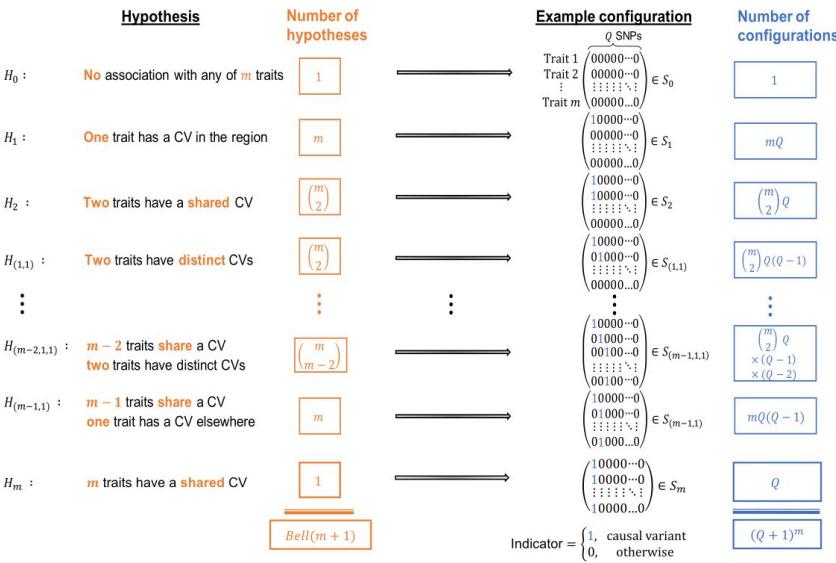
Multi-trait colocalization

1. Subset to variants in 95% credible set
2. Perform MT-molQTL-GWAS
 - a. Input (summary statistics)
 - i. T2D GWAS
 - ii. OA GWAS
 - iii. molQTL (disease/tissue specific)
 1. eQTL performed gene-wise
 2. pQTL performed protein-wise
 - b. R package *HyPrColoc*
 - i. *prior.1 = 1e-10*
 - ii. *prior.2 = 0.7*



HyPrColoc

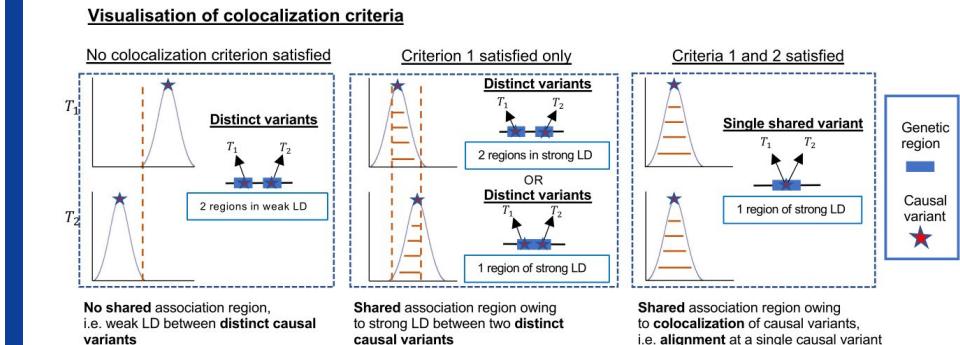
Hypothesis Prioritization
for multi-trait Colocalization
[Foley et al. 2021]



horts. Although the samples overlap, we assumed independence between the datasets, as instructed by the developers of the *HyPrColoc* package.

tal individuals for each study can be found in Table 1. For type 2 diabetes, the GWAS meta-analysis unadjusted for BMI from the DIAMANTE consortium was used.¹⁶ It includes data from 898,130 individuals (74,124 affected individuals) of European ancestry.

Methods



Outline of the main HyPrColoc approximation





University of
Pittsburgh®

School of
Public Health



TYPE 2 DIABETES KNOWLEDGE PORTAL

Genes “high confidence” if score ≥ 4

Methods



Genes “high confidence” if score labeled so

If genes were labeled “high confidence” as defined by T2D Knowledge Portal or GO consortium, but failed other analyses, “score” was updated to 1



1. *Genetic overlap of T2D and OA*
 2. *Colocalization analysis*
 3. *KO mouse phenotypes*
 4. *Rare & syndromic phenotypes*
 5. *Differential gnxp*
 6. *Mutli-trait colocalization (eQTL & pQTL)*
- ## 7. Scoring potential effector genes
8. *Mutli-trait colocalization (adiposity measures)*
 9. *Pathway analysis*
 10. *Classification of high-confidence genes*
 11. *"Druggable genome"*
 12. *Causal inference analysis*
 13. *Two-step MR*
 14. *Tissue-specific approaches*

Methods

Scoring potential effector genes

1. Subset to loci showing evidence of colocalization between T2D & ≥ 1 OA phenotype ($PP4 > 0.8$)
2. Build 1-Mb window around lead variant of credible set
3. Score each disease
 - a. Incorporate prior information
 - i. T2D Knowledge portal
 - ii. GO consortium
 - b. Create "missense variant score"
4. Potential effector genes: ≥ 1 line of evidence for both T2D and OA
(High-confidence effector genes: ≥ 3)
 - a. Grouped genes by OA localization
 - i. "Equally Knee & Hip Associated":
 1. T2D
 2. One of OA:
 - a. Any site
 - b. Knee and/or hip
 - c. TJR (total joint replacement)
 - ii. "Mostly Knee Associated"
 1. T2D
 2. OA
 - a. Knee and/or hip
 - iii. "Mostly Hip associated"
 1. T2D
 2. OA
 - a. Hip and/or THR (total hip replacement)



University of
Pittsburgh®

School of
Public Health

eQTL + pQTL + DEG + KO + OMIM
eQTL + DEG + KO + OMIM

$$\frac{\text{eQTL} + \text{pQTL} + \text{DEG} + \text{KO} + \text{OMIM}}{\text{eQTL} + \text{DEG} + \text{KO} + \text{OMIM}} = \text{OA Score}$$

$$+ * \text{Missense Score} = \text{T2D Score}$$

$$\text{Total Score}$$

Gene	Score
TCF7L2	4
TMEM119	4
TMEM176A	4

*Missense Score only included if OA score or T2D score ≥ 1

Scoring

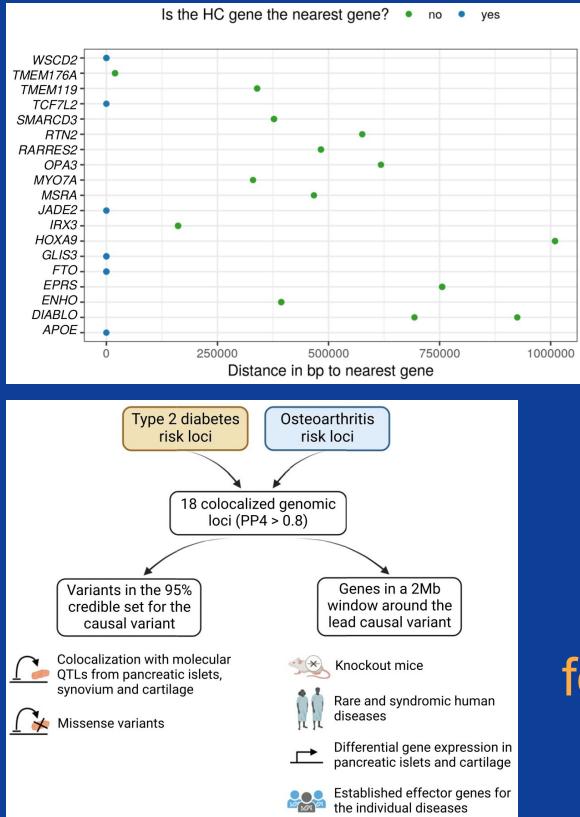
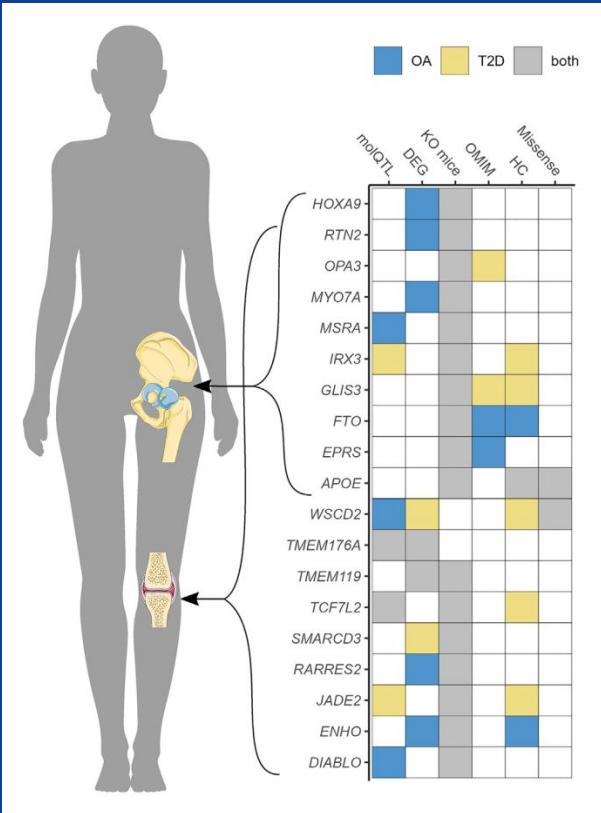
Classification:

“Potential” genes

≥ 1

“High Confidence” genes =
“Potential” & ≥ 3

Gene	Score	Gene	Score	Gene	Score	Gene	Score	Gene	Score	Gene	Score
APOE	3	EPRS	3	HOXA9	3	MSRA	3	RARRES2	3	WSCD2	3
DIABLO	3	FTO	3	IRX3	3	MY07A	3	RTN2	3		
ENHO	3	GLIS3	3	JADE2	3	OPA3	3	SMARCD3	3		



Results

906 genes investigated

- 72 “likely” effector genes
- 19 “high-confidence” effector genes
 - 17 confirmed in KO mice
 - 11 previously unreported
 - 2 contained missense variants
 - 6 nearest gene to lead variant [top right image]

12/18 colocalizing T2D/OA regions showed evidence for colocalizing with molQTL [bottom right image]



- ✓ Genetic overlap of T2D and OA
- ✓ Colocalization analysis
- ✓ KO mouse phenotypes
- ✓ Rare & syndromic phenotypes
- ✓ Differential gnxp
- ✓ Multi-trait colocalization (eQTL & pQTL)
- ✓ Scoring potential effector genes
- ❑ Multi-trait colocalization (adiposity measures)
- ❑ Pathway analysis
- ❑ Classification of high-confidence genes
“Druggable genome”
- ❑ Causal inference analysis
- ❑ Two-step MR
- ❑ Tissue-specific approaches

Recap

From previous recap

- Confirmed genetic correlation of T2D & OA
 - Knee (robust)
 - Hip (less robust)
- Used GWAS from OA (various sub-types) and T2D as input for colocalization
 - 51 total signals
 - 18 unique signals

Since previous recap

- Gathered ancillary data to triangulate evidence with colocalization analysis
- Resolved the 18 unique signals (containing 902 genes) into
 - 72 “likely” effector genes
 - 19 “high-confidence” effector genes
 - Localized “high-confidence” genes according to knee or both



Finished

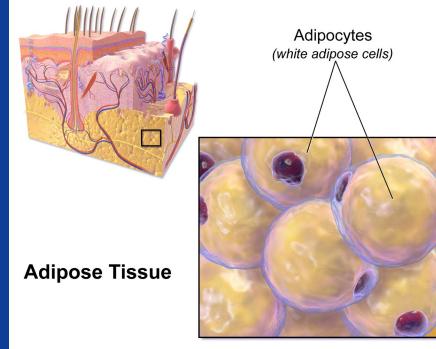
- ✓ Genetic overlap of T2D and OA
- ✓ Colocalization analysis
- ✓ KO mouse phenotypes
- ✓ Rare & syndromic phenotypes
- ✓ Differential gnxp
- ✓ Multi-trait colocalization (eQTL & pQTL)
- ✓ Scoring potential effector genes

New Order

- “Druggable genome”
- Multi-trait colocalization (adiposity measures)
- Classification of high-confidence genes
- Pathway analysis
- Causal inference analysis
- Two-step MR
- Tissue-specific approaches

Moving Forward

We have 72 “likely” effector genes and
19 “high-confidence” effector genes



How to make
these
actionable?

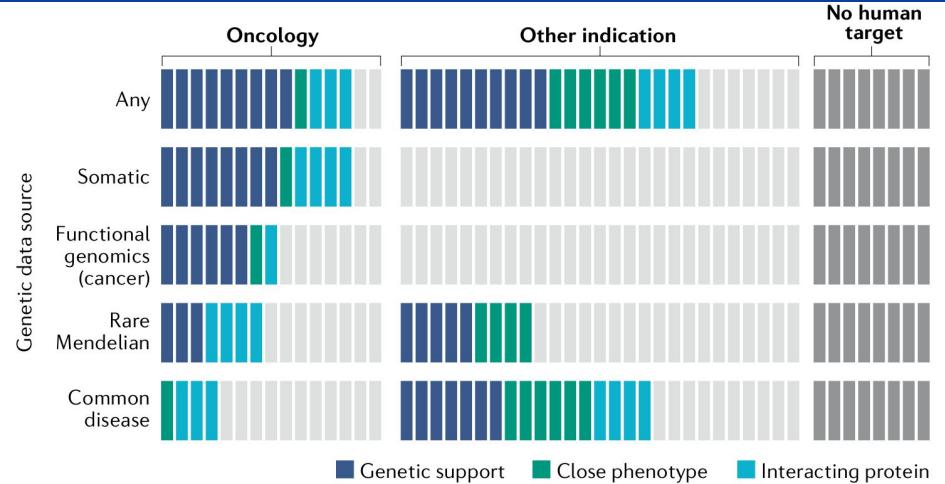
How does
BMI/adiposity fit
into the picture?



Ochoa et
al. (2022)

REGULATORY WATCH

Human genetics evidence
supports two-thirds of the 2021
FDA-approved drugs



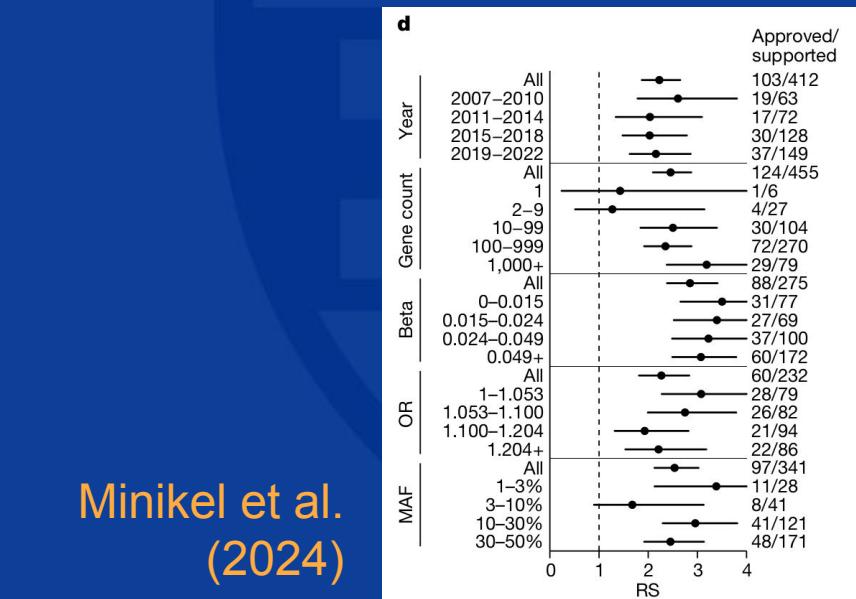
Drug Repurposing

Analysis

Refining the impact of genetic evidence on
clinical success

<https://doi.org/10.1038/s41586-024-07316-0>

Received: 5 July 2023



Minikel et al.
(2024)



1. *Genetic overlap of T2D and OA*
2. *Colocalization analysis*
3. *KO mouse phenotypes*
4. *Rare & syndromic phenotypes*
5. *Differential gnxp*
6. *Mutli-trait colocalization (eQTL & pQTL)*
7. Scoring potential effector genes
8. **“Druggable genome”**
9. *Mutli-trait colocalization (adiposity measures)*
10. *Classification of high-confidence genes*
11. *Pathway analysis*
12. *Causal inference analysis*
13. *Two-step MR*
14. *Tissue-specific approaches*

Methods

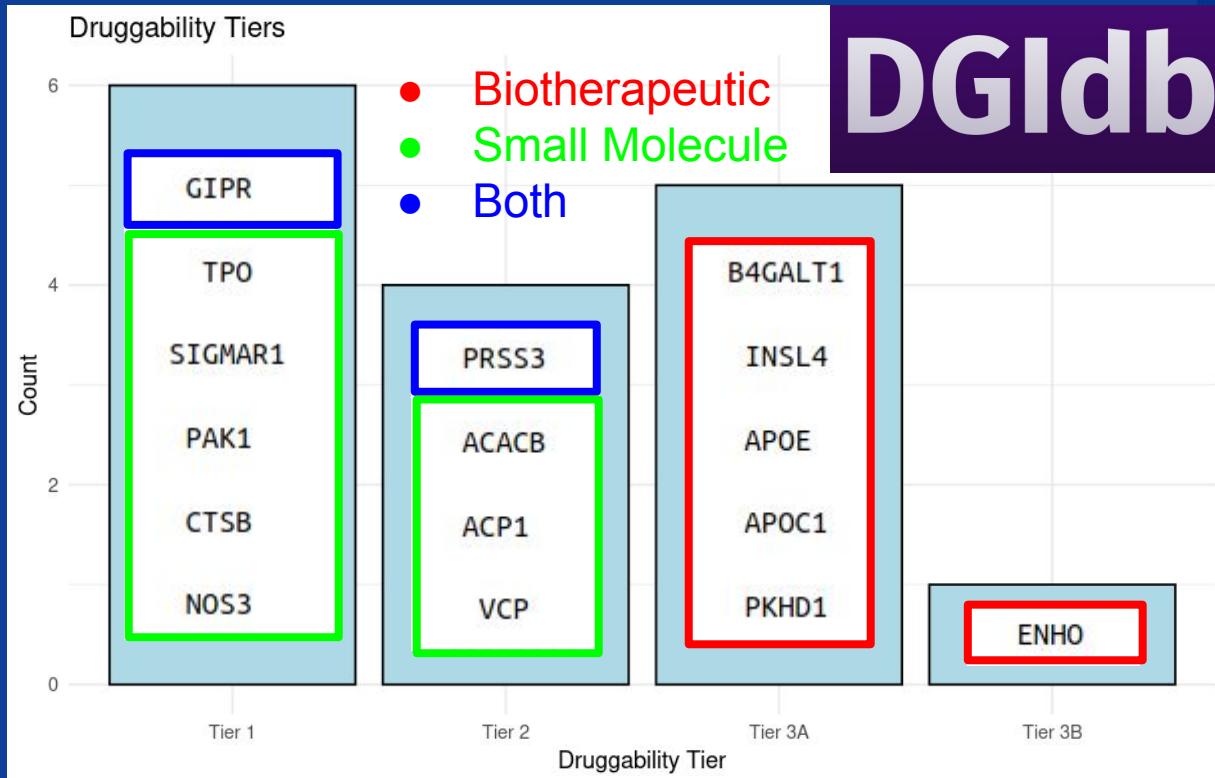
Queried Druggable Genome database for
72 “likely” effector genes

Three tiers of “targets”:

1. Clinical-phase/already-approved (4,479)
2. Known bioactive drug-like small-molecule binding partners and genes with 50% identity (75% of sequence) (682)
3. Extracellular, secreted, distant similarity to approved targets, members of druggable gene families (2,370)



Results



Three tiers of “targets”:

1. Clinical-phase/already approved
2. Known bioactive drug-like small-molecule binding partners and genes with 50% identity (75% of sequence)
3. Extracellular, secreted, distant similarity to approved targets, members of druggable gene families



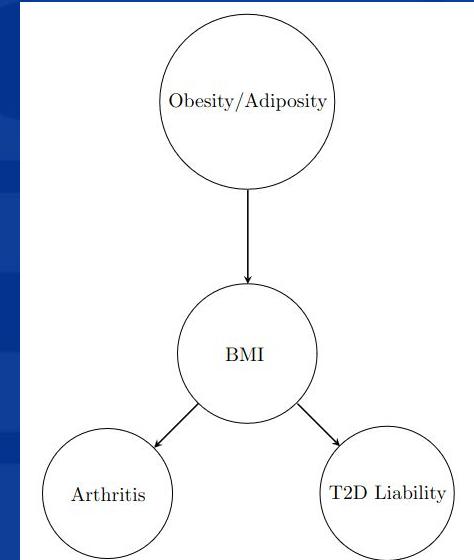
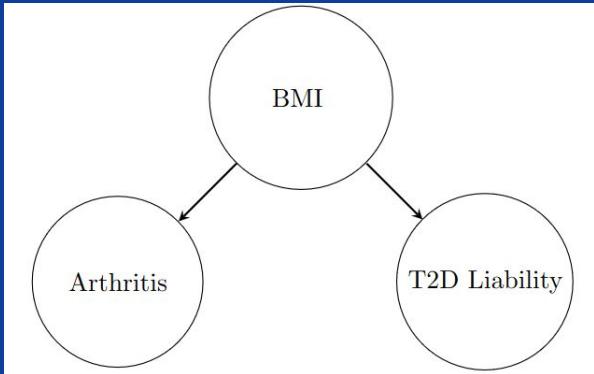
Adiposity

BMI causally affects
OA/T2D

In reality, BMI is proxy for
true cause

Other potential proxies:
WHR

Whole-body fat mass
Body fat percentage





1. *Genetic overlap of T2D and OA*
2. *Colocalization analysis*
3. *KO mouse phenotypes*
4. *Rare & syndromic phenotypes*
5. *Differential gnxp*
6. *Mutli-trait colocalization (eQTL & pQTL)*
7. *Scoring potential effector genes*
8. *"Druggable genome"*
- 9. Mutli-trait colocalization
(adiposity measures)**
10. *Classification of high-confidence genes*
11. *Pathway analysis*
12. *Causal inference analysis*
13. *Two-step MR*
14. *Tissue-specific approaches*

Methods

Subset to loci that showed evidence of colocalization between T2D and OA (95% credible interval)

Perform additional colocalization on measures of adiposity

Same analysis (using *HyPrColoc*, with the same priors)



Data combined from GIANT Consortium and UKBB

- BMI (n=806,834)
- WHR (n=679,734)
- Whole-body fat mass (n=330,762)
- Body fat percentage (n=331,117)

Adiposity Measures

Subset to loci that showed evidence of colocalization between T2D and OA (95% credible interval)

Perform additional colocalization on measures of adiposity

Same analysis (using *HyPrColoc*, with the same priors)



Results

18 genomic regions colocalize (T2D/OA)

16 show association ≥ 1 adiposity measure

Potential alternative mechanisms
(other than obesity)

2 show no association with any adiposity
measure

Reminder (Adiposity measures):

BMI
WHR
Whole-body fat mass
Body fat percentage

“High
confidence”
effector
genes

TMEM176A
RARRES2
SMARCD3
GLIS3



Methods

Classified
high-confidence genes
based on association of
obesity

1. OMIM
2. PhenoScanner
3. All remaining

10. Classification of high-confidence genes

11. Pathway analysis
12. Causal inference analysis
13. Two-step MR
14. Tissue-specific approaches



University of
Pittsburgh®

School of
Public Health

Methods

High-confidence effector genes (association with obesity)

The OMIM logo, which features a green circular emblem with a portrait of a man's head and the acronym "OMIM" in green serif letters to the left.

Level 1: *FTO*

A box with a green border containing the word "Remaining" in large, bold, green sans-serif font.

Level 3: *ENHO*, *EPRS*, *MYO7A*,
OPA3, *RARRES2*, *RTN2*,
SMARCD3, *TMEM119*, *TMEM176A*



PhenoScanner v2

A database of human genotype-phenotype associations

The Ensembl logo, featuring the word "Ensembl" in blue and red lowercase letters next to a stylized "e!" symbol.

Level 2:
APOE, *DIABLO*,
GLIS3, *HOXA9*,
IRX3, *JADE2*,
MRSA, *TCF7L2*,
WSCD2



1. *Genetic overlap of T2D and OA*
 2. *Colocalization analysis*
 3. *KO mouse phenotypes*
 4. *Rare & syndromic phenotypes*
 5. *Differential gnxp*
 6. *Mutli-trait colocalization (eQTL & pQTL)*
 7. Scoring potential effector genes
 8. *“Druggable genome”*
 9. *Mutli-trait colocalization (adiposity measures)*
 10. *Classification of high-confidence genes*
- ## 11. Pathway analysis
12. *Causal inference analysis*
 13. *Two-step MR*
 14. *Tissue-specific approaches*

Methods

Gene-set enrichment analysis (GSA) performed on likely & high-confidence effector genes (stratified by knee or hip OA)

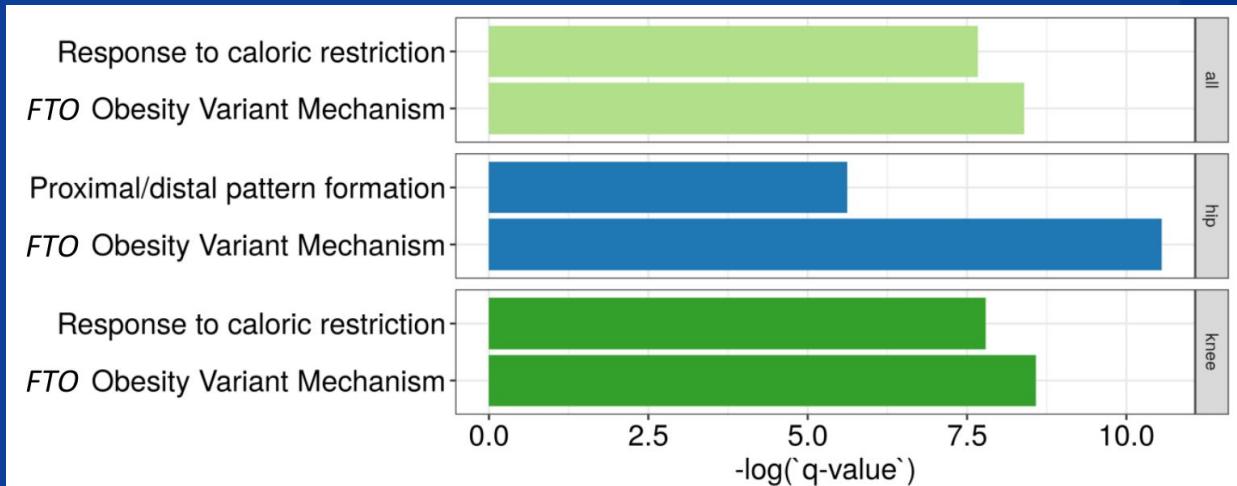
ConsensusPathDB used to examine functional annotation by testing enrichment in networks:

- Reactome
- KEGG
- WikiPathways
- Gene Ontology

Significance threshold: FDR < 0.05



Gene-set enrichment analysis (GSA)
performed on likely & high-confidence
effector genes (stratified by knee or hip
OA)



Results

Figure S4). These results provide biological support for the link between obesity and both diseases and for the association between bone development and hip osteoarthritis.⁷

ConsensusPathDB used to examine functional annotation by testing enrichment in networks:

- Reactome
- KEGG
- WikiPathways
- Gene Ontology



- ✓ Genetic overlap of T2D and OA
- ✓ Colocalization analysis
- ✓ KO mouse phenotypes
- ✓ Rare & syndromic phenotypes
- ✓ Differential gnxp
- ✓ Multi-trait colocalization (eQTL & pQTL)
- ✓ Scoring potential effector genes
- ✓ Multi-trait colocalization (adiposity measures)
- ✓ Pathway analysis
- ✓ Classification of high-confidence genes
- ✓ “Druggable genome”
- ❑ Causal inference analysis
- ❑ Two-step MR
- ❑ Tissue-specific approaches

Recap #1

- Confirmed genetic correlation of T2D & OA
 - Knee (robust)
 - Hip (less robust)
- Used GWAS from OA (various sub-types) and T2D as input for colocalization
 - 51 total signals
 - 18 unique signals

Recap #2

- Gathered ancillary data to triangulate evidence with colocalization analysis
- Resolved the 18 unique signals (containing 902 genes) into
 - 72 “likely” effector genes
 - 19 “high-confidence” effector genes
 - Localized “high-confidence” genes according to knee or both

Recap since recap #2:

Assessed “druggability”, incorporated adiposity measures, gene-set enrichment analysis



Moving Forward

Finished

- ✓ Genetic overlap of T2D and OA
- ✓ Colocalization analysis
- ✓ KO mouse phenotypes
- ✓ Rare & syndromic phenotypes
- ✓ Differential gnxp
- ✓ Multi-trait colocalization (eQTL & pQTL)
- ✓ Scoring potential effector genes
- ✓ “Druggable genome”
- ✓ Multi-trait colocalization (adiposity measures)
- ✓ Classification of high-confidence genes
- ✓ Pathway analysis

Home Stretch

- Causal inference analysis
- Two-step MR
- Tissue-specific approaches



Terminology

Methods Section Title	Causal	Bi-Directional	Two-Sample	TwoSampleMR package	Two-Step	Tissue-Specific Data
<i>Causal inference analysis</i>	✓	✓	✓	✓	✗	✗
<i>Two-step MR</i>	✓	✗	✓	✓	✓	✓ pancreas, synovium, cartilage,
<i>Tissue-specific effects</i>	✓	✗	✓	?	✗	✓ brain & adipose

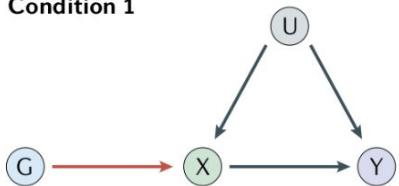
Bi-directional = perform analysis twice, switching exposure and outcome

Two-Sample = different exposure & outcome samples

Two-Step = special case (covered later)



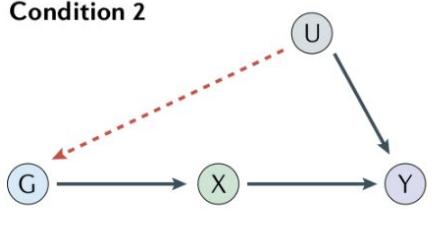
Condition 1



“Relevance”

Generally assessed through
 F -statistic (hope for $F > 10$)

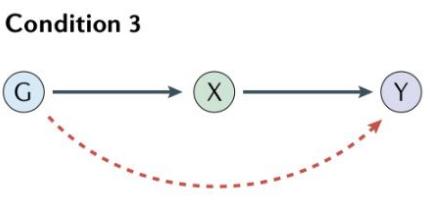
Condition 2



“Independence” (or “exchangeability”)

No unmeasured confounders

Condition 3



“Exclusion”

Instrument only affects
outcome through exposure

MR Basics

Primary Concern

Weak instruments

Primary Concern

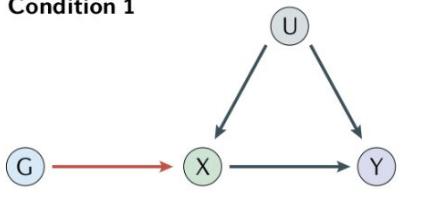
Population stratification (but
also possibly IGE, AM,
transmission distortion)

Primary Concern

(Horizontal) Pleiotropy



Condition 1



“Relevance”

Generally assessed through
 F -statistic

Condition 2

“Independence” (or “exchangeability”)

No unmeasured confounders

Condition 3

“Exclusion”

Instrument only affects
outcome through exposure

MR Basics

These three conditions are sufficient
to test null hypothesis (of no causal
effect)

However, to estimate causal effect an
additional **point-estimate-identifying**
assumption is necessary

Two frequently used:

1. Homogeneity of effect
2. Monotonicity in IV-exposure
association



MR Sensitivity

Steiger Filtering:

Filters instruments (SNPs) to ensure stronger association with exposure than outcome

Weighted Median:

Outlier-removal method commonly used with summary-level data (requires that > half SNPs are valid instruments)

F-testing:

Straightforward *F*-test, ensures sufficient instrument strength (rule of thumb: $F > 10$)

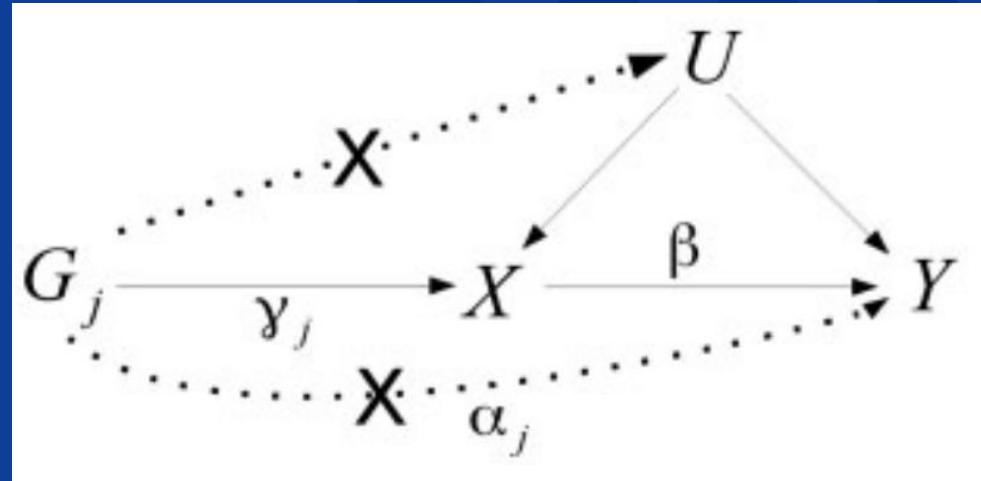
Heterogeneity Testing (Q-statistic):

Analogous to Cochran's Q from meta-analysis

MR-Egger

- InSIDE assumption
 - Wordy, but boils down to

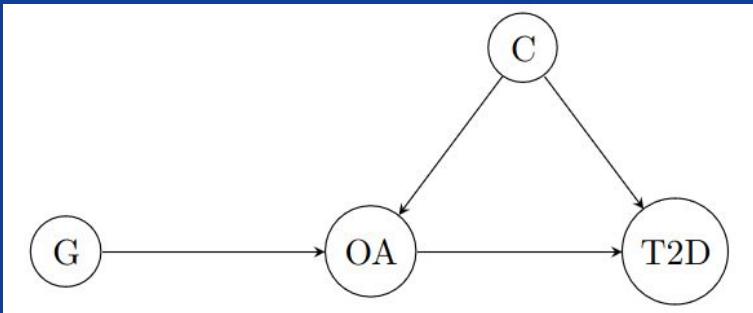
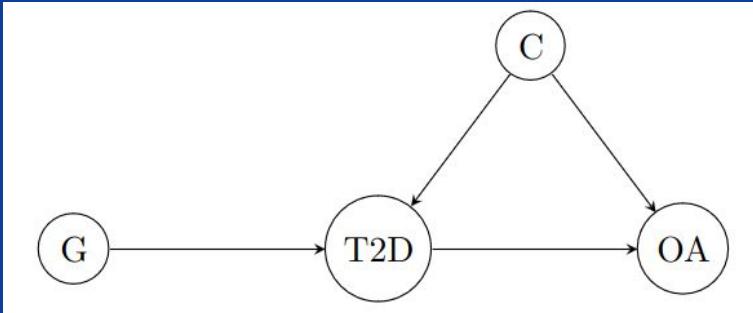
$$\text{Corr}(\gamma_j, \alpha_j) = 0$$





Methods

Bi-directional MR: *TwoSampleMR* R package



Instrument selection:

- Genome-wide significant
- LD-clumped over 10-Mb window
- Steiger filtering performed

Methods used

- IVW
- WM

Sensitivity analyses

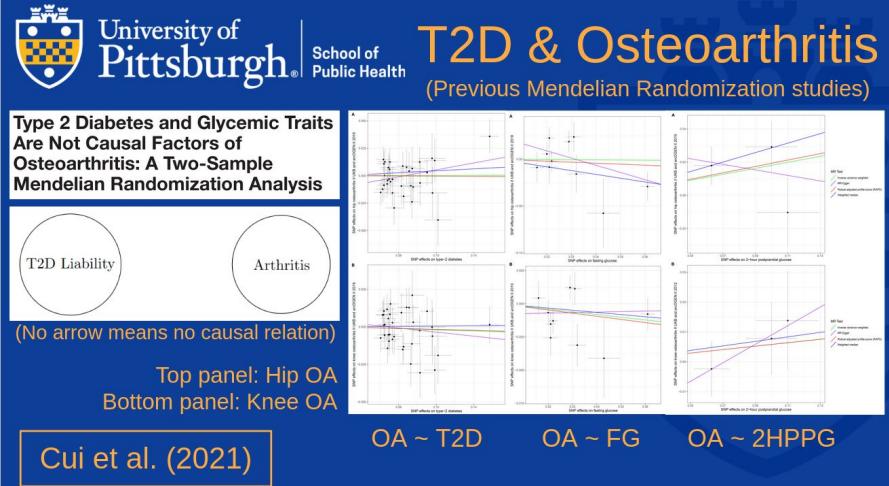
- Q-statistics
- MR-Egger

TwoSampleMR 0.6.2



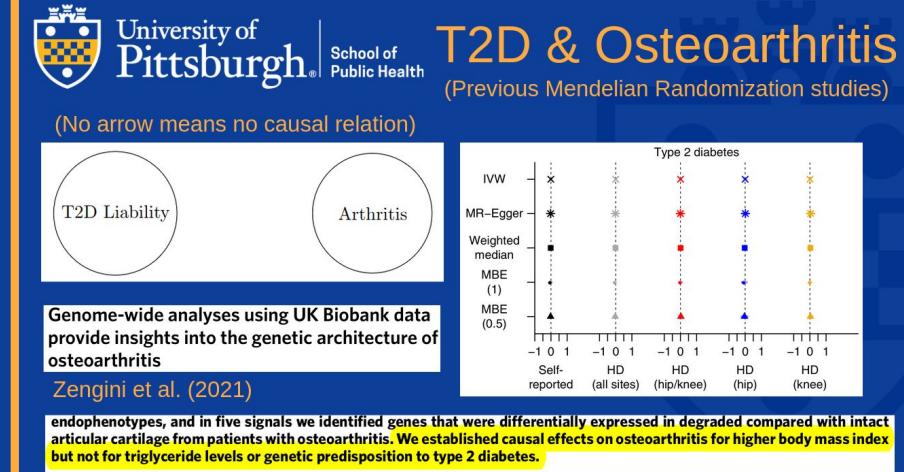
TwoSampleMR 0.6.2

Recall from earlier:



Causal inference analyses using MR showed evidence for a non-causal relationship between the two diseases (Table S6), consistent with smaller-scale studies in the literature.¹²

Results





1. *Genetic overlap of T2D and OA*
2. *Colocalization analysis*
3. *KO mouse phenotypes*
4. *Rare & syndromic phenotypes*
5. *Differential gnxp*
6. *Mutli-trait colocalization (eQTL & pQTL)*
7. *Scoring potential effector genes*
8. *Mutli-trait colocalization (adiposity measures)*
9. *Pathway analysis (methods)*
10. *Classification of high-confidence genes*
11. *“Druggable genome”*
12. *Causal inference analysis*

13. Two-step MR

14. *Tissue-specific approaches*

Methods

Two-step MR between adiposity measures and T2D or OA

1. Assess causality of adiposity gnxp in tissues
2. Two-sample MR between each gene and T2D or OA

cis eQTLs of each high-confidence effector gene in disease-relevant tissues used as mediators (excluding independent eQTLs from risk variants of each adiposity measure)

Sensitivity analyses

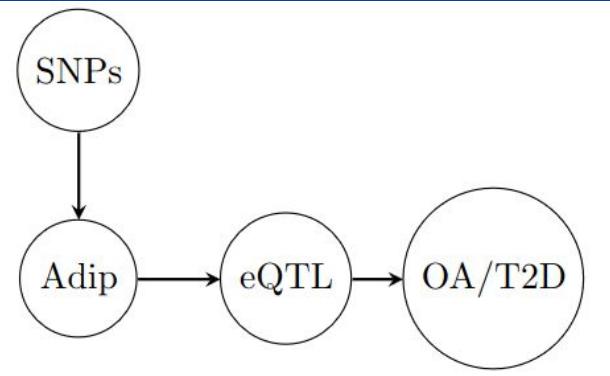
- Steiger filtering
- Tested heterogeneity
- MR-Egger
- F-statistics



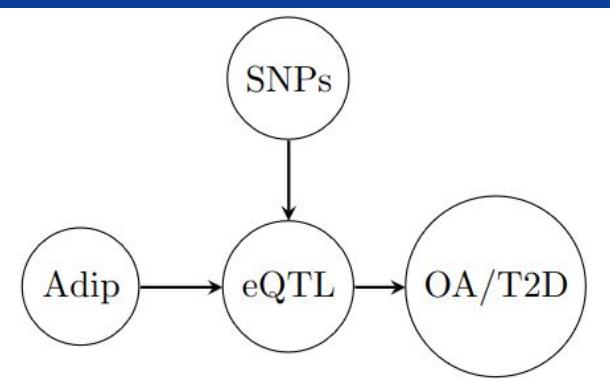
Two-Step MR

Two-step MR between adiposity measures and T2D or OA

Step 1



Step 2



1. Assess causality of adiposity for gnxp in tissues for high-confidence effector genes (excluding independent eQTLs from risk variants of each adiposity measure: $R^2 < 1e-3$, 10Mb window)
 - a. OA
 - i. Intact cartilage (n=95)
 - ii. Degenerated cartilage (n=87)
 - iii. Synovium (n=77)
 - b. T2D
 - i. Pancreatic islets (383 controls/37 cases)

Sensitivity analyses

- Steiger filtering
- Tested heterogeneity
- MR-Egger
- F-statistics



Altogether: evidence of causal relationship between adiposity & 9 high-confidence effector genes.

effector genes (Table S7). For example, we found that all measures of adiposity have a causal effect on higher expression of *IRX3* in synovium or pancreatic islets and on lower expression of *RTN2* in osteoarthritis cartilage. For the high-

Results

exposure	Tissue	Gene	Effect_Direction
bmi	PancreaticIslets	IRX3	+
body_fat_per	PancreaticIslets	IRX3	+
whole_body_fat_mass	PancreaticIslets	IRX3	+
body_fat_per	Synovium	IRX3	+
whole_body_fat_mass	Synovium	IRX3	+
whr	Synovium	IRX3	+
bmi	LowGradeCartilage	RTN2	-
body_fat_per	LowGradeCartilage	RTN2	-
whole_body_fat_mass	LowGradeCartilage	RTN2	-
whr	LowGradeCartilage	RTN2	-



1. *Genetic overlap of T2D and OA*
2. *Colocalization analysis*
3. *KO mouse phenotypes*
4. *Rare & syndromic phenotypes*
5. *Differential gnxp*
6. *Mutli-trait colocalization (eQTL & pQTL)*
7. *Scoring potential effector genes*
8. *Mutli-trait colocalization (adiposity measures)*
9. *Pathway analysis*
10. *Classification of high-confidence genes*
11. *“Druggable genome”*
12. *Causal inference analysis*
13. *Two-step MR*

14. **Tissue-specific approaches**

Methods

Assessed tissue-specific role of BMI

1. OA (brain eQTLs) (86 SNPs)
2. T2D (adipose eQTLs) (140 SNPs)

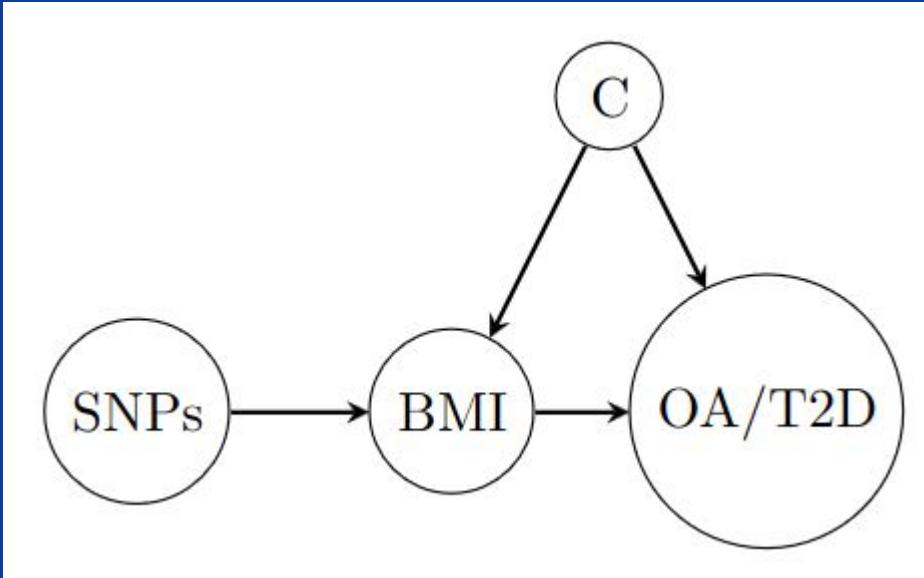
Summary-level MR Wald Ratio:

$$\frac{\hat{\beta}_{\text{OA}/\text{T2D}}}{\hat{\beta}_{\text{BMI}}}$$

Sensitivity analyses:

MR-Egger
WM

Z-test to assess difference between adipose & brain



SNPs subset:

- Brain: n=1194, loci=140
- Adipose: 1257, loci = 86

Methods

Assessed tissue-specific role of BMI

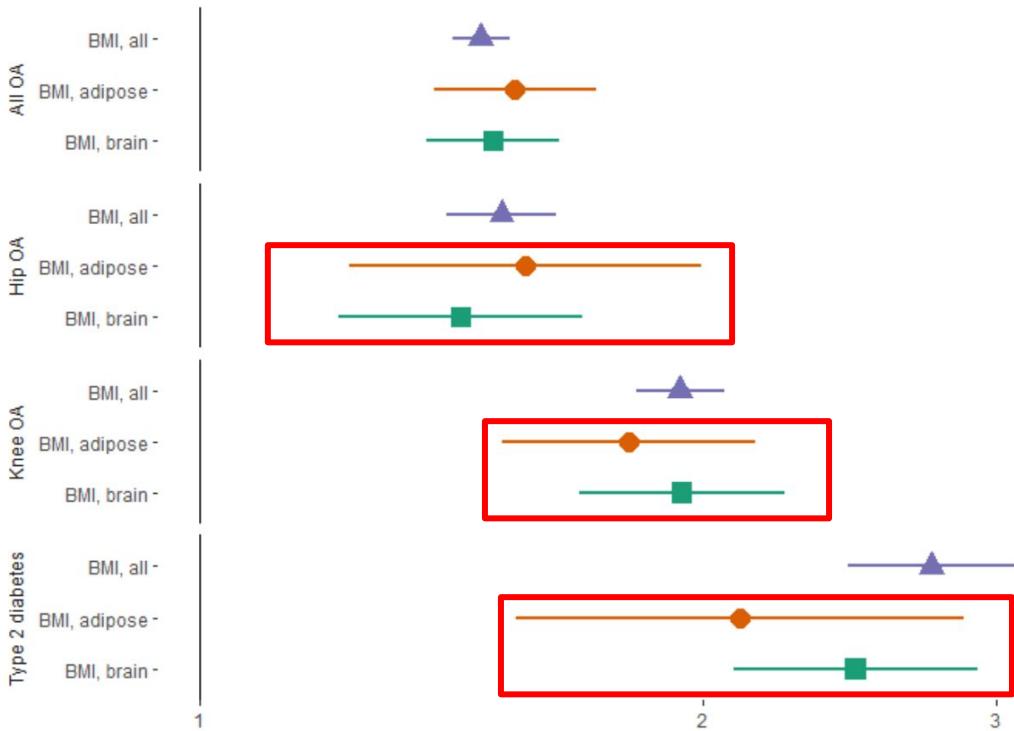
Summary-level MR Wald Ratio:

$$\frac{\hat{\beta}_{\text{SNPs-OA/T2D}}}{\hat{\beta}_{\text{SNPs-BMI}}}$$

Sensitivity analyses:

MR-Egger
WM

Z-test to assess difference between
adipose & brain

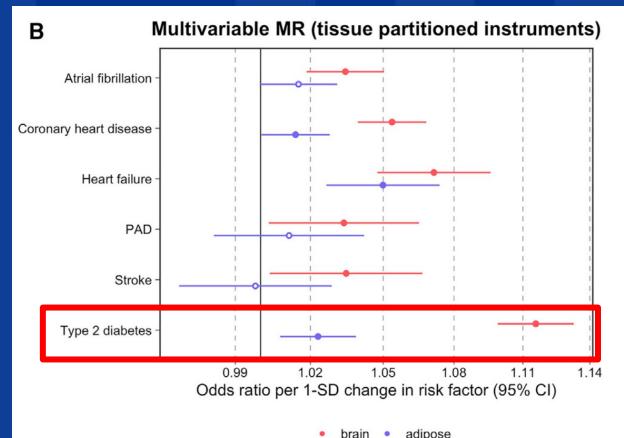


Results

Causal role of BMI-associated variants with tissue-specific effects

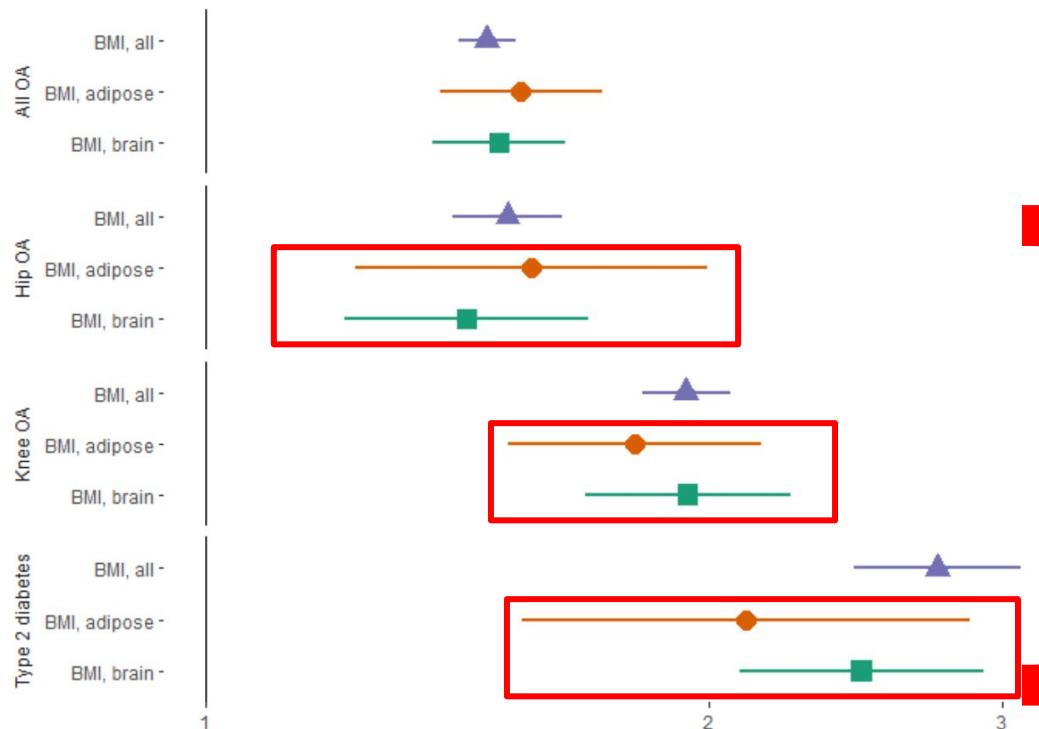
- Hip: stronger in adipose
- Knee: stronger in brain
- T2D: stronger in adipose

CIs overlap, but T2D result has been seen before





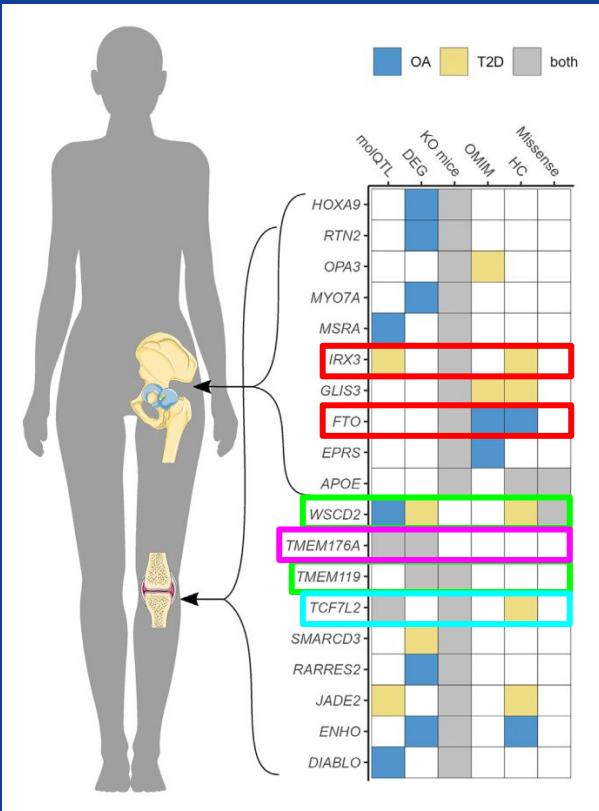
Results



Causal role of BMI-associated variants with tissue-specific effects

- Hip: stronger in adipose
- Knee: stronger in brain
- T2D: stronger in adipose

eQTLs (Table S12). Our results suggest a similar biological underpinning of the adiposity effect captured by BMI on type 2 diabetes and knee osteoarthritis but potentially different processes for hip osteoarthritis.



Results

Biological Results

1. *FTO* & *IRX3*
2. *TCF7L2*
3. *TMEM119* & *WSCD2*
4. *TMEM176A*

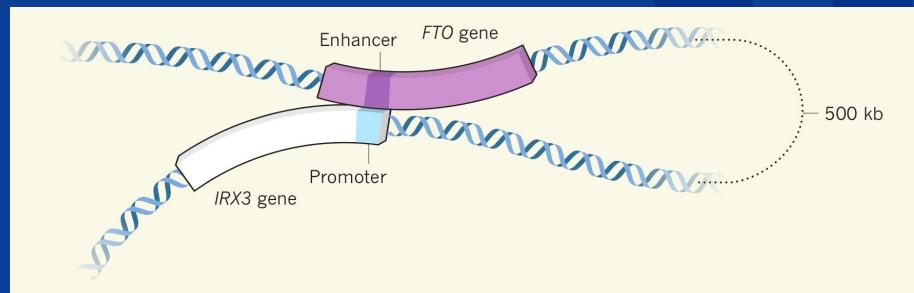
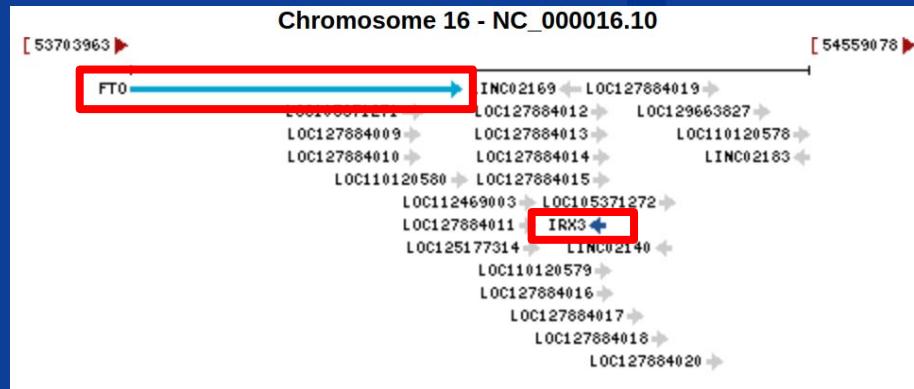
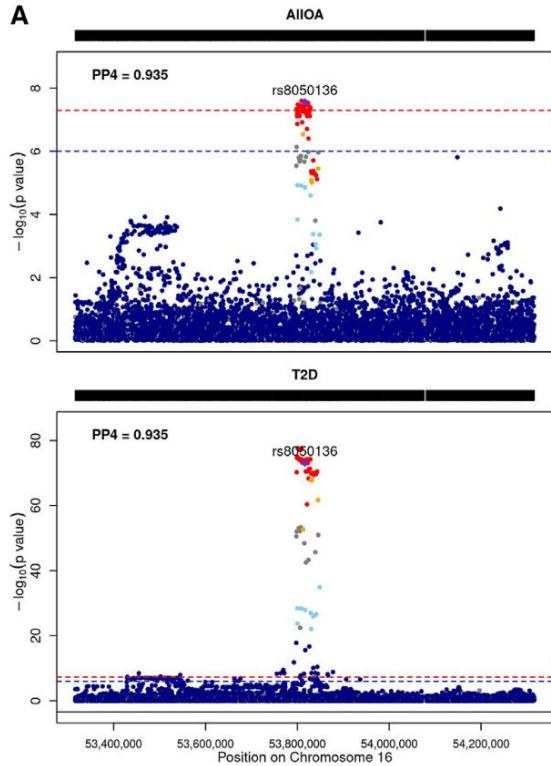


University of
Pittsburgh®

School of
Public Health

Results

FTO
&
IRX3



Gorken & Ren (2014)

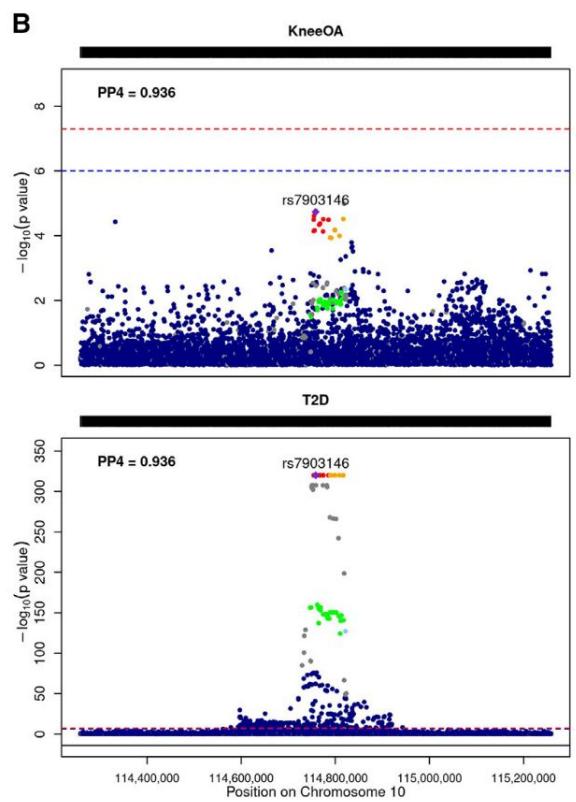
Causal effect of increased expression on of *IRX3* risk of T2D
OR=1.16
(1.08,1.25)



Results

TCF7L2

- One of the highest scoring effector genes (score = 4)
- Among leading signals for T2D risk (remains after adjusting for BMI)
- Shown to regulate genes related to cartilage destruction
- Key effector gene of Wnt/ β -catenin signaling pathway



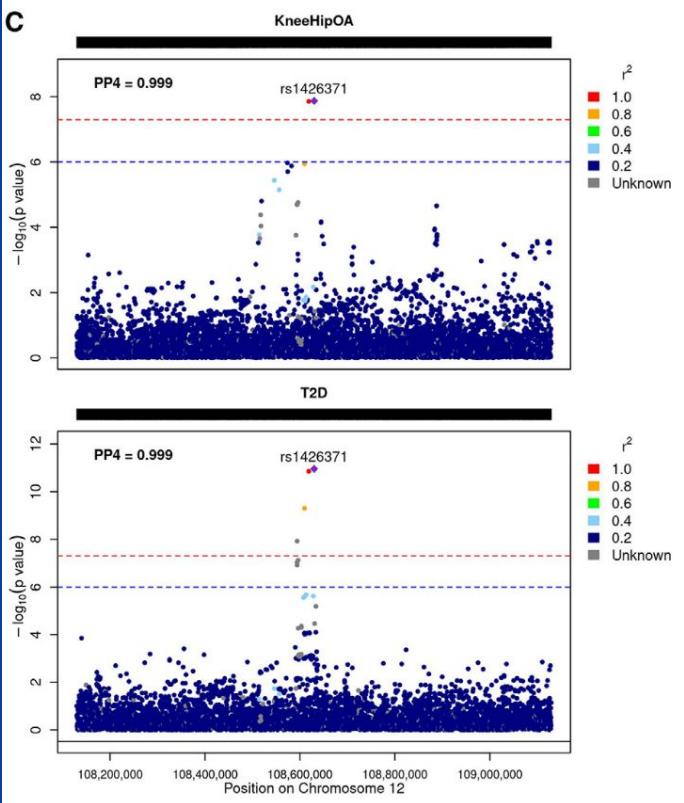


Results

TMEM119 & WSCD2

95% credible set contains two SNPs:

1. rs1426371 (intrinsic)
2. rs3764002 (missense)



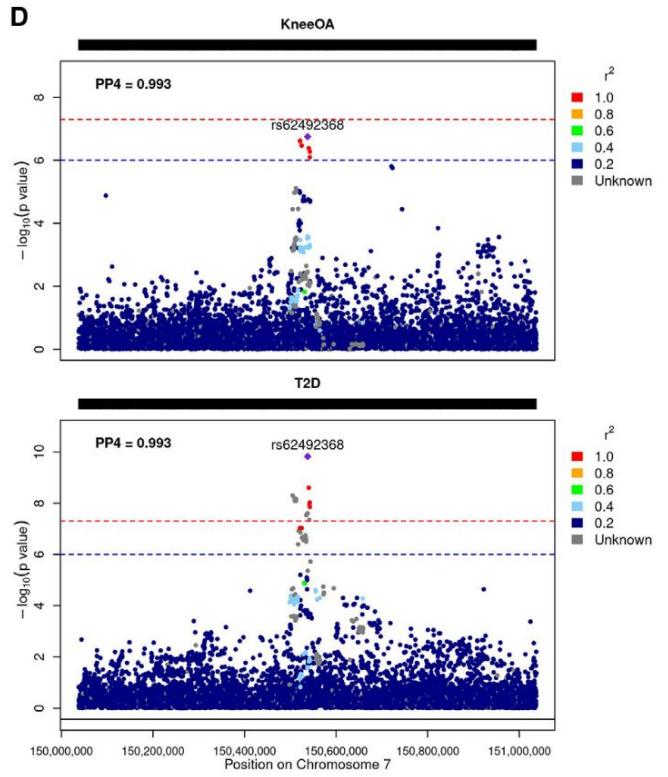
	rs1426371			rs3764002		
	EA	beta	pval	EA	beta	pval
Knee OA	A	-0.0515	8.86E-10	T	-0.0508	8.90E-10
Knee and/or hip OA	A	-0.0408	1.34E-08	T	-0.0402	1.39E-08
All OA	A	-0.0246	3.62E-06	T	-0.0242	3.77E-06
TJR	A	-0.0412	2.67E-05	T	-0.0403	3.27E-05
T2D	A	-0.05	1.10E-11	T	-0.049	1.40E-11



Results

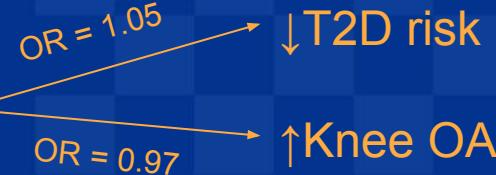
TMEM176A

Previously unreported for OA or
T2D



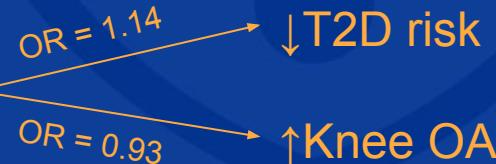
Intact OA cartilage

↓ *TMEM176A* expression



Pancreatic Islets

↓ *TMEM176A* expression

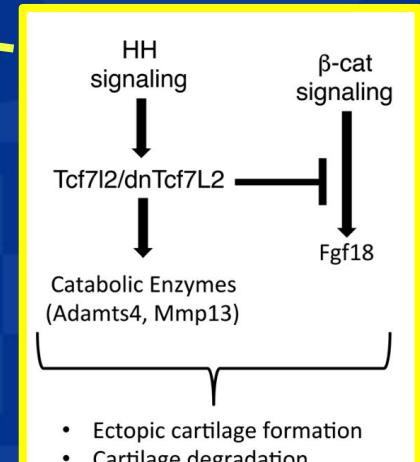
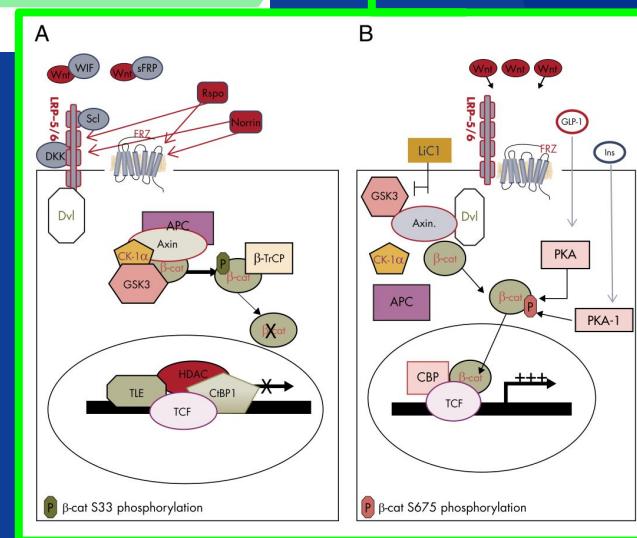




Mechanistic Insights

an alternative biological pathway to increased BMI. It has been shown that isoforms of *TCF7L2* regulate the expression of genes related to cartilage destruction in human chondrocytes.⁵² *TCF7L2* is a key effector gene of the Wnt/β-catenin signaling pathway. This pathway plays a role in both type 2 diabetes, through glucose homeostasis, and osteoarthritis, through cartilage and bone formation.^{53,54}

Association of *TCF7L* with T2D persists after controlling for BMI suggests alternative pathway.



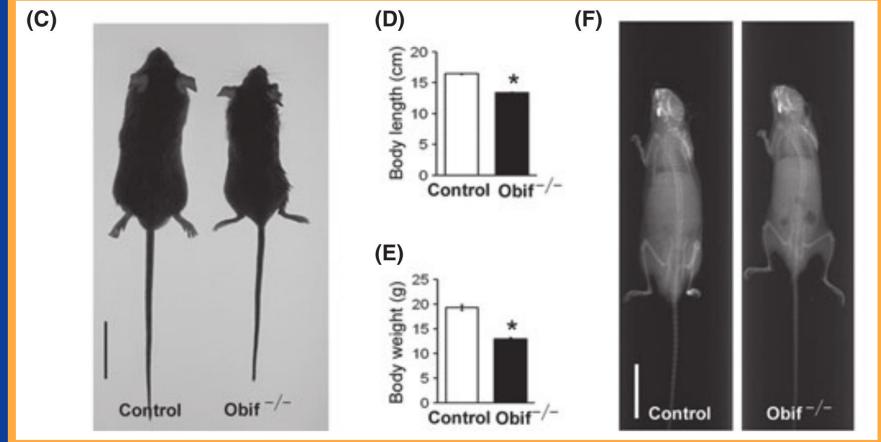
Rockel et al. (2016)

Jin (2016)



ciated with bone fracture.⁶⁶ One possible link between bone and lipid metabolism is the fact that osteoblasts and adipocytes share a common progenitor cell in adult bone marrow with a degree of plasticity that can lead to an imbalance between the two cell lineages.⁶⁷ In support of this link, differentiation regulation of osteoblasts is highlighted by one of the identified high-confidence effector genes, *TMEM119*. In summary, we highlight three potential biological mechanisms underpinning the comorbidity between type 2 diabetes and osteoarthritis: obesity, imbalance between osteoblasts and adipocytes differentiation in adult bone marrow and the Wnt/β-catenin signaling pathway.

Mechanistic Insights



have been relatively well elucidated; however, the exact roles of cell-extrinsic molecules are less clear. We previously identified human and mouse *Obif*, an osteoblast induction factor, also known as *Tmem119*, which encodes a single transmembrane protein. OBIF is predominantly expressed in osteoblasts in mouse. While



University of
Pittsburgh®

School of
Public Health

Main Takeaways

Genome-wide genetic correlation between T2D & OA, notably subtypes related to knee

18 robust colocalization signals between T2D & OA

Pathway enrichment:

Knee–lipid metabolism
Hip–skeletal formation

Biological mechanisms

1. Obesity
2. Osteoblast & adipocyte imbalance (bone marrow differentiation)
3. Wnt/β-catenin signaling pathway

19 high-confidence effector genes (targets for future research)

Adiposity plays cell-type specific risk



University of
Pittsburgh®

School of
Public Health

FIN.



Bibliography

- Arruda, A. L., Hartley, A., Katsoulas, G., Smith, G. D., Morris, A. P., & Zeggini, E. (2023). Genetic underpinning of the comorbidity between type 2 diabetes and osteoarthritis. *The American Journal of Human Genetics*, 110(8), 1304–1318. <https://doi.org/10.1016/j.ajhg.2023.06.010>
- B. Blom, A., L. Van Lent, P., M. Van Der Kraan, P., & B. Van Den Berg, W. (2010). To Seek Shelter from the Wnt in Osteoarthritis? Wnt-Signaling as a Target for Osteoarthritis Therapy. *Current Drug Targets*, 11(5), 620–629. <https://doi.org/10.2174/138945010791011901>
- Bezerra De Souza, L. D., Oliveras-Fabregas, A., Espelt, A., Bosque-Prous, M., De Camargo Cancela, M., Teixidó-Compañó, E., & Jerez-Roig, J. (2021). Multimorbidity and its associated factors among adults aged 50 and over: A cross-sectional study in 17 European countries. *PLOS ONE*, 16(2), e0246623. <https://doi.org/10.1371/journal.pone.0246623>
- Boer, C. G., Hatzikotoulas, K., Southam, L., Stefnásdóttir, L., Zhang, Y., Coutinho De Almeida, R., Wu, T. T., Zheng, J., Hartley, A., Teder-Laving, M., Skoglund, A. H., Terao, C., Zengini, E., Alexiadis, G., Bärysenka, A., Björnsdóttir, G., Gabrielsen, M. E., Gilly, A., Ingvarsson, T., ... Minnich, L. J. (2021). Deciphering osteoarthritis genetics across 826,690 individuals from 9 populations. *Cell*, 184(18), 4784–4818.e17. <https://doi.org/10.1016/j.cell.2021.07.038>
- Bowden, J., Davey Smith, G., & Burgess, S. (2015). Mendelian randomization with invalid instruments: Effect estimation and bias detection through Egger regression. *International Journal of Epidemiology*, 44(2), 512–525. <https://doi.org/10.1093/ije/dvq080>
- Cui, Z., Feng, H., He, B., Xing, Y., Liu, Z., & Tian, Y. (2021). Type 2 Diabetes and Glycemic Traits Are Not Causal Factors of Osteoarthritis: A Two-Sample Mendelian Randomization Analysis. *Frontiers in Genetics*, 11, 597876. <https://doi.org/10.3389/fgene.2020.597876>
- De S. V. Arnuda, A. L., Hartley, A., Katsoulas, G., Smith, G. D., Morris, A. P., & Zeggini, E. (2023). *Insights into the comorbidity between type 2 diabetes and osteoarthritis*. <https://doi.org/10.1101/2023.03.28.23287861>
- Foley, C. N., Staley, J. R., Breen, P. G., Sun, B. B., Kirk, P. D. W., Burgess, S., & Howson, J. M. M. (2021). A fast and efficient colocalization algorithm for identifying shared genetic risk factors across multiple traits. *Nature Communications*, 12(1), 764. <https://doi.org/10.1038/s41467-020-20885-8>
- He, L., Culmbo, I., Loika, Y., Arbeeny, K. G., Bagley, O., Duan, M., Yashin, A. I., & Kulminski, A. M. (2018). Causal effects of cardiovascular risk factors on onset of major age-related diseases: A time-to-event Mendelian randomization study. *Experimental Gerontology*, 107, 74–86. <https://doi.org/10.1016/j.exger.2017.09.019>
- Jin, T. (2016). Current Understanding on Role of the Wnt Signaling Pathway Effector TCF7L2 in Glucose Homeostasis. *Endocrine Reviews*, 37(3), 254–277. <https://doi.org/10.1210/er.2015-1146>
- Katsoulas, G., Steinberg, J., Tuerlings, M., Coutinho De Almeida, R., Southam, L., Swift, D., Meulenbelt, I., Wilkinson, J. M., & Zeggini, E. (2022). A molecular map of long non-coding RNA expression, isoform switching and alternative splicing in osteoarthritis. *Human Molecular Genetics*, 31(12), 2090–2105. <https://doi.org/10.1093/hmg/ddac017>
- Lee, J. J., McGuire, M., Iacono, W. G., & Chow, C. C. (2018). The accuracy of LD Score regression as an estimator of confounding and genetic correlations in genome-wide association studies. *Genetic Epidemiology*, 42(8), 783–795. <https://doi.org/10.1002/gepi.22161>
- Leyden, G. M., Shapland, C. Y., Davey Smith, G., Sanderson, E., Greenwood, M. P., Murphy, D., & Richardson, T. G. (2022). Harnessing tissue-specific genetic variation to dissect putative causal pathways between body mass index and cardiometabolic phenotypes. *The American Journal of Human Genetics*, 109(2), 240–252. <https://doi.org/10.1016/j.ajhg.2021.12.013>
- Li, M.-X., Yeung, J. M. Y., Cherny, S. S., & Sham, P. C. (2012). Evaluating the effective numbers of independent tests and significant p-value thresholds in commercial genotyping arrays and public imputation reference datasets. *Human Genetics*, 131(5), 747–756. <https://doi.org/10.1007/s00439-011-1118-2>
- Louati, K., Vidal, C., Berenbaum, F., & Sellam, J. (2015). Association between diabetes mellitus and osteoarthritis: Systematic literature review and meta-analysis. *RMD Open*, 1(1), e000077. <https://doi.org/10.1136/rmdopen-2015-000077>
- Mahajan, A., Tallan, D., Thurmer, M., Robertson, N. R., Torres, J. M., Rayner, N. W., Payne, A. J., Steinthorsdóttir, V., Scott, R. A., Grarup, N., Cook, J. P., Schmidt, E. M., Wuttke, M., Sarnowski, C., Magi, R., Nano, J., Gieger, C., Trompet, S., Lecoeur, C., ... McCarthy, M. I. (2018). Fine-mapping type 2 diabetes loci to single-variant resolution using high-density imputation and islet-specific epigenome maps. *Nature Genetics*, 50(11), 1505–1513. <https://doi.org/10.1038/s41588-018-0241-6>
- McKinney, J. D. (n.d.). *Photonics Illuminates the future of radar*.
- Minikel, E. V., Painter, J. L., Dong, C. C., & Nelson, M. R. (2024). Refining the impact of genetic evidence on clinical success. *Nature*, 629(8012), 624–629. <https://doi.org/10.1038/s41586-024-07316-0>
- Mizuhashi, K., Kanamoto, T., Ito, M., Morishi, T., Muranishi, Y., Omori, Y., Terada, K., Komori, T., & Furukawa, T. (2012). OBIF, an osteoblast induction factor, plays an essential role in bone formation in association with osteoblastogenesis. *Development, Growth & Differentiation*, 54(4), 474–480. <https://doi.org/10.1111/j.1440-169X.2012.01333.x>
- Nguyen, H., Manolova, G., Daskalopoulou, C., Vitoratou, S., Prince, M., & Prina, A. M. (2019). Prevalence of multimorbidity in community settings: A systematic review and meta-analysis of observational studies. *Journal of Comorbidity*, 9, 2235042X1987093. <https://doi.org/10.1177/2235042X19870934>
- Ochoa, D., Karim, M., Ghousaini, M., Hulcoop, D. G., McDonagh, E. M., & Dunham, I. (2022). Human genetics evidence supports two-thirds of the 2021 FDA-approved drugs. *Nature Reviews Drug Discovery*, 21(8), 551–551. <https://doi.org/10.1038/d41573-022-00120-3>
- Rockel, J. S., Yu, C., Whetstone, H., Craft, A. M., Reilly, K., Ma, H., Tsushima, H., Puvilindran, V., Al-Jazrawe, M., Keller, G. M., & Alman, B. A. (2016). Hedgehog inhibits β -catenin activity in synovial joint development and osteoarthritis. *Journal of Clinical Investigation*, 126(5), 1649–1663. <https://doi.org/10.1172/JCI80205>
- Schizophrenia Working Group of the Psychiatric Genomics Consortium, Bulik-Sullivan, B. K., Loh, P.-R., Finucane, H. K., Ripke, S., Yang, J., Patterson, N., Daly, M. J., Price, A. L., & Neale, B. M. (2015). LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nature Genetics*, 47(3), 291–295. <https://doi.org/10.1038/ng.3211>
- Skou, S. T., Mair, F. S., Fortin, M., Guthrie, B., Nunes, B. P., Miranda, J. J., Boyd, C. M., Patil, S., Mtenga, S., & Smith, S. M. (2022). Multimorbidity. *Nature Reviews Disease Primers*, 8(1), 48. <https://doi.org/10.1038/s41572-022-00376-4>
- Steinberg, J., Southam, L., Roumeliotis, T. I., Clark, M. J., Jayasuria, R. L., Swift, D., Shan, K. M., Butterfield, N. C., Brooks, R. A., McCaskie, A. W., Bassett, J. H. D., Williams, G. R., Choudhary, J. S., Wilkinson, J. M., & Zeggini, E. (2021). A molecular quantitative trait locus map for osteoarthritis. *Nature Communications*, 12(1), 1309. <https://doi.org/10.1038/s41467-021-21593-7>
- Vihuela, A., Varshney, A., Van De Bunt, M., Prasad, R. B., Asplund, O., Bennett, A., Boehnke, M., Brown, A. A., Erdos, M. R., Fadista, J., Hansson, O., Hatem, G., Howald, C., Iyengar, A. K., Johnson, P., Krus, U., MacDonald, P. E., Mahajan, A., Manning Fox, J. E., ... McCarthy, M. I. (2020). Genetic variant effects on gene expression in human pancreatic islets and their implications for T2D. *Nature Communications*, 11(1), 4912. <https://doi.org/10.1038/s41467-020-18581-8>
- Wigger, L., Barovic, M., Brunner, A.-D., Marzetta, F., Schöninger, E., Mehl, F., Kipke, N., Friedland, D., Burdet, F., Kessier, C., Lesche, M., Thorens, B., Bonifacio, E., Legido-Quigley, C., Barber Saint Hilaire, P., Delerive, P., Dahl, A., Klose, C., Gerl, M. J., ... Solimena, M. (2021). Multi-omics profiling of living human pancreatic islet donors reveals heterogeneous beta cell trajectories towards type 2 diabetes. *Nature Metabolism*, 3(7), 1017–1031. <https://doi.org/10.1038/s42255-021-00420-9>
- Yengo, L., Vedantam, S., Marouli, E., Sidorenko, J., Bartell, E., Sakae, S., Graff, M., Eliassen, A. U., Jiang, Y., Raghavan, S., Miao, J., Arias, J. D., Graham, S. E., Mukamel, R. E., Spracklen, C. N., Yin, X., Chen, S.-H., Ferreira, T., Highland, H. H., ... Hirschhorn, J. N. (2022). A saturated map of common genetic variants associated with human height. *Nature*, 610(7933), 704–712. <https://doi.org/10.1038/s41586-022-05275-v>
- Zeggini, E., Hatzikotoulas, K., Tachmazidou, I., Steinberg, J., Hartwig, F. P., Southam, L., Hacking, S., Boer, C. G., Styrkarsdóttir, U., Gilly, A., Suveges, D., Killian, B., Ingvarsson, T., Jonsson, H., Babis, G. C., McCaskie, A., Utterlinden, A. G., Van Meurs, J. B. J., Thorsteinsdóttir, U., ... Zeggini, E. (2018). Genome-wide analyses using UK Biobank data provide insights into the genetic architecture of osteoarthritis. *Nature Genetics*, 50(4), 549–558. <https://doi.org/10.1038/s41588-018-0079-y>



Resource

Deciphering osteoarthritis genetics
across 826,690 individuals from 9 populations

1.3e-08 threshold

ORIGINAL INVESTIGATION

Evaluating the effective numbers of independent tests
and significant *p*-value thresholds in commercial genotyping
arrays and public imputation reference datasets

Boer et al. (2021)

Significance threshold

The testing of $M = 11$ osteoarthritis phenotypes in this study needed to be taken into account in the interpretation of genome-wide statistical significance. Applying a Bonferroni correction would be inherently conservative as this method assumes independence among the tests considered. Therefore, we first used LD Score regression method (Bulik-Sullivan et al., 2015a, 2015b) (<https://github.com/bulik/ldsc/>) with genome-wide meta-analysis summary statistics to estimate the genetic correlation matrix between the 11 osteoarthritis traits (Table S6) and then calculated the effective number of independent traits (M_{eff}) from the eigenvalues λ_i of the correlation matrix (Li et al., 2012):

$$M_{\text{eff}} = M - \sum i=1^M [I(\lambda_i > 1)(\lambda_i - 1)]$$

For the $M = 11$ osteoarthritis phenotypes in this study, $M_{\text{eff}} = 4.6565$. The threshold corrected for the effective number of traits to report genome-wide significance is $p < 1.3 \times 10^{-8}$.

Li et al. (2012)



6 vs. 11 Traits

```
> rg_mat <- read_csv("/home/dylan/comp/supp_tables/ldsc_rg_matrix.csv")[2:7,2:7] %>%  
+   as_tibble  
New names:  
• `` -> `...1`  
Rows: 8 Columns: 9  
─ Column specification ──────────────────────────────────────────────────  
Delimiter: ","  
chr (1): ...1  
dbl (8): All_OA, Knee_OA, Hip_OA, KneeHip OA, TKR, THR, TJR, T2D  
  
i Use `spec()` to retrieve the full column specification for this data.  
i Specify the column types or set `show_col_types = FALSE` to quiet this message.  
> rg_mat  
# A tibble: 6 × 6  
  All_OA Knee_OA Hip_OA `KneeHip OA`  TKR    THR  
  <dbl> <dbl> <dbl>      <dbl> <dbl> <dbl>  
1 0.944  1     0.570      0.943 0.945 0.440  
2 0.790  0.570  1         0.817 0.509 0.963  
3 0.983  0.943  0.817     1     0.886 0.720  
4 0.886  0.945  0.509     0.886 1     0.431  
5 0.676  0.440  0.963     0.720 0.431 1  
6 0.908  0.792  0.891     0.941 0.819 0.873  
> eigen <- eigen(rg_mat)$values  
> M <- eigen %>% length  
> M_eff <- M - sum((eigen > 1) * (eigen - 1))  
> alpha <- 5e-8/M_eff  
> alpha  
[1] 2.291086e-08
```

Probably could have gone with 2.3e-8, by likely Type II error is more of a concern here.

[Back to presentation](#)



ARTICLE

<https://doi.org/10.1038/s41467-020-20885-8>

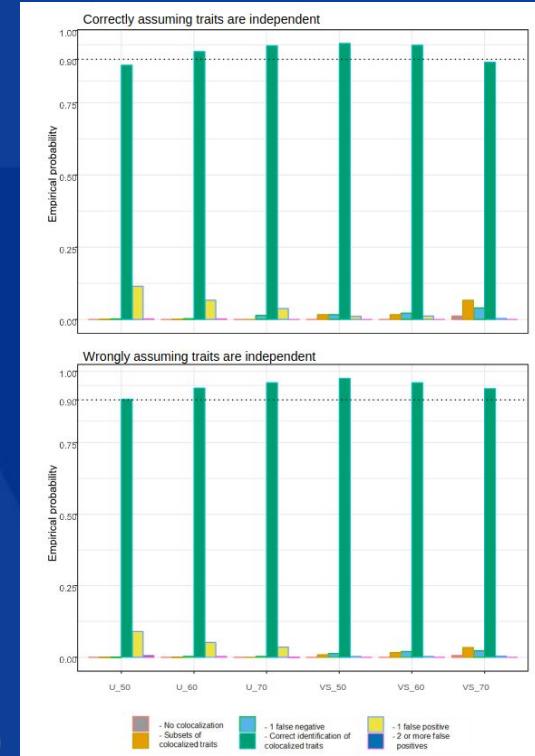
OPEN

A fast and efficient colocalization algorithm for identifying shared genetic risk factors across multiple traits

Christopher N. Foley^{1,2}, James R. Staley^{2,3}, Philip G. Breen⁴, Benjamin B. Sun², Paul D. W. Kirk¹, Stephen Burgess^{1,2} & Joanna M. M. Howson^{1,2,5,6}

Fig. S11 and Supplementary Table S2). Our results indicated that scenario (a), i.e. ignoring all correlation by treating studies as independent and traits as a-priori exchangeable, even when there is complete sample overlap (i.e. participants are the same in all studies), gives reasonable results and in our assessment was comparable to scenario (c) (Supplementary Fig. S10 and Tables S2–S3). We discuss the theoretical reasons for this in

Sample Overlap



Back to presentation

Supplementary Fig. S10