eurac research

Leveraging observational population-based studies for molecular target identification

Cristian Pattaro

Eurac Research, Institute for Biomedicine, Bolzano, Italy

Milano, May 8th, 2026

cristian.pattaro@eurac.edu @cpattaro.bsky.social

Outline

1) Contextualization
2) Causality. Instrumental variable and Mendelian randomization analysis
3) Examples
    a) Mirror game! Phospholipids, BMI and Atrial Conduction
    b) Banana skins. An example on complement activation
    c) Give me hope! A molecular target for desmoplakin
4) Conclusions

Contextualization

Observational population-based studies and biobanks

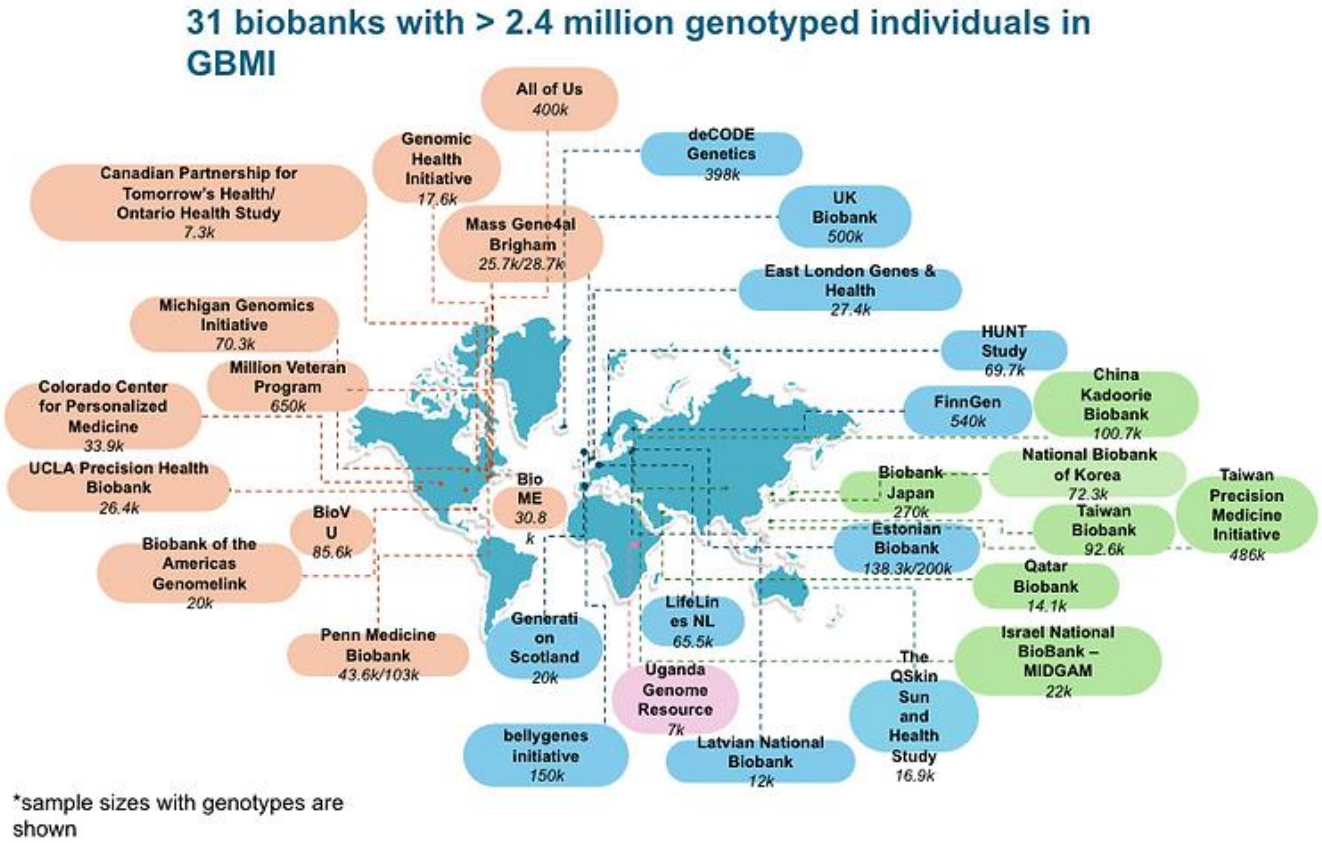
AMERICAN JOURNAL OF PUBLIC HEALTH *March, 1951*

Epidemiological Approaches to Heart Disease: The Framingham Study*

THOMAS R. DAWBER, M.D., GILCIN F. MEADORS, M.D., M.P.H., AND FELIX E. MOORE, JR.

National Heart Institute, National Institutes of Health, Public Health Service, Federal Security Agency, Washington, D. C.

THE use of the word "epidemiology" and the concept of what epidemiology as a discipline may encompass has varied widely since the days of Peter Panum and John Snow. There are today many differing definitions of the word, cal diagnosis. Thus, today, the epidemiological approach is used to explore certain relationships in health and disease which, with present technological methods, cannot be observed directly. In addition to the many studies of the infec-



https://www.globalbiobankmeta.org/

Clinical information

Socio-demographic data

Genomics

Proteomics

Glycomics

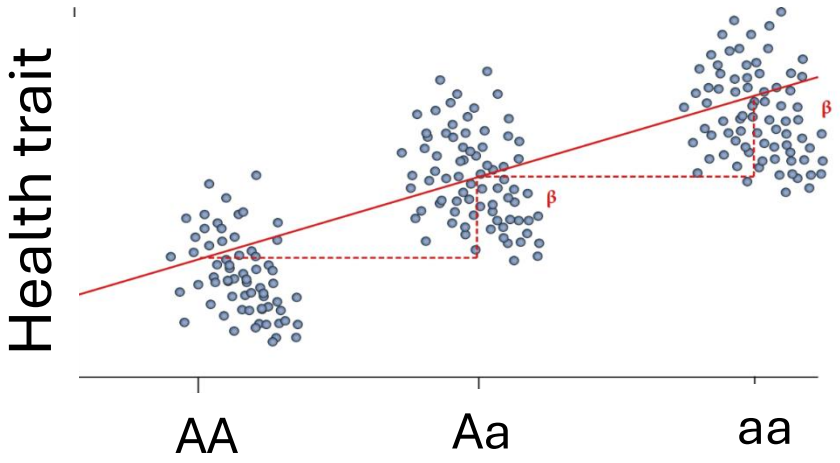
Exposomics

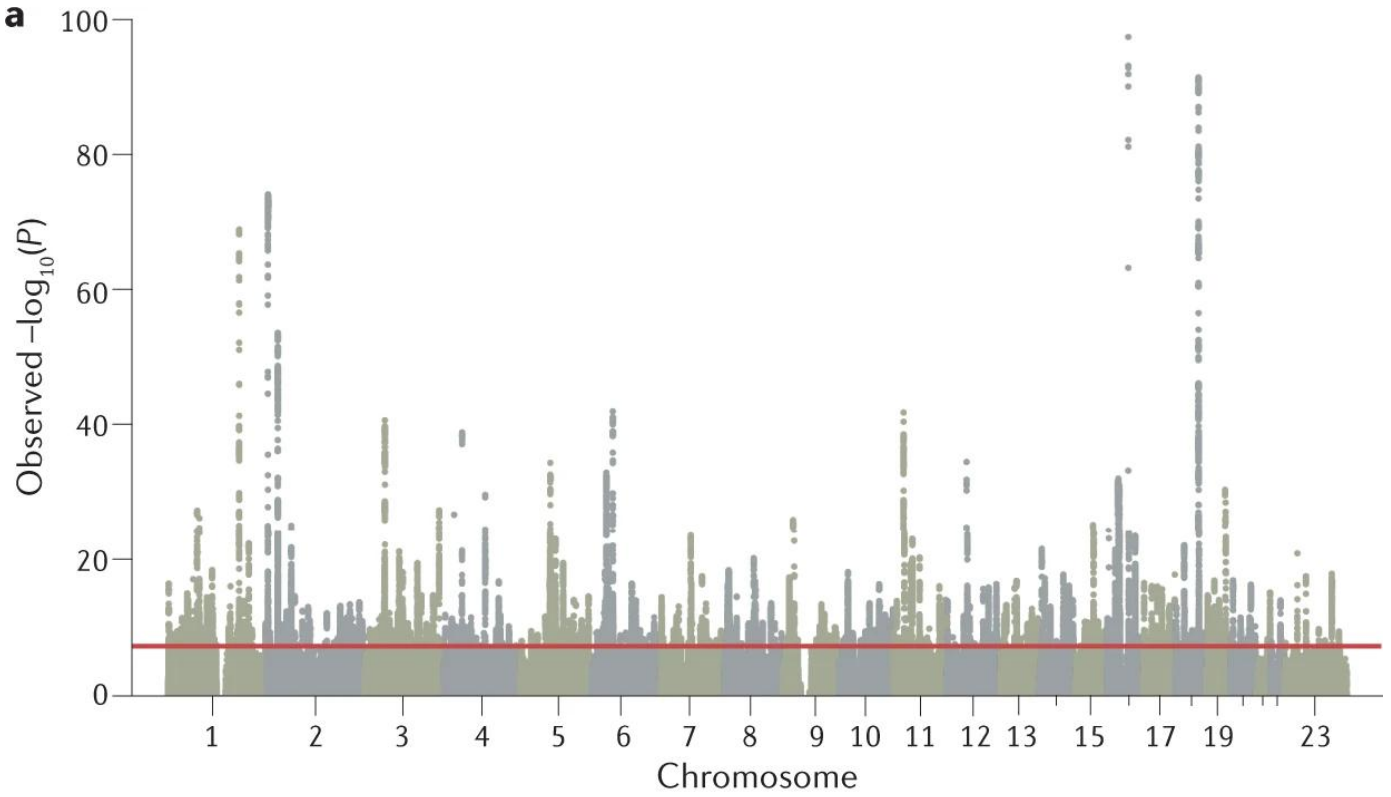
Transcriptomics

Methylomics

Metabolomics

Genome-wide association studies (GWAS)





adapted from O'Seaghdha & Fox, Nat Rev Nephrol (2012)

Uffelmann et al, Nat Rev Methods Primers 2021; 1:59

Proteomics

1

nature

Explore content · About the journal · Publish with us

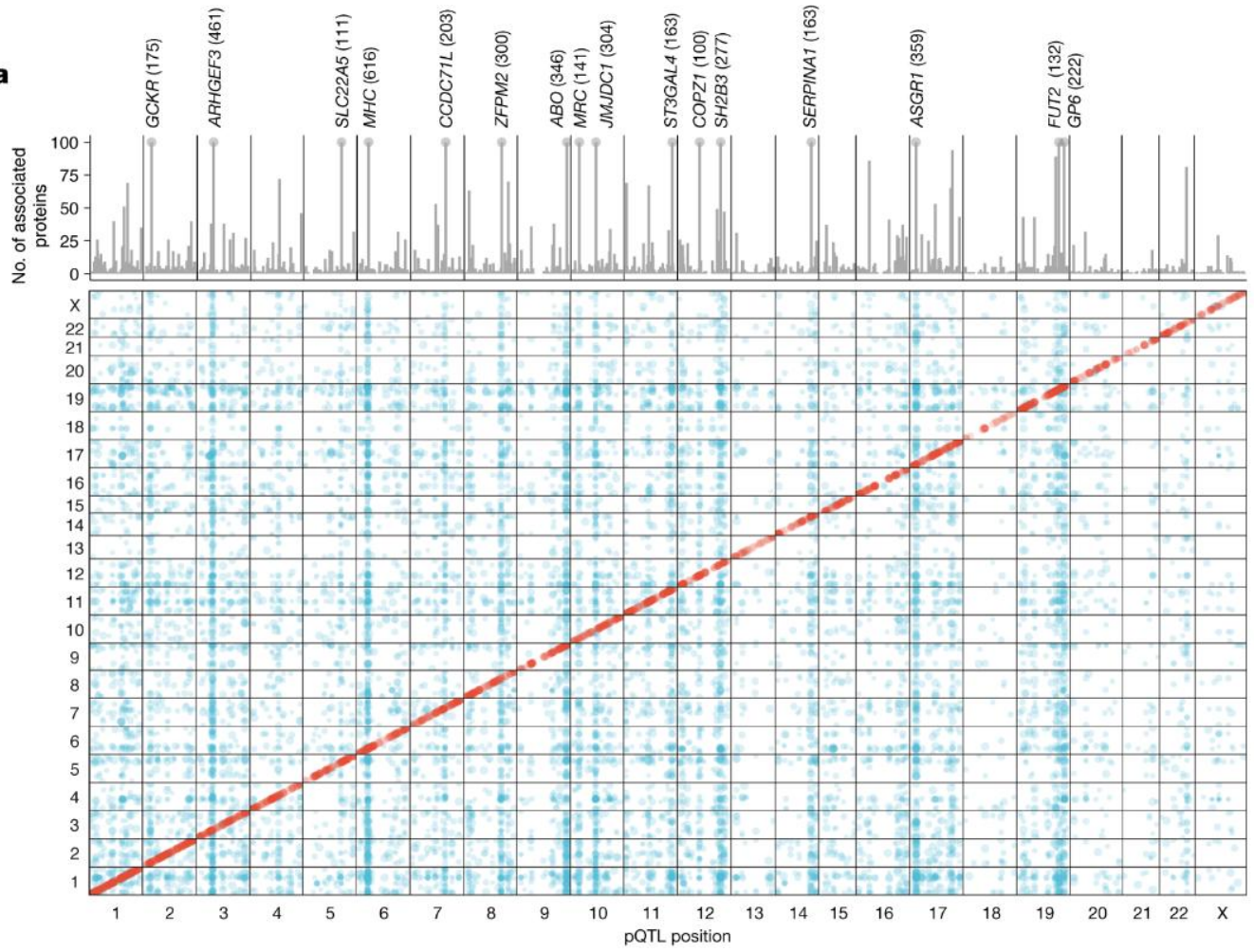
nature › articles › article

Article | Open access | Published: 04 October 2023

Plasma proteomic associations with genetics and health in the UK Biobank

Benjamin B. Sun, Joshua Chiou, Matthew Traylor, Christian Benner, Yi-Hsiang Hsu, Tom G. Richardson, *Nature* **622**, 329–338 (2023) | Cite this article

GWAS of 2922 protein targets measured in plasma from 54,219 individuals



2

nature genetics

Explore content · About the journal · Publish with us · Subscribe

nature › nature genetics › articles › article

Article | Published: 02 December 2021

Large-scale integration of the plasma proteome with genetics and disease

Egil Ferkingstad, Patrick Sulem, Bjarni A. Atlason, Gardar Sveinbjornsson, Magnus I. Magnusson, Edda L. *Nature Genetics* **53**, 1712–1721 (2021) | Cite this article

GWAS of 4907 protein targets measured in plasma from 35,559 Icelandic individuals

tested plasma protein levels for association with 373 diseases and other traits and identified 257,490 associations, identifying 938 genes encoding potential drug targets

3

Science

Current Issue · First release papers · Archive · About · Submit manus

HOME › SCIENCE › VOL. 374, NO. 6569 › MAPPING THE PROTEO-GENOMIC CONVERGENCE OF HUMAN DISEASES

RESEARCH ARTICLE | DISEASE GENOMICS

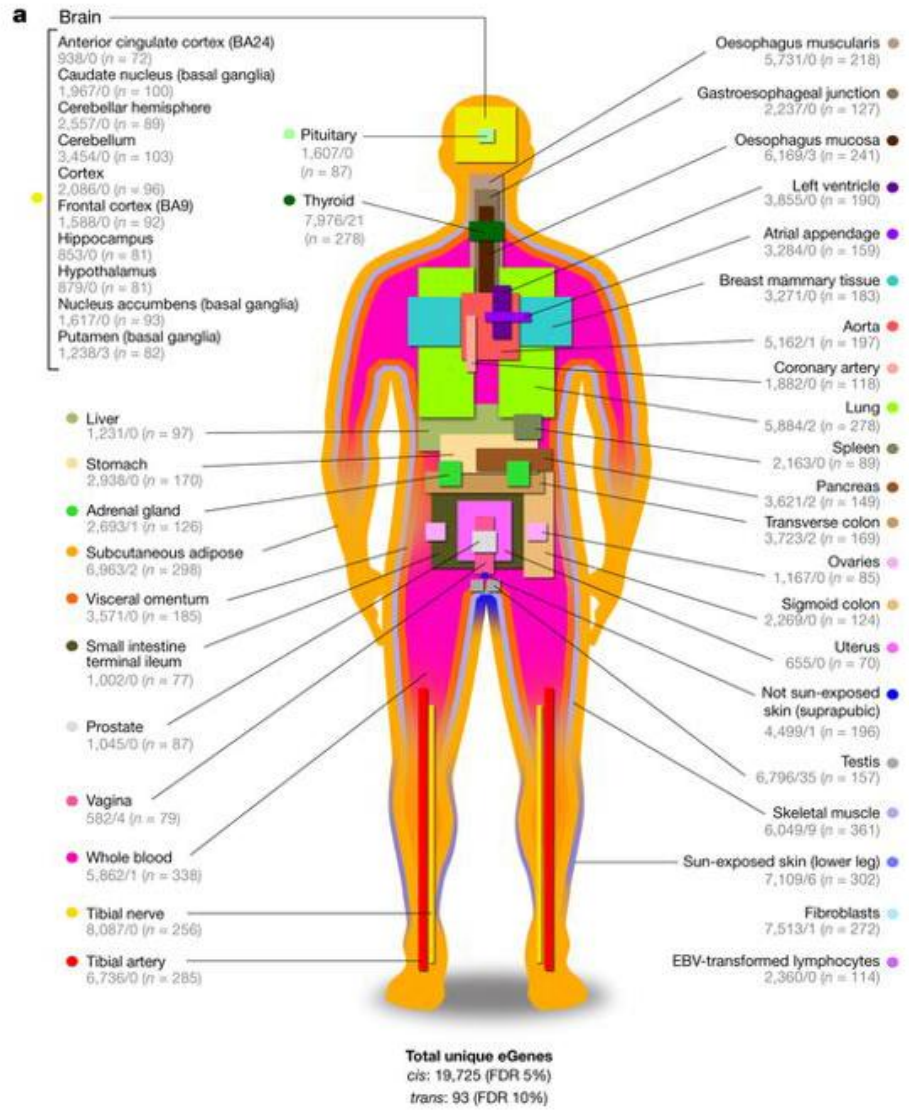
Mapping the proteo-genomic convergence of human diseases

MAIK PIETZNER, ELEANOR WHEELER, JULIA CARRASCO-ZANINI, ADRIAN CORTES, MINE KOPRULU, MARIA A. WÖRHEIDE, ERIN OERTON, JAMES COOK, ISOBEL D. STEWART, [...] CLAUDIA LANGENBERG +12 authors Authors Info & Affiliations

SCIENCE · 14 Oct 2021 · Vol 374, Issue 6569 · DOI: 10.1126/science.abj1541

GWAS of 4775 protein targets measured in plasma from 10,708 individuals

Tissue biobanks



GTEx Consortium, *Nature* 2017: **550**: 204-13

GTEx Portal

https://www.gtexportal.org/home/documentationPage

GTEx is a database of genotype ←→ gene expression associations

- Measured expression of >30,000 genes
- in 53 human tissues
- Sample size: 70 to ~500 samples per tissue

Identification of SNVs associated with gene expression

- within each tissue/organ
- within each cell type within organ

Strengths

Data richness & depth. Large sample size.

Limitations

Observational. Prone to selection bias.

?

Can they be used to identify molecular targets for treatment & prevention? How?


Baby
Yes
No
Stork
Yes
N_{11}
N_{12}
No
N_{21}
N_{22}

(unobserved)
confounder

Exposure

The question is not on the association. It is on the direction of the effect. We question the causality issue.

Outcome

?

Intermediate outcome

Outcome

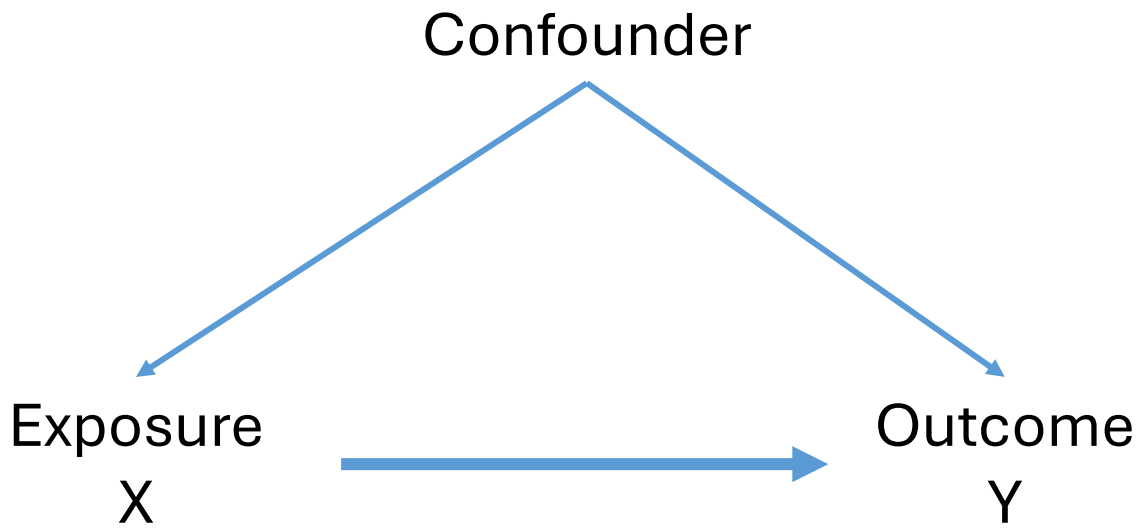
Exposure

Observational studies:

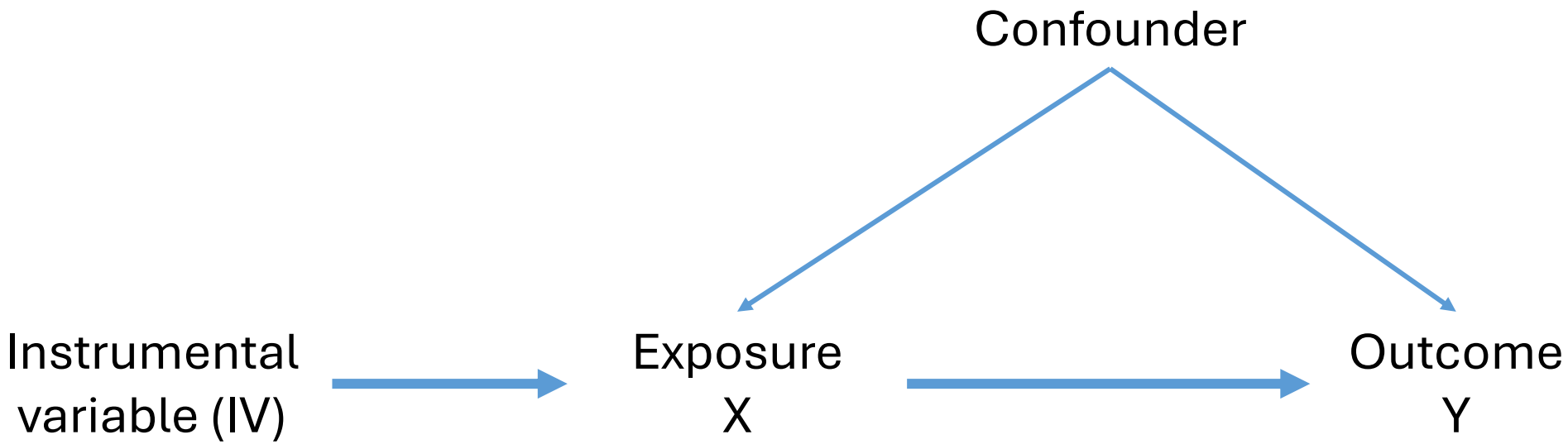
- Suitable to assess associations
- Limited possibilities to assess causality

Causality. Instrumental variable and Mendelian randomization analysis

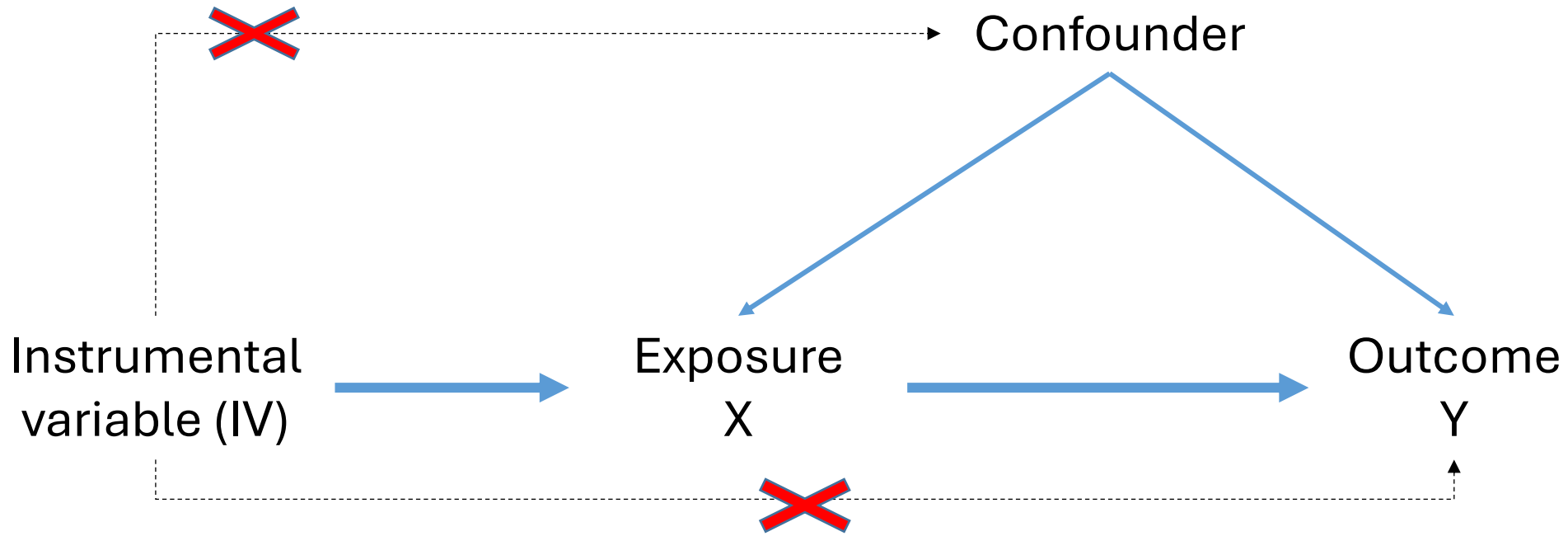
Instrumental variable analysis



Instrumental variable analysis



Instrumental variable analysis

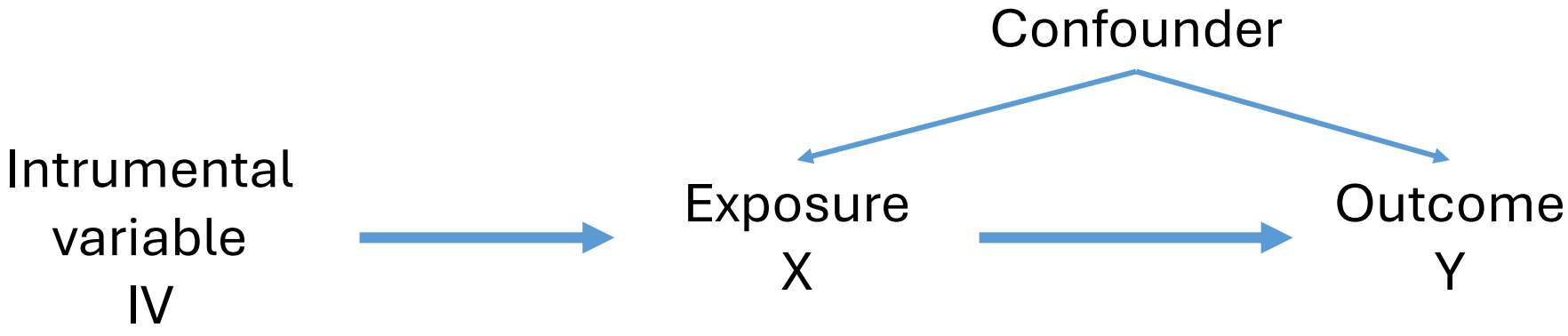


Core assumptions

1) IV is **strongly** associated with the Exposure X
2) The **association between IV and X is not confounded**
3) The is **no independent pathway** from IV to the Outcome Y other than through X

Instrumental variable analysis

the 2-stage least square (2SLS) regression



$$\boldsymbol{Y} = \alpha + \beta \boldsymbol{X} + \varepsilon$$

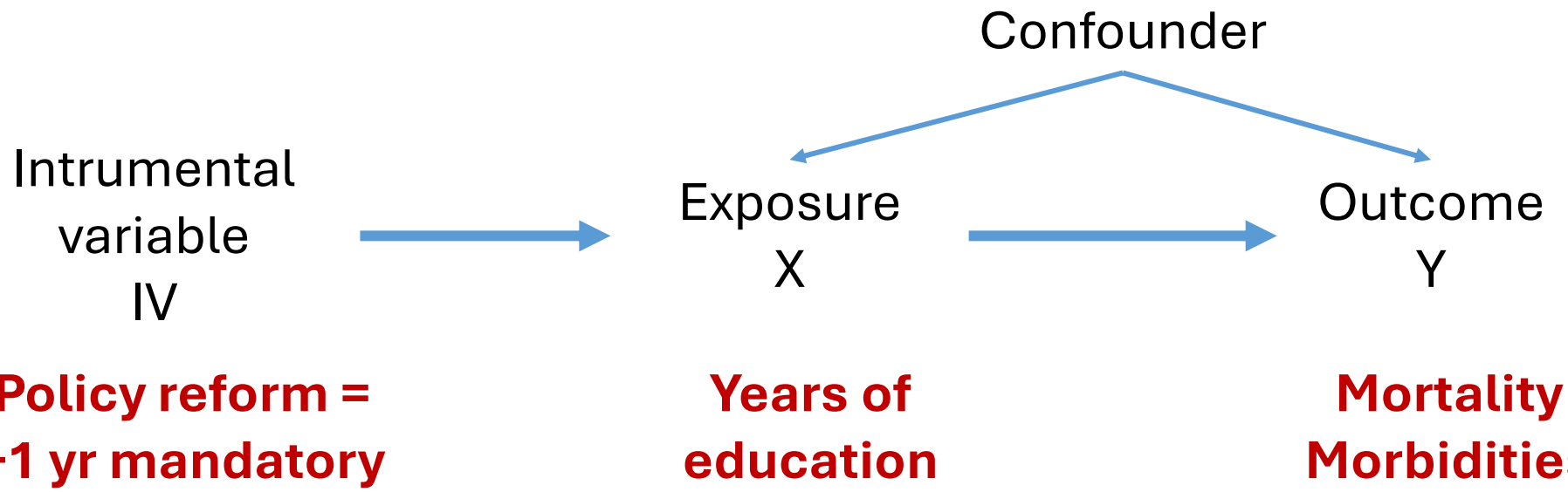
STAGE 1: $\boldsymbol{X} = \delta + \gamma \boldsymbol{IV} + u$

$$\widehat{\boldsymbol{X}} = \hat{\delta} + \hat{\gamma} \boldsymbol{IV}$$

STAGE 2: $\boldsymbol{Y} = \alpha' + \beta' \widehat{\boldsymbol{X}} + \varepsilon'$

Instrumental variable analysis

the 2-stage least square (2SLS) regression - EXAMPLE



Int J Epidemiol. 2023 Dec 25;52(6):1878-1886. doi: 10.1093/ije/dyad104.

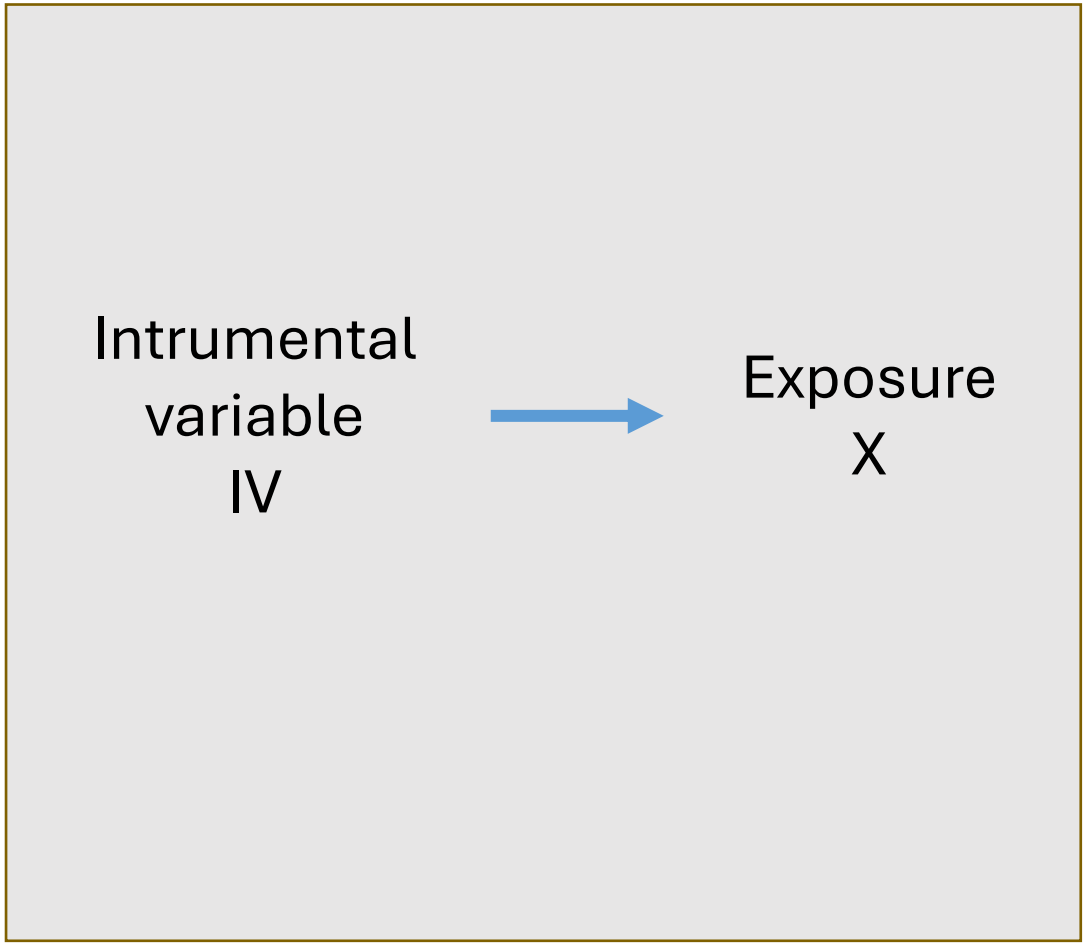
The causal effects of education on adult health, mortality and income: evidence from Mendelian randomization and the raising of the school leaving age

Neil M Davies [1 2 3 4], Matt Dickson [5], George Davey Smith [4 6], Frank Windmeijer [4 7], Gerard J van den Berg [8 9]

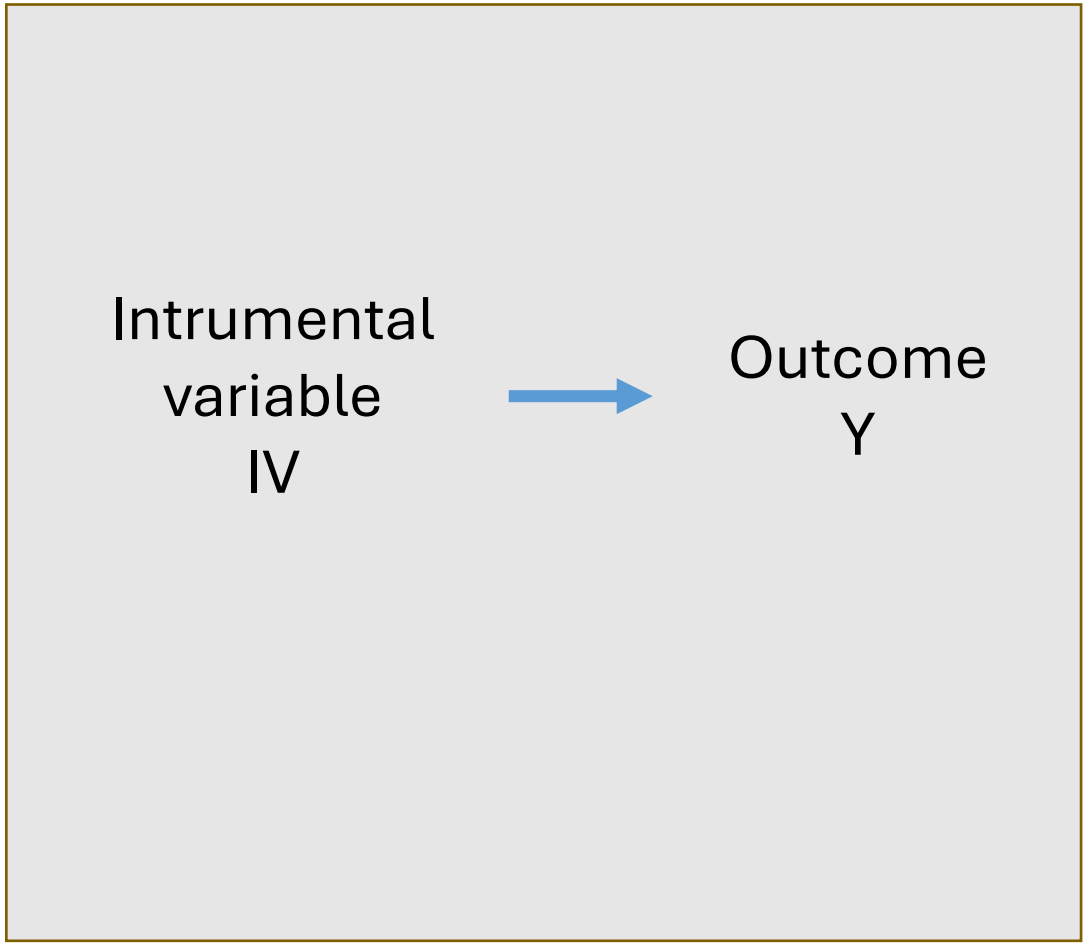
Instrumental variable analysis

the 2-sample case

Dataset 1

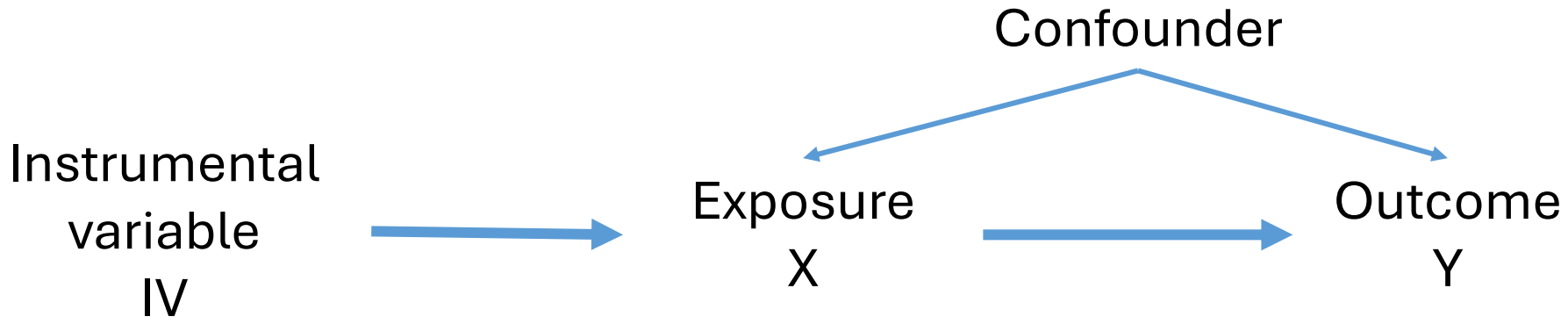


Dataset 2



2-sample instrumental variable analysis

the Wald-ratio estimator (W) for summary data



$$\boldsymbol{Y} = \alpha + \beta \boldsymbol{X} + \varepsilon$$

1. $\boldsymbol{X} = \delta + \gamma \boldsymbol{IV} + \varepsilon$ *(Stage 1)*

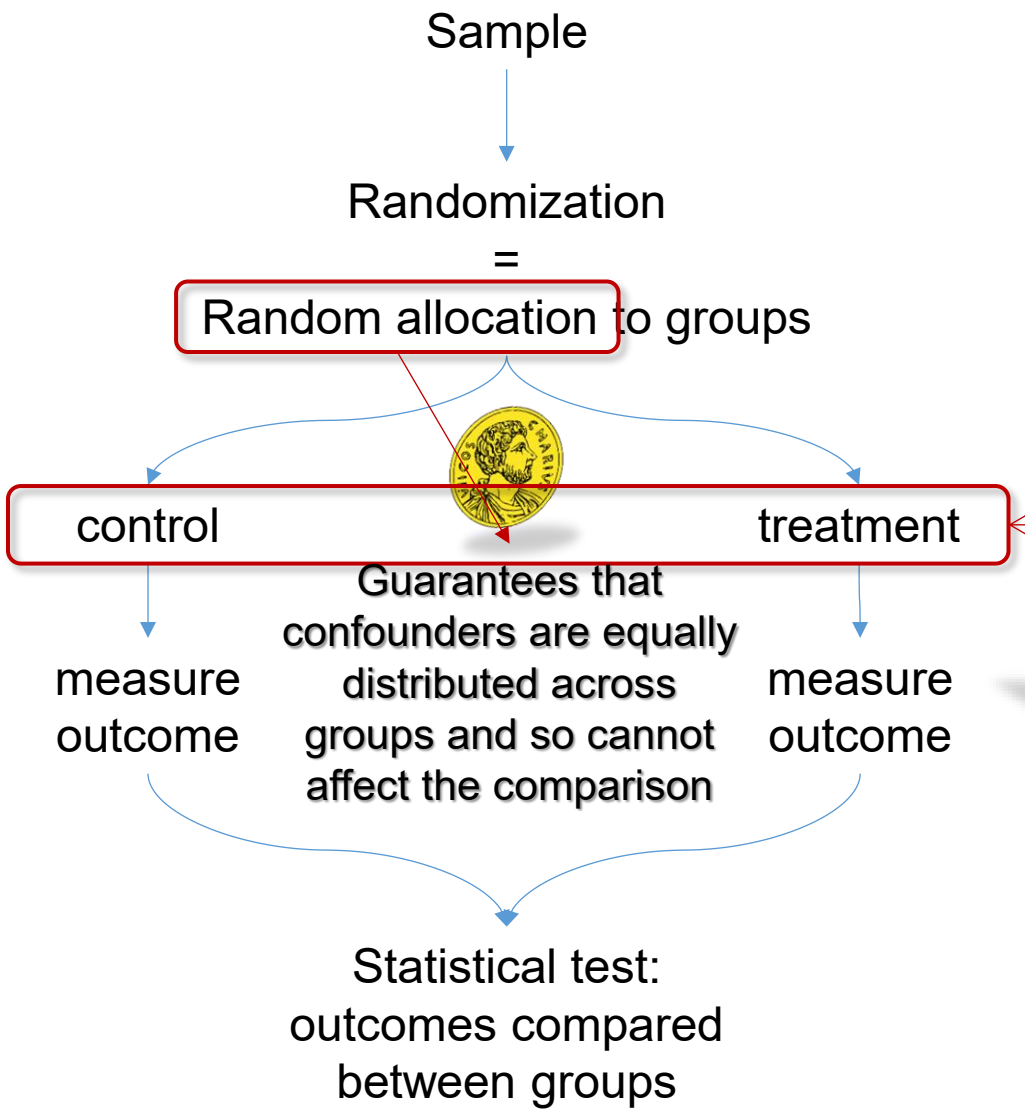
2. $\boldsymbol{Y} = \eta + \theta \boldsymbol{IV} + u$

3. $\hat{\beta} = W = \frac{\hat{\theta}}{\hat{\gamma}}$

The variance can be estimated via delta method, using a Taylor series expansion:

$$var(\hat{W}) \approx \frac{var(\hat{\theta})}{\hat{\gamma}^2} + -2\frac{\hat{\theta}}{\hat{\gamma}^3} cov(\hat{\theta}, \hat{\gamma}) + \frac{\hat{\theta}^2}{\hat{\gamma}^4} var(\hat{\gamma})$$

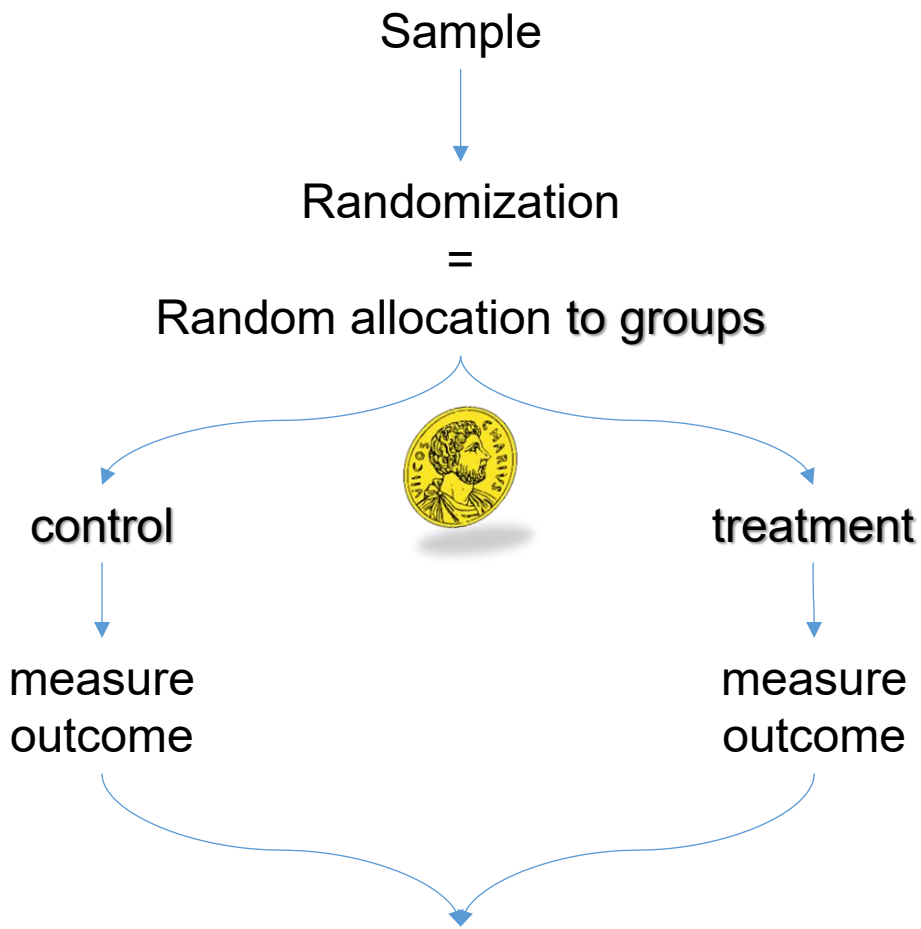
Randomized controlled trial



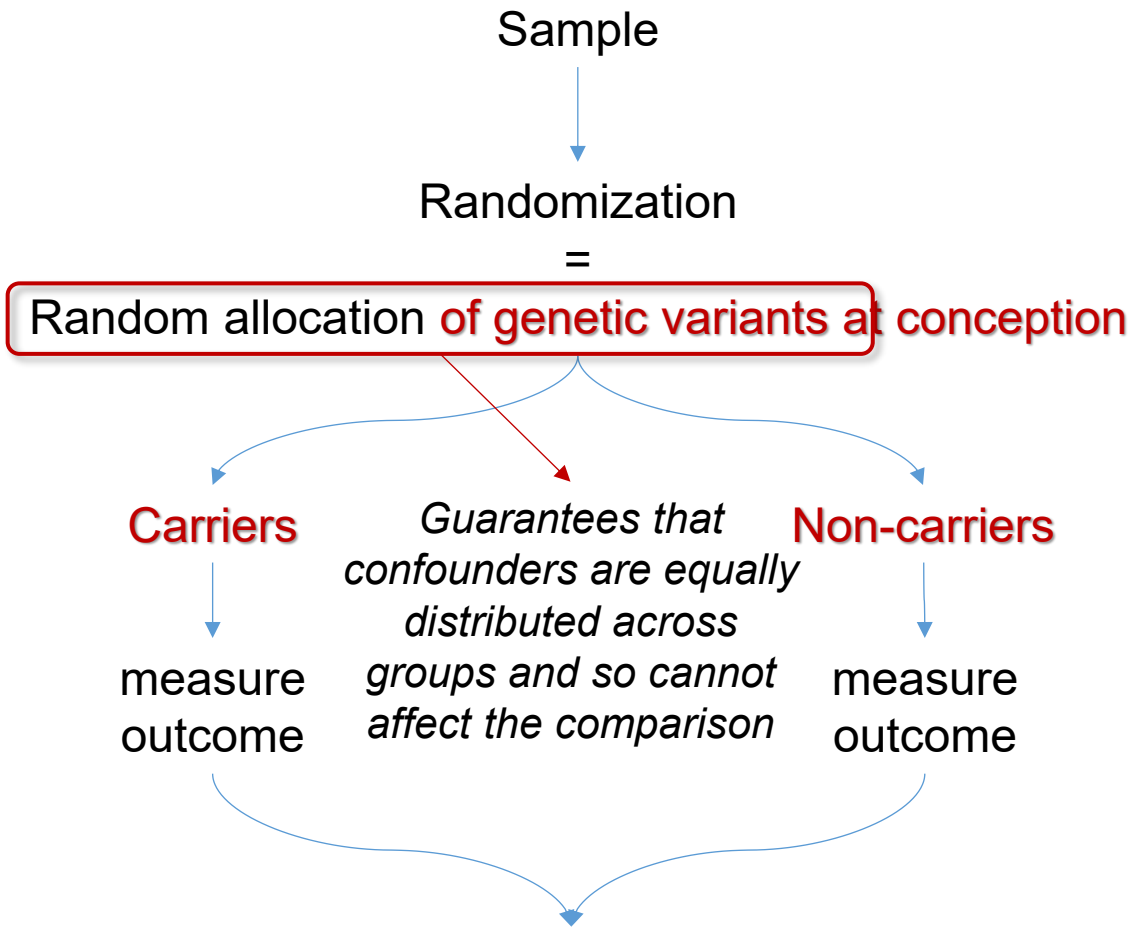
intervention

Typically, the target of the intervention is a biological mechanism, e.g. reducing or increasing the expression of a gene (amount of gene produced), reducing or increasing the levels of a certain protein, etc. (molecular targets)

Randomized controlled trial



Mendelian randomization

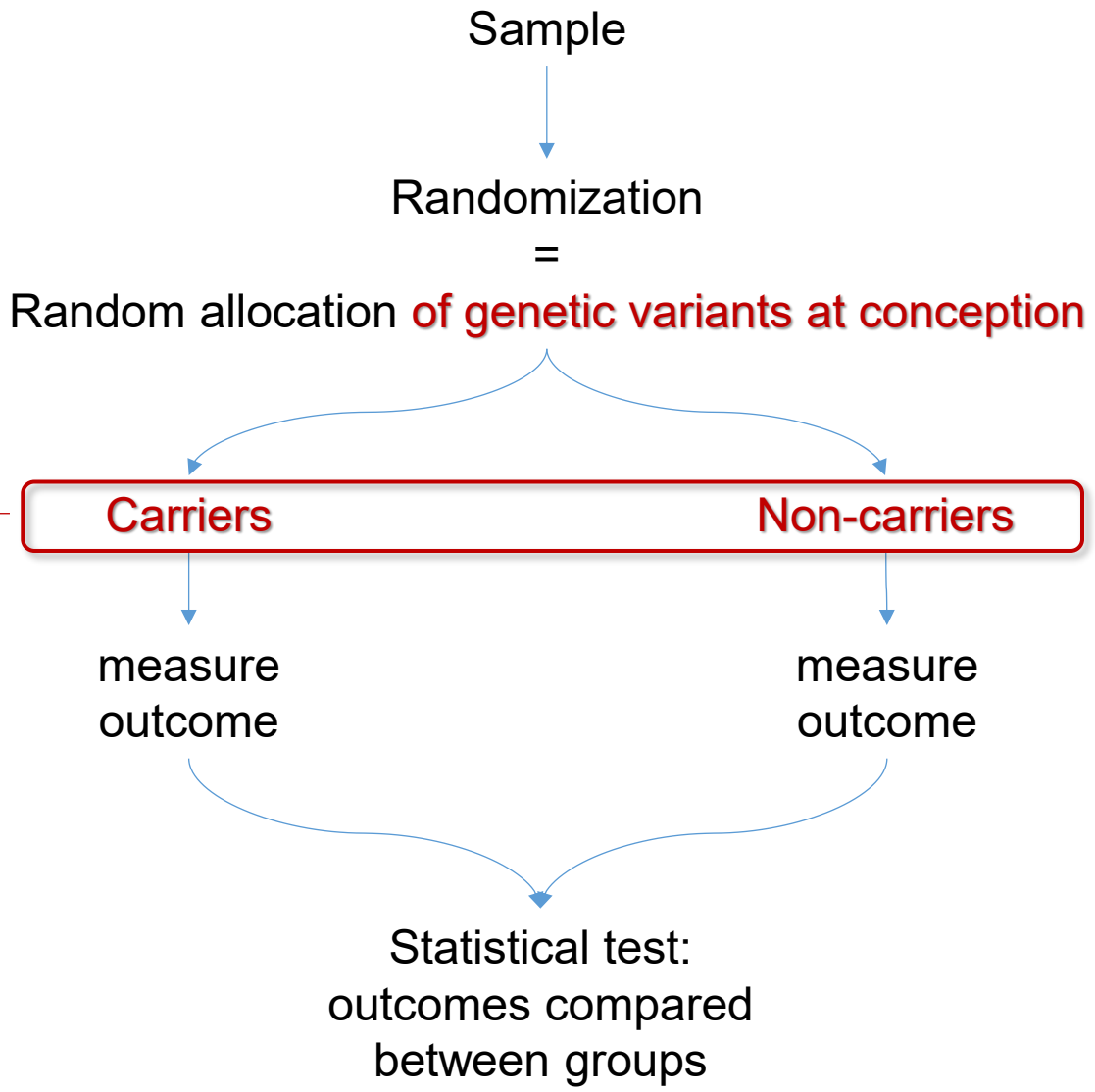


intervention

Different genetic status → different gene expression levels, protein levels, disease susceptibility → mimics the drug effect

similar to

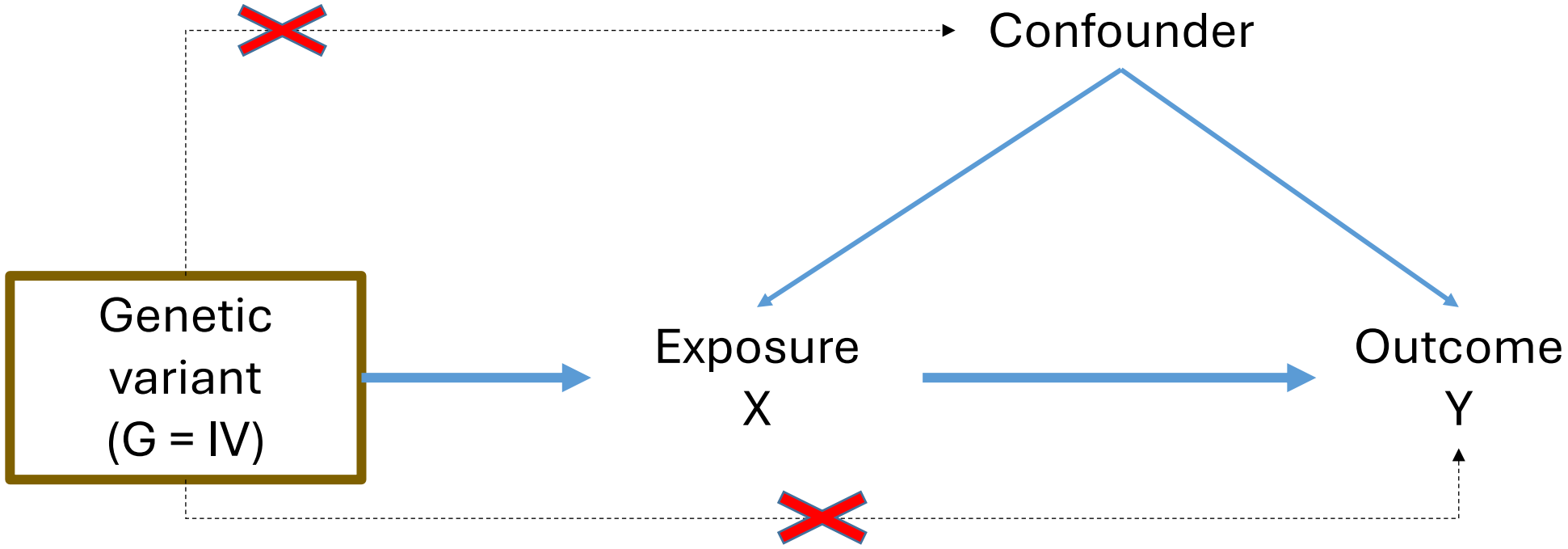
Mendelian randomization



Examples

1. Does lowering protein X reduce the risk of disease Y?
2. Does higher BMI increase risk of atrial fibrillation?
3. Does knock-out of ANGPTL3 gene reduces the risk of Coronary Artery Disease?

Mendelian randomization = instrumental variable analysis technique



Core assumptions

1) G is **strongly** associated with the Exposure X
2) The **association between G and X is not confounded**
3) The is **no independent pathway** from G to the Outcome Y other than through X

Examples

Mirror game!

Phospholipids, BMI and Atrial Conduction

Lipidomics, Atrial Conduction, and Body Mass Index

Evidence From Association, Mediation, and Mendelian Randomization Models

Fabiola Del Greco M ✉, **Luisa Foco**, **Alexander Teumer**, **Niek Verweij**, **Giuseppe Paglia**, **Viviana Meraviglia**, **Roberto Melotti**, **Vladimir Vukovic**, **Werner Rauhe**, **Peter K. Joshi**, **Ayse Demirkan**, **Stephan B. Felix**, **Maik Pietzner**, **M. Abdullah Said**, **Yordi J. van de Vegte**, **Pim van der Harst**, **Lifelines Cohort Study***, **Alan F. Wright**, **Andrew A. Hicks**, **Harry Campbell**, **Marcus Dörr**, **Harold Snieder**, **James F. Wilson**, **Peter P. Pramstaller** ✉, **Alessandra Rossini**, **Cristian Pattaro** ✉ **See fewer authors**

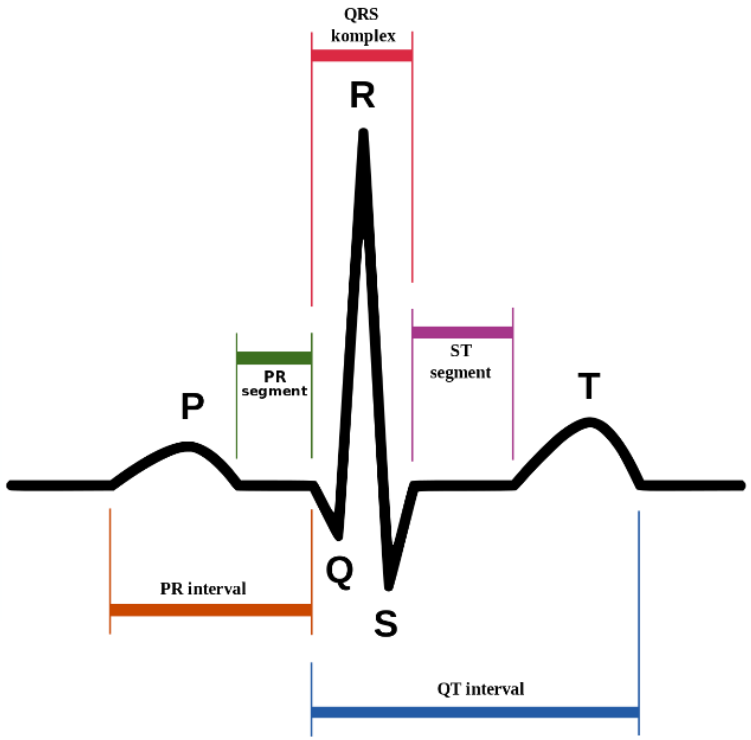
Originally published 15 Jul 2019 | https://doi.org/10.1161/CIRCGEN.118.002384 | Circulation: Genomic and Precision Medicine. 2019;12:e002384

Fabiola Del Greco

Alessandra Rossini

Luisa Foco

- **Serum lipid levels** are suspected to have a causal role on atrial fibrillation (they should be proarrhythmic influencers).
- However, limited knowledge on the specific mechanisms connecting lipid alterations with **atrial conduction**.
- **P wave** = reliable non-invasive marker atrial conduction, associated with atrial fibrillation (AF) risk



By SinusRhythmLabels.svg: Created by Agateller (Anthony Atkielski), converted to svg by atom.derivative work: Kychot (talk) - SinusRhythmLabels.svg, Copyrighted free use, https://commons.wikimedia.org/w/index.php?curid=7875754

Population-based studies

ORCADES, Orkney islands (**REPLICATION**)

N = 951

P wave from 10s ECG

Age 53(15)

F: 56%

MICROS, South Tyrol (**DISCOVERY**)

N = 839

P wave from 10s ECG

Age 44 (16)

F: 52%

151 sphingo & phospho-lipids

measured in the same lab

$$\text{P-wave duration} \sim \alpha_i + \beta_i L_i + \varepsilon_i$$

$$i=1..151$$

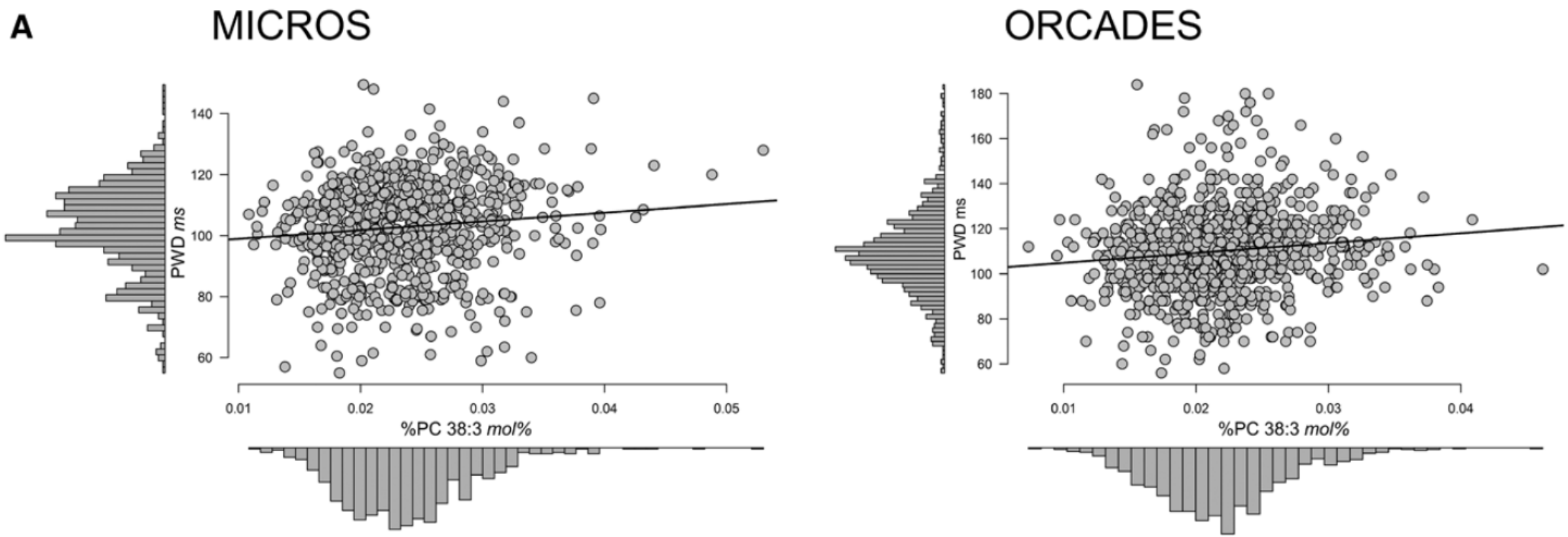
Del Greco et al, Circ Genom Prec Med 2019

Lipid	Discovery: MICROS		Replication: ORCADES		Meta-Analysis		
	b (SE)	2-Sided *P*-value	b (SE)	1-Sided *P*-value	b (SE)	2-Sided *P*-value	I^2%
%PE 36:3	−1.25 (0.40)	1.8×10^{-3}	−0.81 (0.53)	0.063	−1.09 (0.32)	6.4×10^{-4}	0
%PE 34:2	−0.57 (0.18)	2.1×10^{-3}	−0.23 (0.22)	0.148	−0.42 (0.14)	2.5×10^{-3}	30
%PE 38:4	0.55 (0.19)	3.0×10^{-3}	0.07 (0.20)	0.361	0.33 (0.14)	0.016	67
%PE 40:6	1.05 (0.37)	4.6×10^{-3}	0.35 (0.28)	0.109	0.61 (0.22)	6.9×10^{-3}	56
%PLPE 16:0/18:2	−1.26 (0.46)	6.5×10^{-3}	−0.68 (0.52)	0.097	−1.00 (0.35)	3.8×10^{-3}	0
%PC 30:1	−16.32 (6.01)	6.7×10^{-3}	−1.50 (9.47)	0.437	−12.06 (5.08)	0.017	43
%PE 40:5	3.91 (1.47)	7.9×10^{-3}	0.90 (1.67)	0.294	2.60 (1.11)	0.019	45
%PC 38:3*	2.27 (0.86)*	8.5×10^{-3}*	3.08 (1.06)*	1.9×10^{-3}*	2.59 (0.67)*	1.1×10^{-4}*	0*
%PE 38:3	3.00 (1.20)	0.013	2.25 (1.58)	0.078	2.73 (0.96)	4.4×10^{-3}	0
%PC 38:4	0.98 (0.41)	0.016	0.53 (0.46)	0.125	0.80 (0.30)	0.103	0
%PLPE 18:0/18:2	−0.65 (0.29)	0.024	−0.51 (0.31)	0.051	−0.58 (0.21)	5.7×10^{-3}	0
%PC 34:2	−0.37 (0.17)	0.025	−0.25 (0.19)	0.093	−0.32 (0.12)	0.010	0
%LPC 20:3	6.10 (2.89)	0.034	0.85 (2.46)	0.366	3.06 (1.88)	0.102	48
%LPC 18:0	0.45 (0.22)	0.042	0.59 (0.26)	0.012	0.51 (0.17)	2.7×10^{-3}	0
%PC 34:3	−6.84 (3.41)	0.045	−1.33 (3.01)	0.329	−3.75 (2.25)	0.097	32
%PC 40:6	3.06 (1.53)	0.045	1.53 (1.11)	0.083	2.06 (0.90)	0.022	0
%LPC 18:3	−16.49 (8.26)	0.046	−3.87 (5.91)	0.256	−8.14 (4.81)	0.090	35
%SPM 18:0	1.39 (0.71)	0.049	1.14 (0.78)	0.072	1.28 (0.53)	0.015	0

P-value < 0.05

1-sided P-value < 0.05/18

The association between PC38:3 and PWD was perfectly replicated in a completely independent population sample



Del Greco et al, Circ Genom Prec Med 2019

Adjustment for BMI did attenuate the PC38:3-PWD association in the same manner in both independent studies (45% effect attenuation), consistently across studies

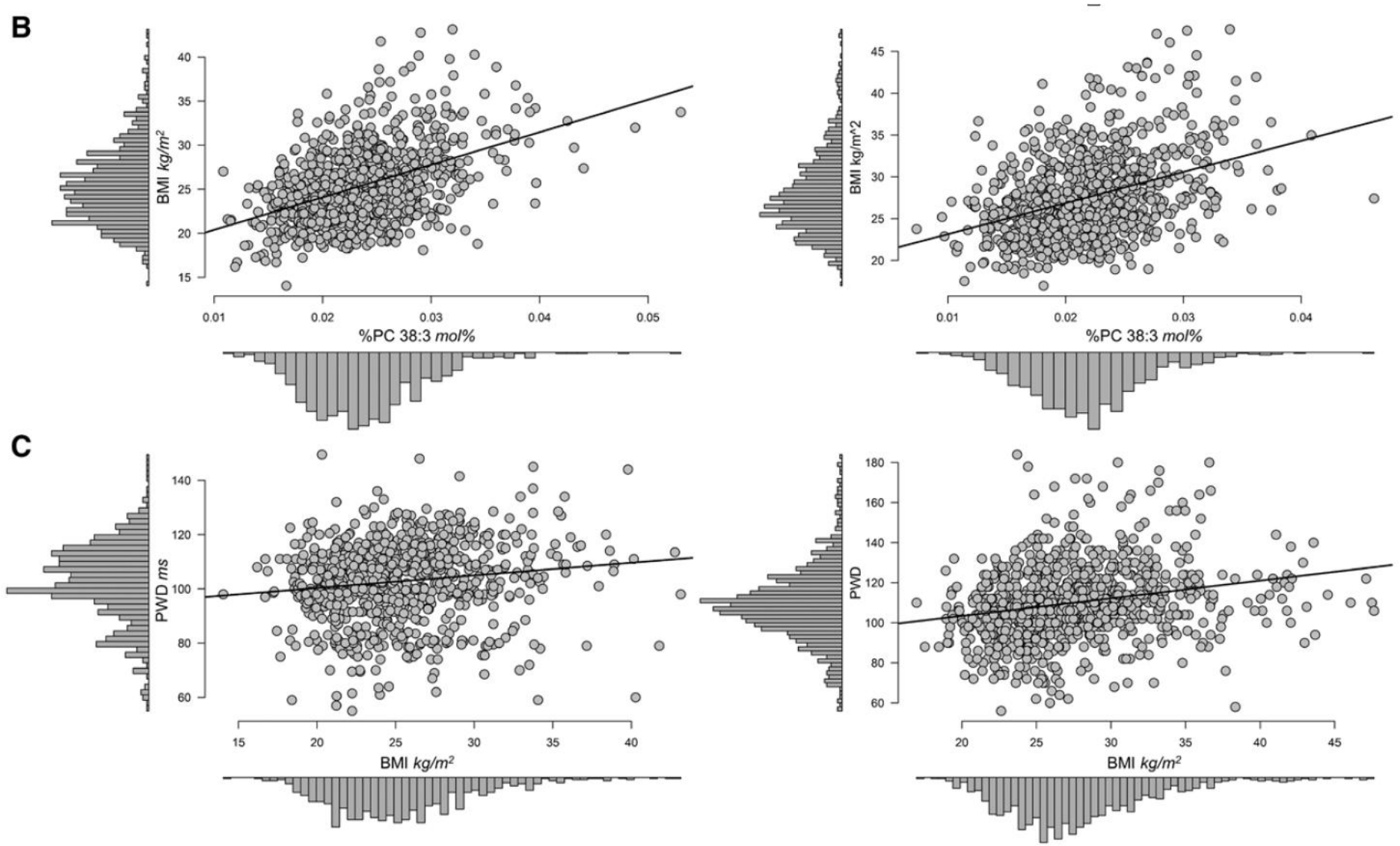
Potential Confounder	MICROS		ORCADES		Meta-Analysis*	
	b (SE)	*P*-value	b (SE)	*P*-value	b (SE)	*P*-value
BMI	1.41 (0.93)	0.131	1.50 (1.14)	0.188	1.45 (0.72)	0.045
LDL	2.13 (0.93)	0.022	2.80 (1.11)	0.012	2.41 (0.71)	7.4×10^{-4}
HDL	2.25 (0.90)	0.013	2.60 (1.12)	0.020	2.39 (0.70)	6.7×10^{-4}
TG	2.16 (0.89)	0.015	2.92 (1.11)	0.009	2.45 (0.69)	4.0×10^{-4}
TC	2.11 (0.88)	0.017	3.04 (1.08)	5.0×10^{-3}	2.48 (0.68)	2.8×10^{-4}
DBP	1.91 (0.87)	0.028	2.86 (1.07)	7.5×10^{-3}	2.29 (0.67)	7.0×10^{-4}
SBP	1.99 (0.87)	0.022	3.01 (1.07)	5.0×10^{-3}	2.40 (0.68)	3.9×10^{-4}
FGlu	2.11 (0.87)	0.015	3.01 (1.06)	5.0×10^{-3}	2.47 (0.67)	2.4×10^{-4}
DM	2.32 (0.86)	7.1×10^{-3}	3.11 (1.07)	3.6×10^{-3}	2.63 (1.32)	8.7×10^{-5}
Multivariable model†	1.45 (1.04)	0.160	1.58 (1.23)	0.200	1.50 (0.79)	0.058

Del Greco et al, Circ Genom Prec Med 2019

BMI was associated with both exposure PC38:3 and outcome PWD → candidate to be either in the causal pathway or a confounder

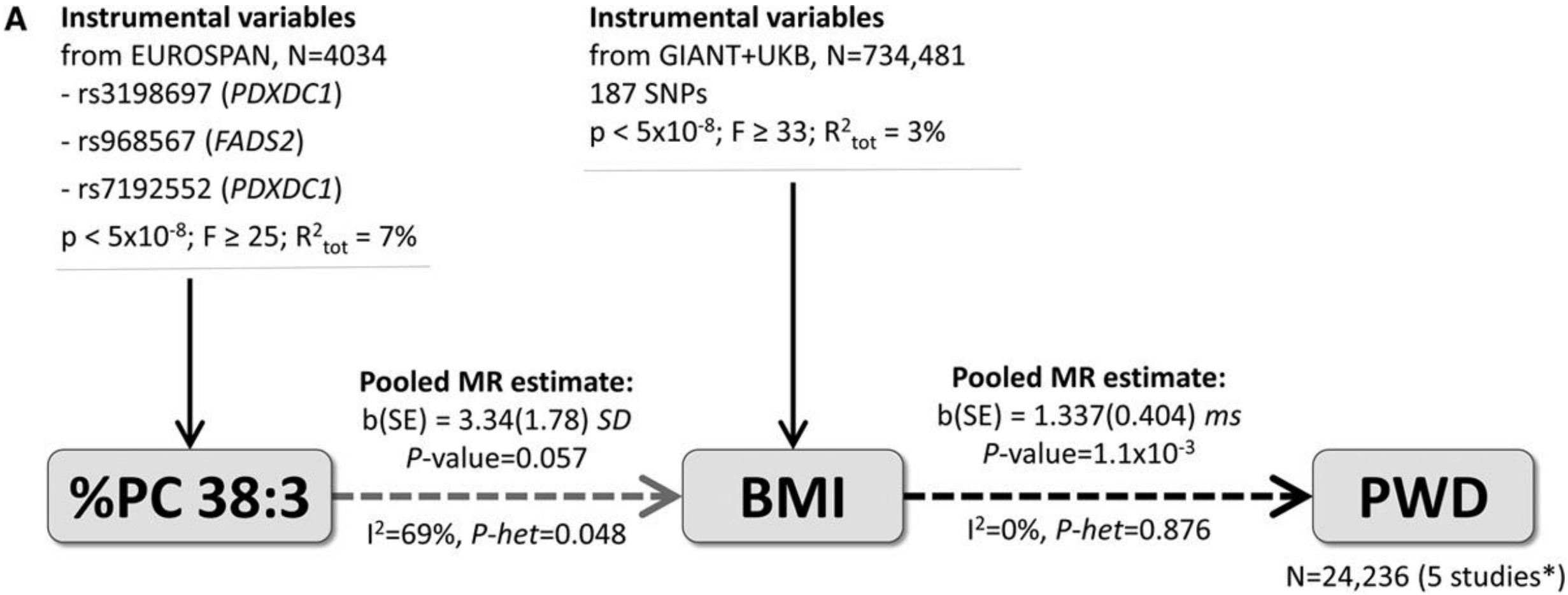
BMI ←→ PC38:3

BMI ←→ PWD

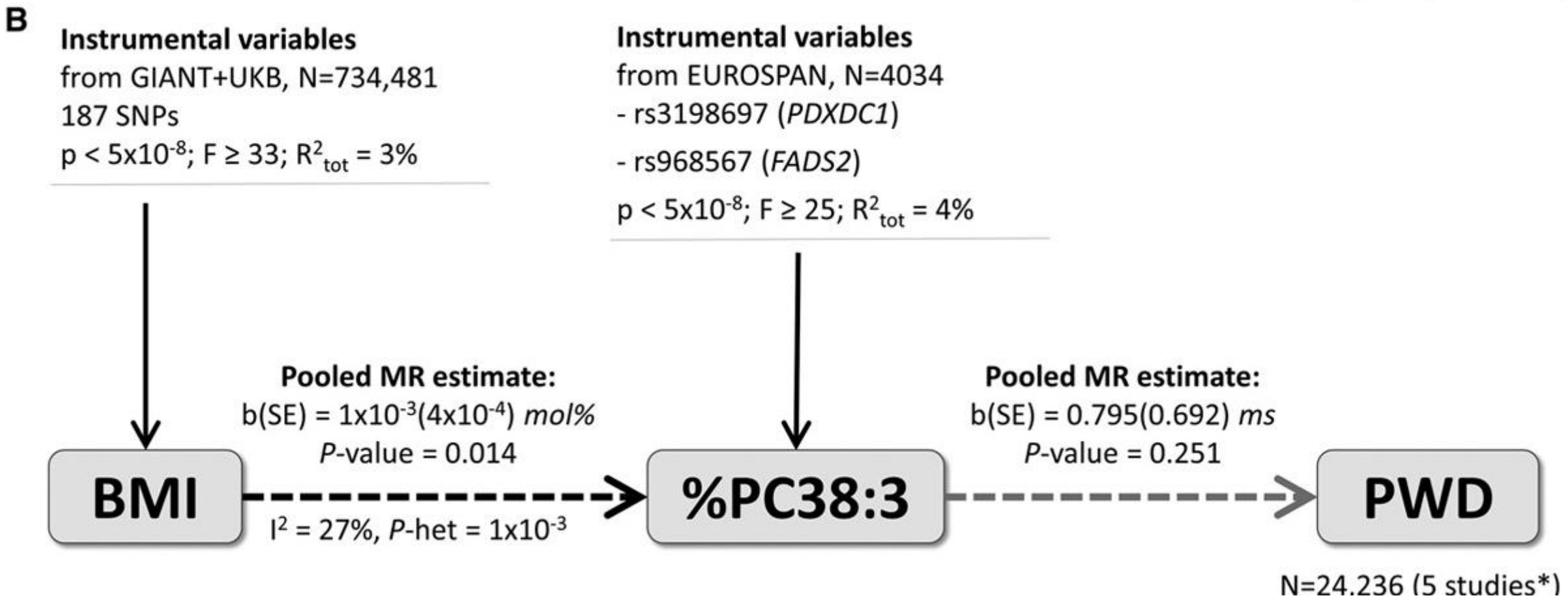


Del Greco et al, Circ Genom Prec Med 2019

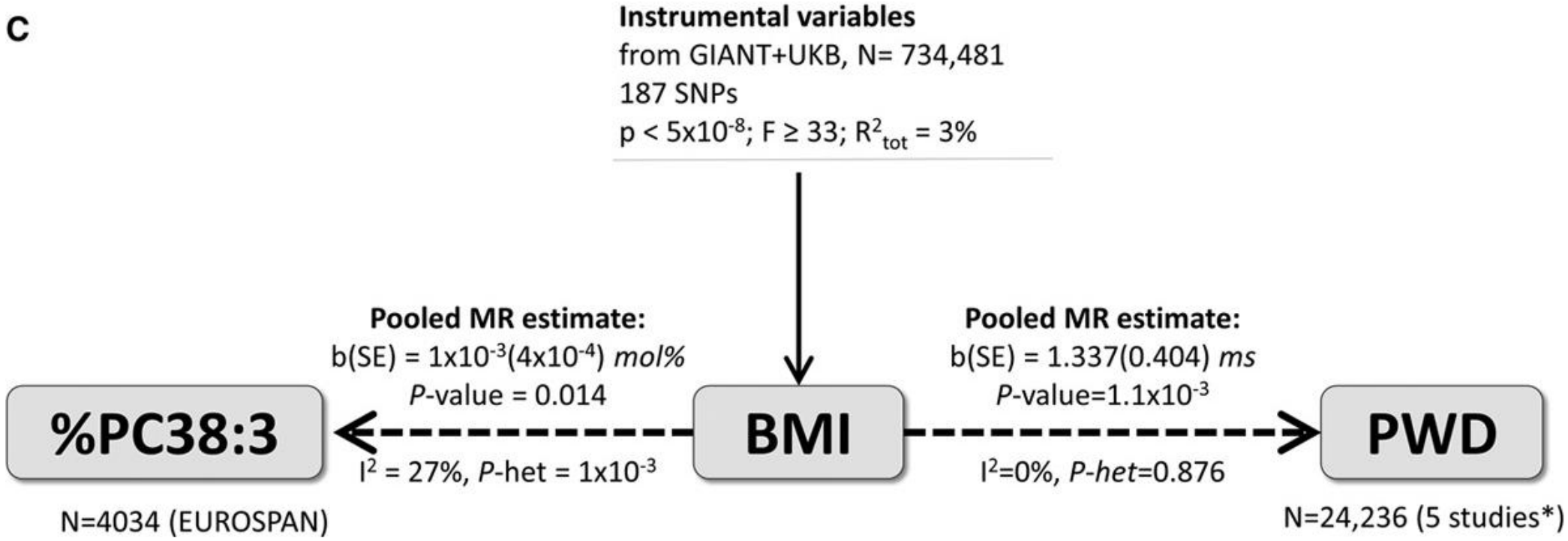
What is the most plausible causal model?



What is the most plausible causal model?



What is the most plausible causal model?

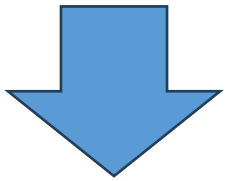


Results support causal effect of BMI on

both PWD (P=8.3×10^{-5}) and %PC 38:3 (P=0.014).

Conclusions

1. Despite “perfect” observational replication → no evidence of causal effect of %PC 38:3 on P-wave duration
2. BMI = the ideal confounder, causally associated with both %PC 38:3 and PWD
3. (BMI is a known causal risk factor for atrial fibrillation)



No laboratory follow-up experiment because: (1) PC38:3 not easily available on the market; (2) high human & instrumental resource investment (>1 year of work); (3) absence of causal evidence

Banana skins!

An example on the complement system

Damia Noce

Luisa Foco

Reinhard Würzner

Cell Reports
A Cell Press journal

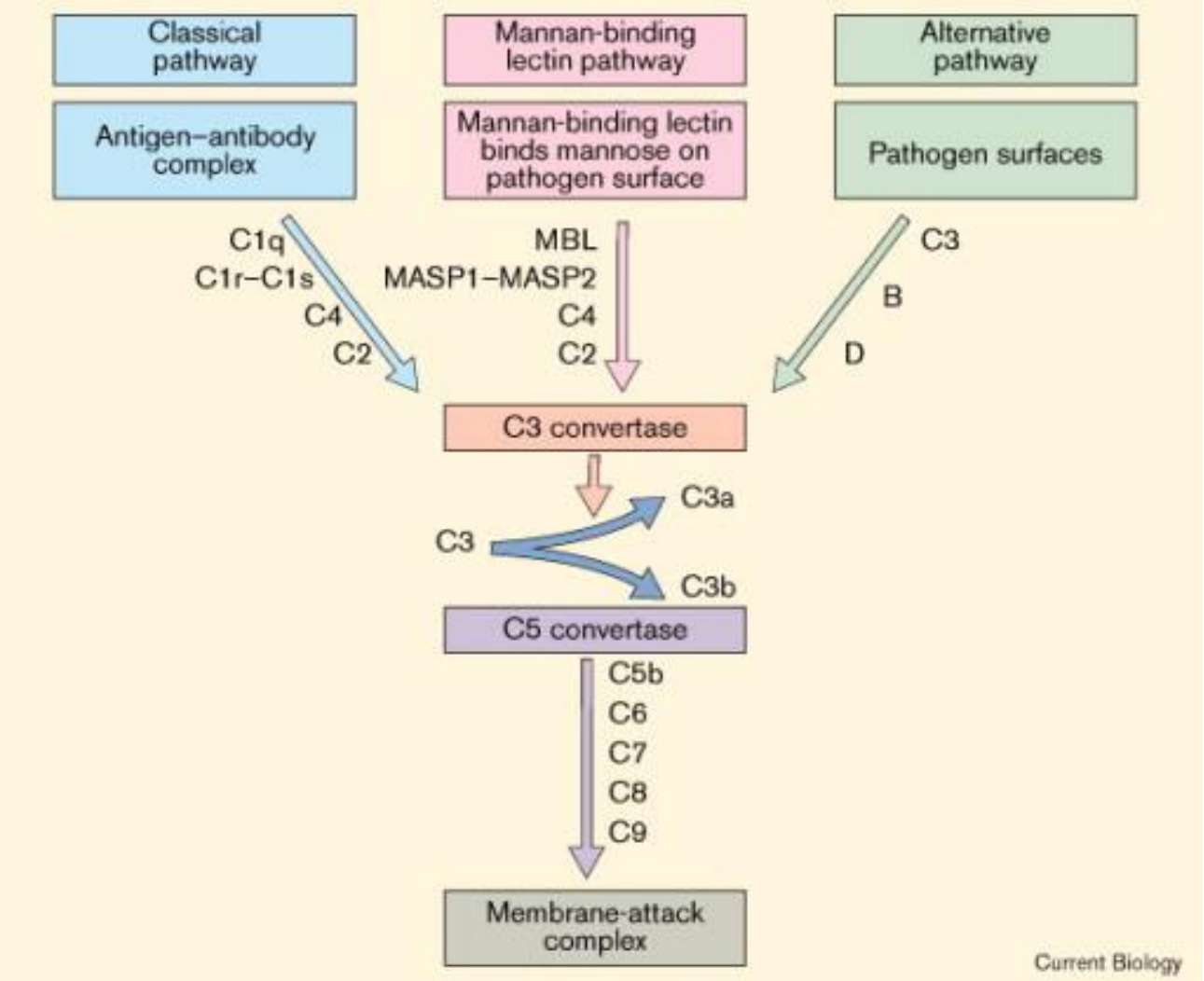
ARTICLE · Volume 43, Issue 1, 113611, January 23, 2024 · *Open Access* · Download Full Issue

Genetic determinants of complement activation in the general population

Damia Noce [1,2,9] · Luisa Foco [1,9] · Dorothea Orth-Höller [2,3,9] · Eva König [1] · Giulia Barbieri [1,4] · Maik Pietzner [5,6] · Dariush Ghasemi-Semeskandeh [1,8] · Stefan Coassin [7] · Christian Fuchsberger [1] · Martin Gögele [1] · Fabiola Del Greco M. [1] · Alessandro De Grandi [1] · Monika Summerer [7] · Eleanor Wheeler [6] · Claudia Langenberg [5] · Cornelia Lass-Flörl [2] · Peter Paul Pramstaller [1] · Florian Kronenberg [7,10] · Reinhard Würzner [2,10] · Cristian Pattaro [1,10,11] Show less

Complement

- essential, evolutionary ancient part of the innate immune system
- transversal roles across all human tissues
- altered activation associated with a broad spectrum of human diseases of both syndromic and complex nature
- functions: immune complex clearing; chemotaxis for recruiting inflammatory cells; opsonization and phagocytosis of foreign particles; cell lysis

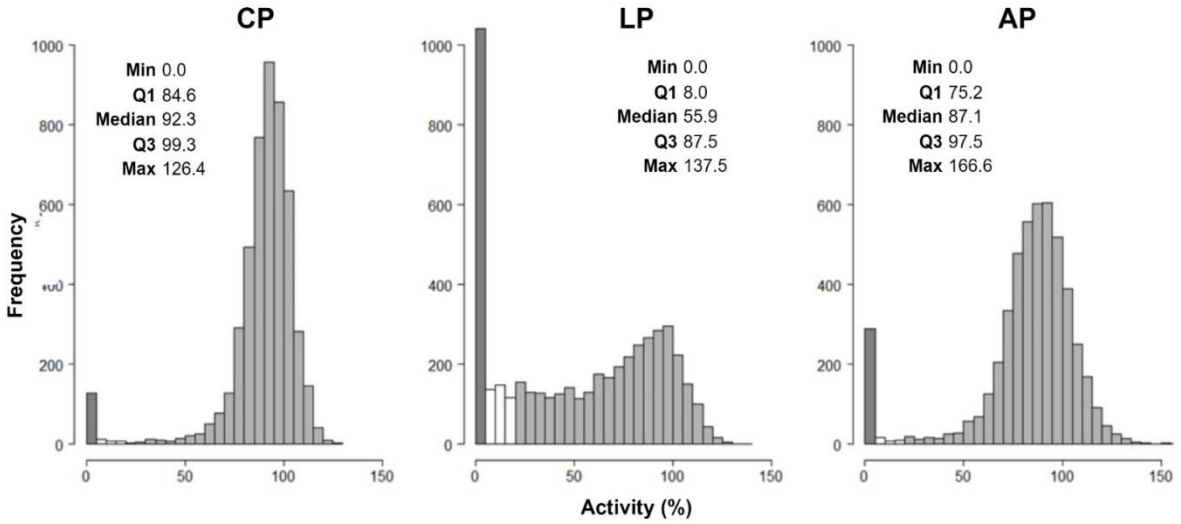


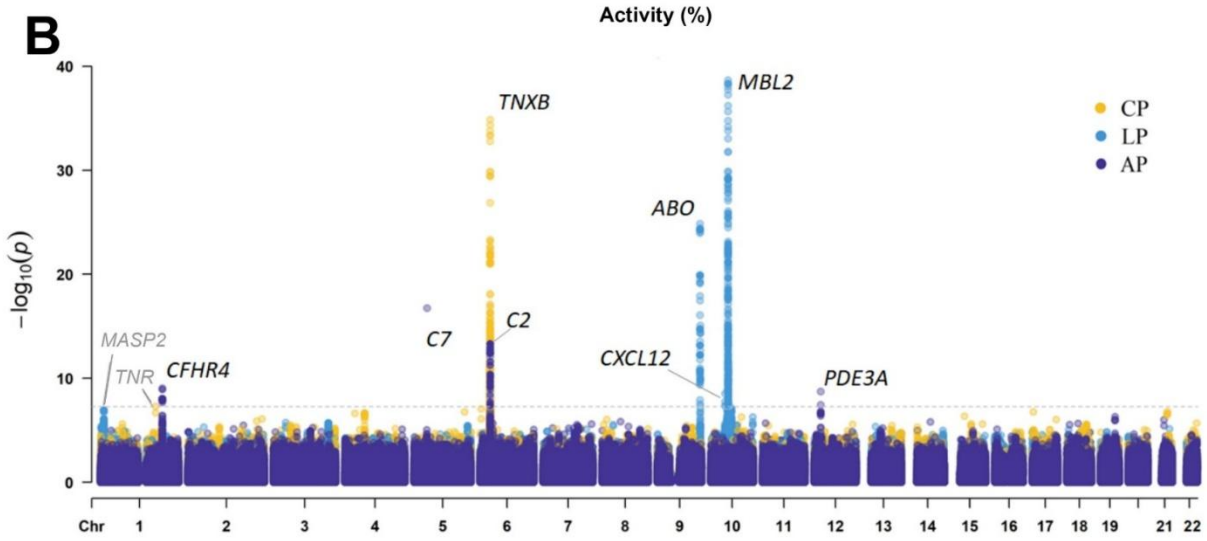
Taylor et al, **The complement system,** Current Biology 1998; 8(8):R259 - R261

Cooperative Health Research in South Tyrol (CHRIS) study

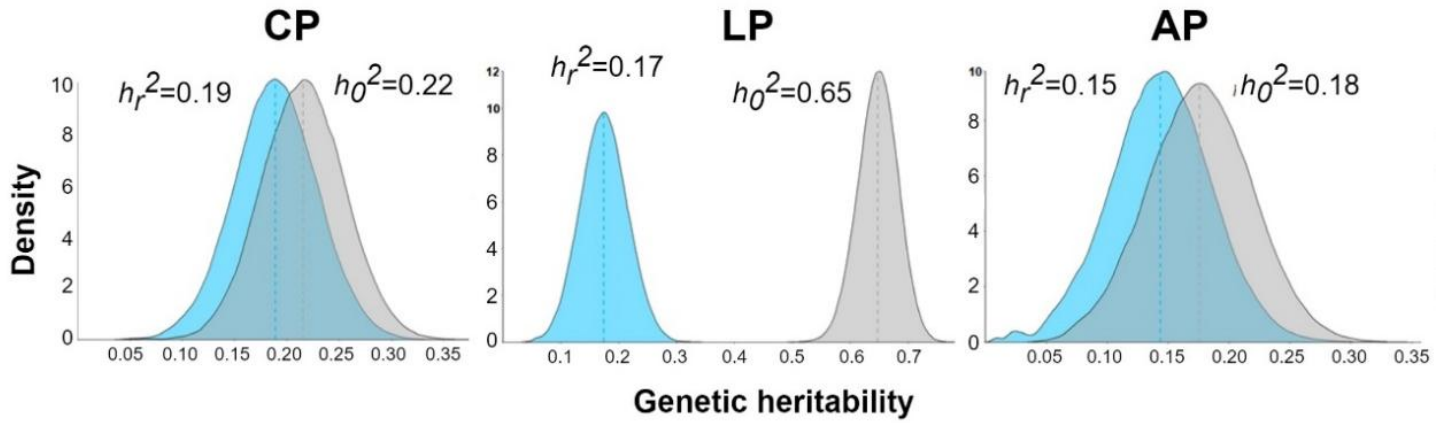
CHRIS
eurac research

Population-based study; N=4990 participants not enriched for any disease





13 independent, pathway-specific variants



explaining up to 74% of complement pathways' genetic heritability

GWAS of CP, LP, AP

Z. Kutalik, E. Porcu
(U Lausanne)

(1)

Functional annotation

Associated variants

Colocalization in *cis*

A. Teumer
(U Greifswald)

(2)

Phenome-wide scan:

- proteomics
- metabolomics
- complex traits & diseases

with gene expression

with protein level

M. Pietzner, C. Langenberg
(Queen Mary U London)

Informing instrumental variable selection

is it a complement protein?

NO

presence of colocalization?

YES

YES

A. Butterworth
(U Cambridge)

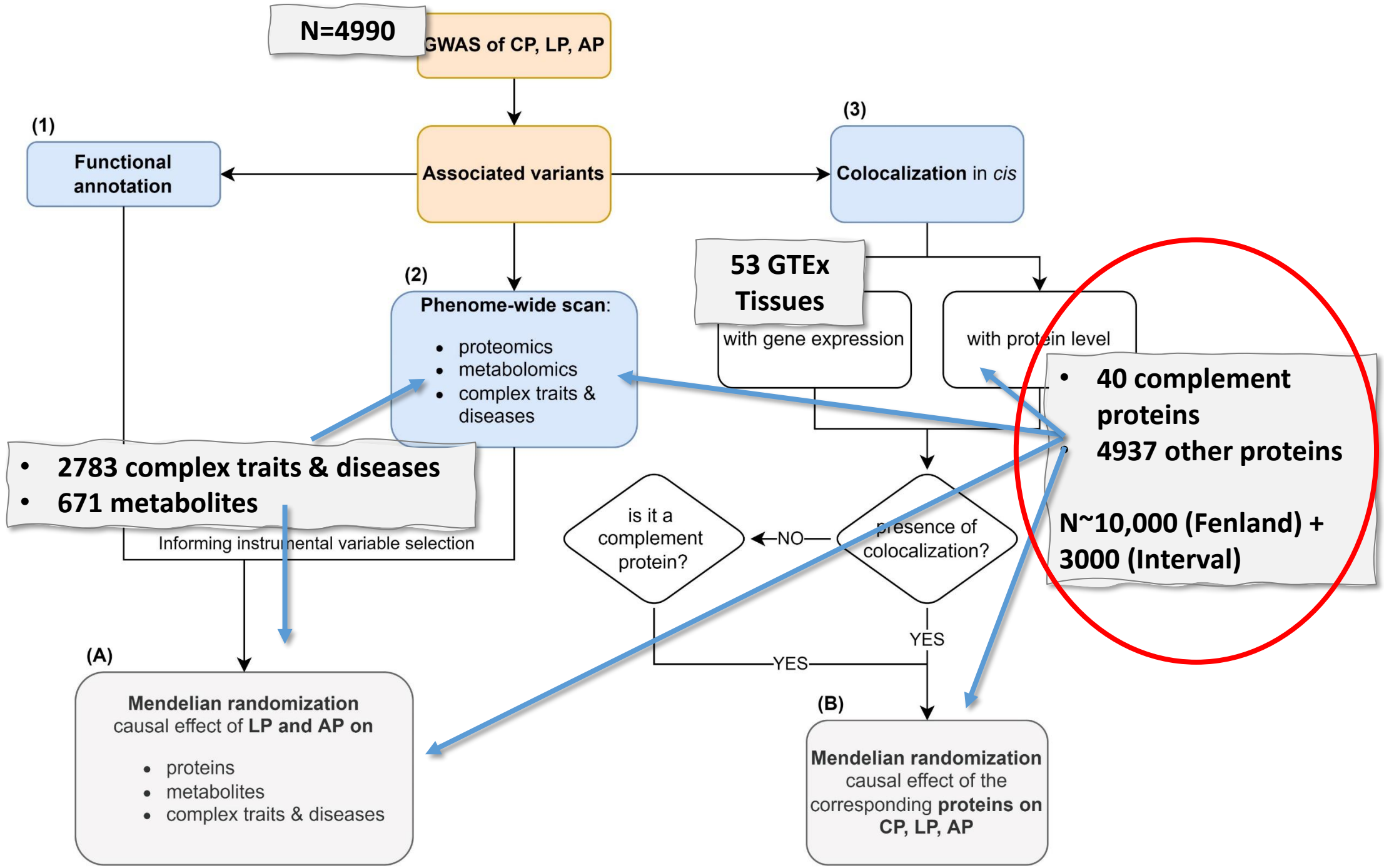
(A)

Mendelian randomization
causal effect of **LP and AP on**

- proteins
- metabolites
- complex traits & diseases

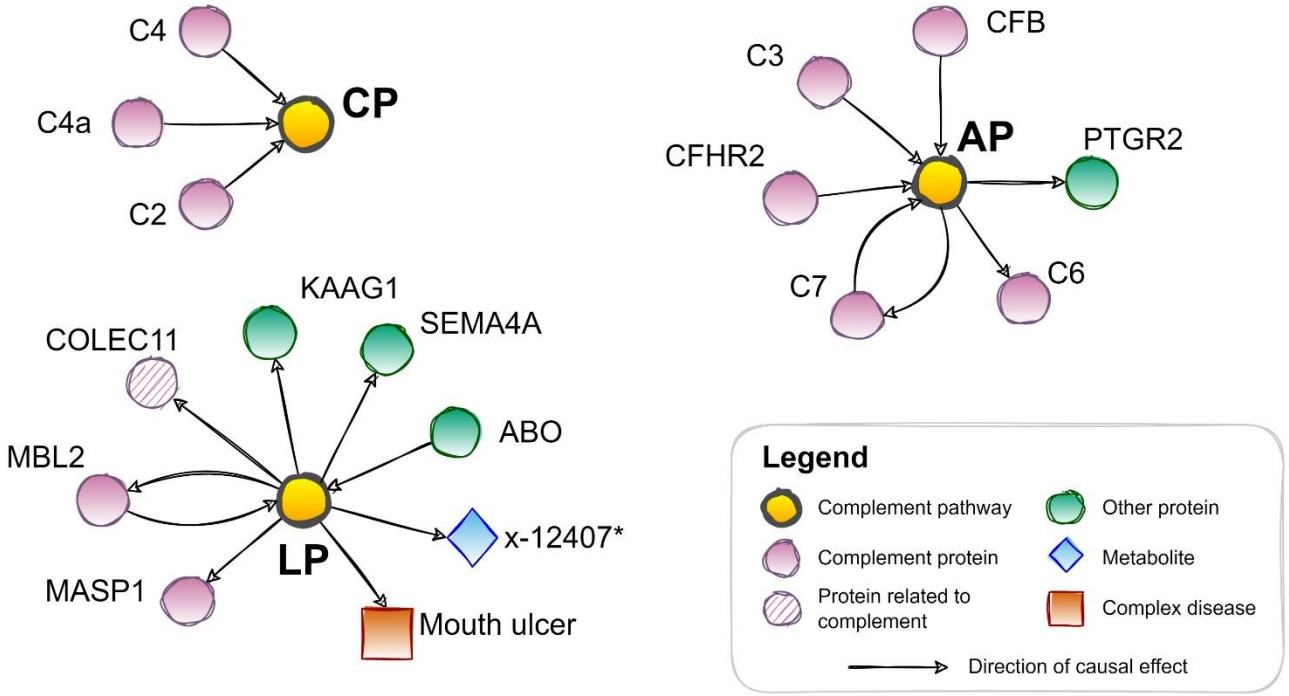
(B)

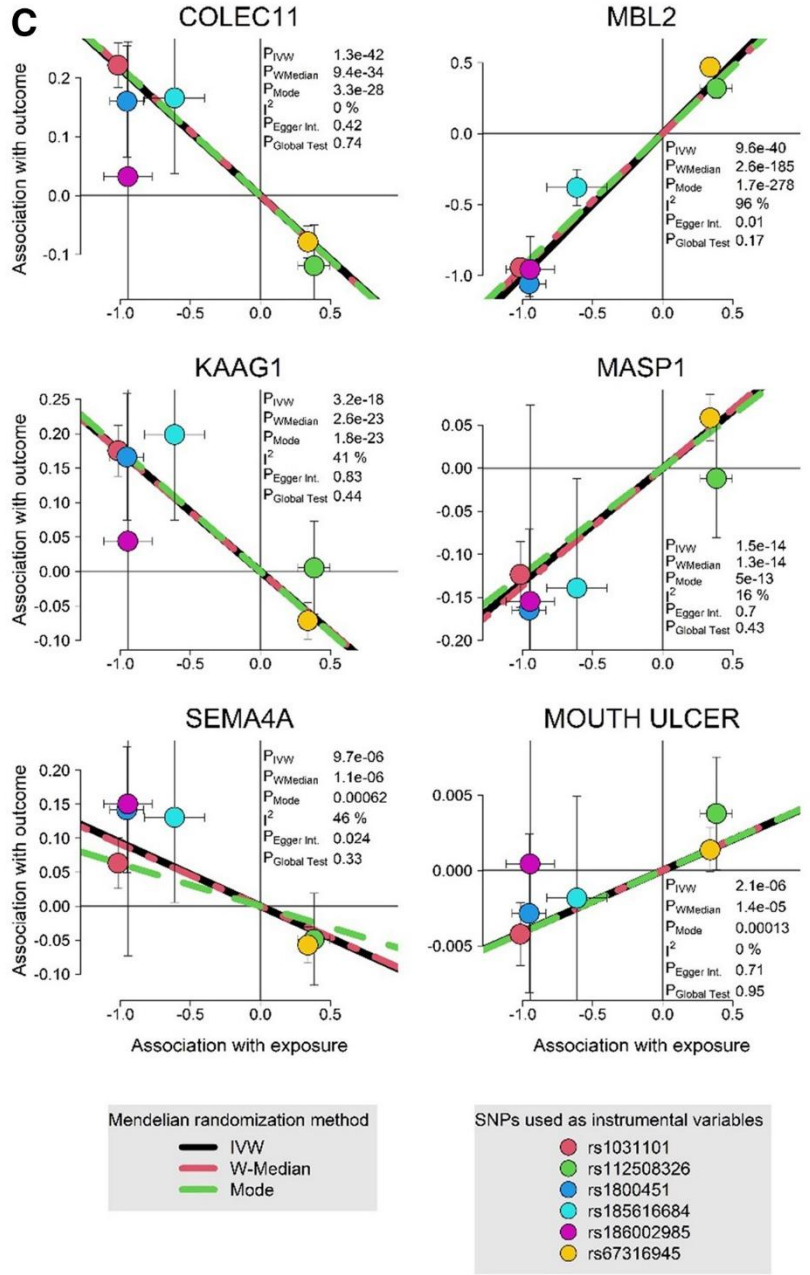
Mendelian randomization
causal effect of the corresponding **proteins on CP, LP, AP**

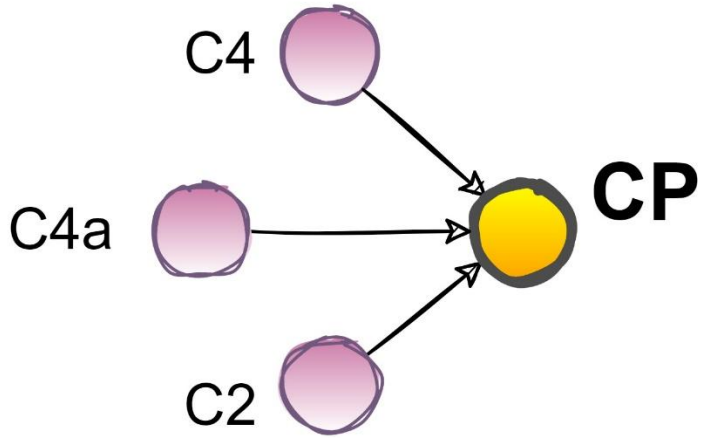

N=4990
GWAS of CP, LP, AP
(1)
Functional annotation
Associated variants
(3)
Colocalization in *cis*
(2)
Phenome-wide scan:
• proteomics
• metabolomics
• complex traits & diseases
53 GTEx Tissues
with gene expression
with protein level
• 40 complement proteins
• 4937 other proteins
N~10,000 (Fenland) + 3000 (Interval)
• 2783 complex traits & diseases
• 671 metabolites
Informing instrumental variable selection
is it a complement protein?
NO
presence of colocalization?
YES
YES
(A)
Mendelian randomization
causal effect of LP and AP on
• proteins
• metabolites
• complex traits & diseases
(B)
Mendelian randomization
causal effect of the corresponding proteins on CP, LP, AP

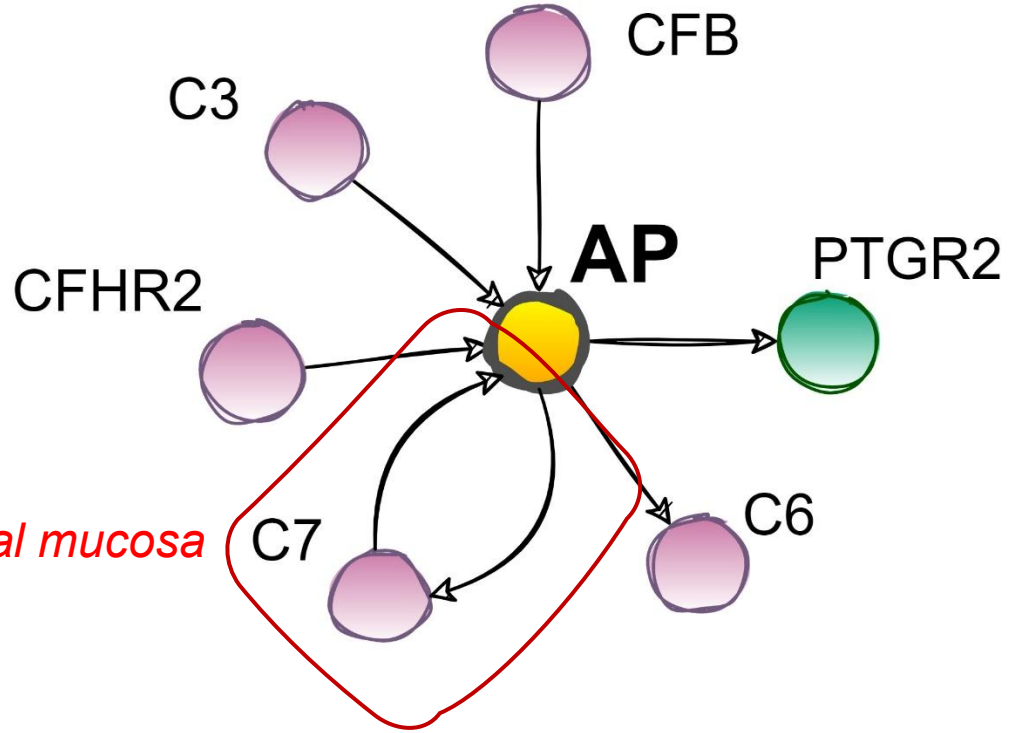
Mendelian randomizations of:

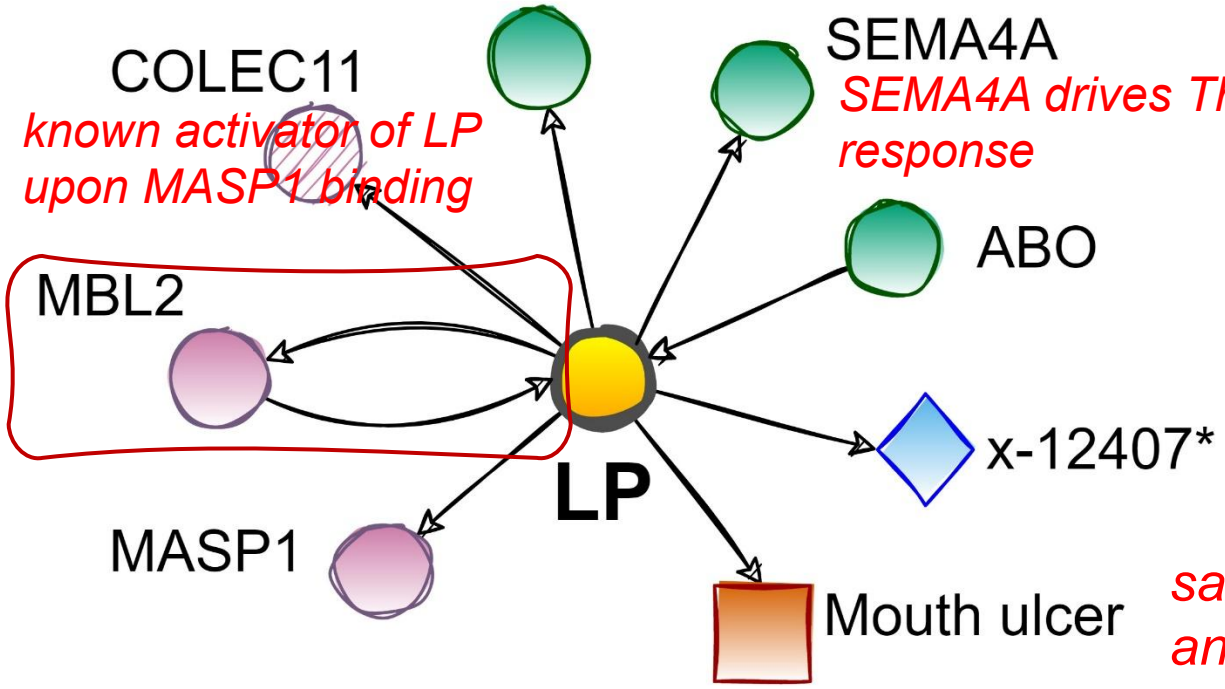
- 66 prioritized proteins (26 non-complement + 40 complement ones) →CP, LP, AP
- LP and AP → 4,979 proteins, 671 metabolites, and 2,783 complex traits





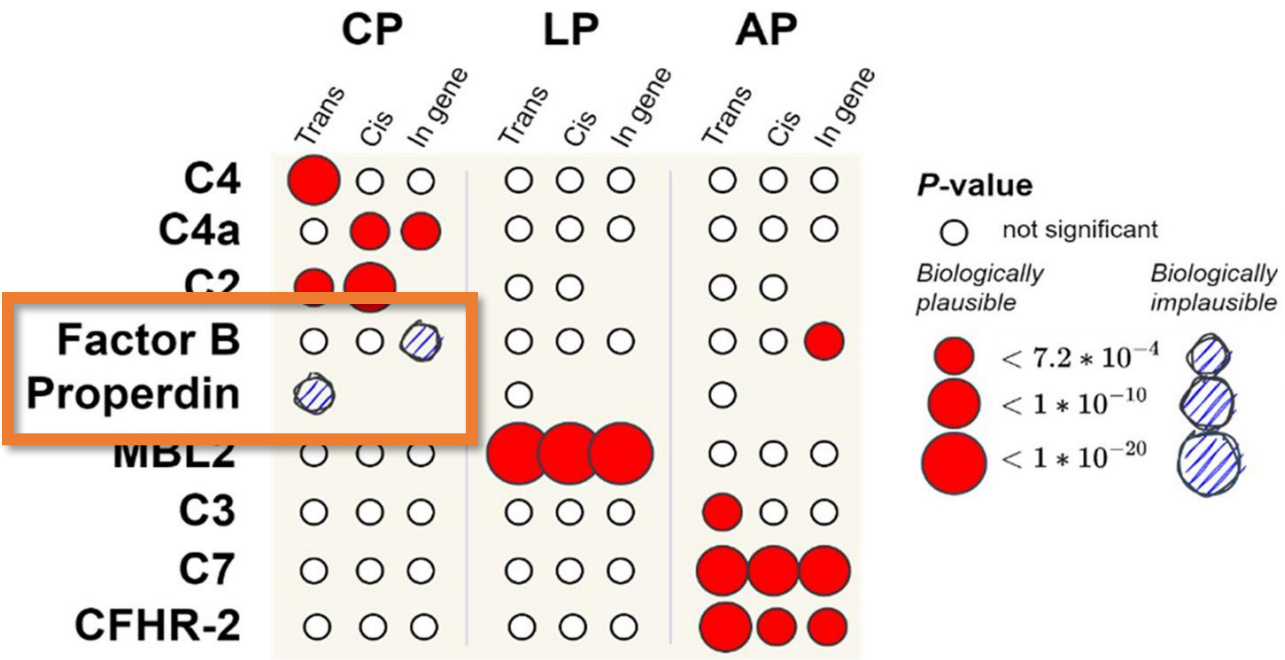


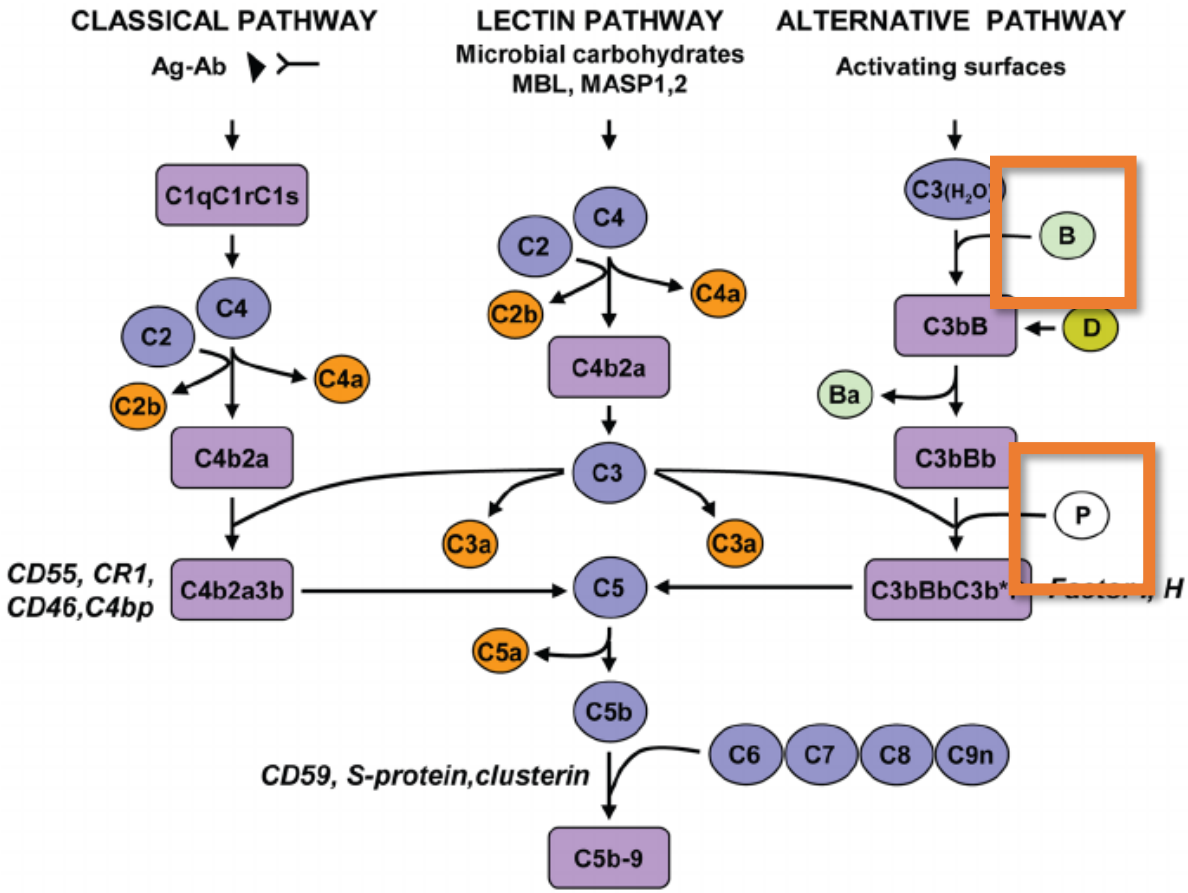




salivary agglutinin (SA) and salivary scavenger and agglutinin (SALSA) proteins can activate LP and inhibit LP activation

Causal relations between complement proteins and complement system activation vs known biology





Take home message

- ✓ Instrumental variable analysis reliably recapitulated most known biology of the complement system
- ✓ …and it identified new pathways
- ✓ …In a minority of cases, it also led to false claims that would have been impossible to confute in the absence of known biology

Even when all MR hypotheses are met, MR results should be treated with caution, awaiting experimental validation.

Give me hope!

A molecular target for desmoplakin

Luisa Foco

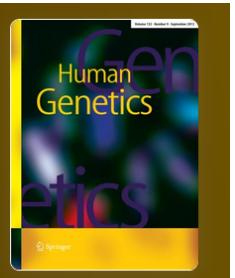
Marzia De Bortoli

Home › Human Genetics › Article

Genomic and molecular evidence that the LncRNA *DSP-AS1* modulates desmoplakin expression

Original Investigation | Open access | Published: 30 July 2025

(2025) Cite this article



Human Genetics

Luisa Foco ✉, Marzia De Bortoli, Fabiola Del Greco M, Laura S. Frommelt, Chiara Volani, Diana A. Riekschnitz, Benedetta M. Motta, Christian Fuchsberger, Thomas Delerue, Uwe Völker, Tianxiao Huan, Martin Gögele, Juliane Winkelmann, Marcus Dörr, Daniel Levy, Melanie Waldenberger, Alexander Teumer, Peter P. Pramstaller, Alessandra Rossini & Cristian Pattaro ✉

Arrhythmogenic Cardiomyopathy (ACM)

Rare inherited **cardiac disease**

→Progressive loss of cardiomyocytes

→fibrofatty replacement of the myocardium

→severe ventricular arrhythmias

→sudden cardiac death

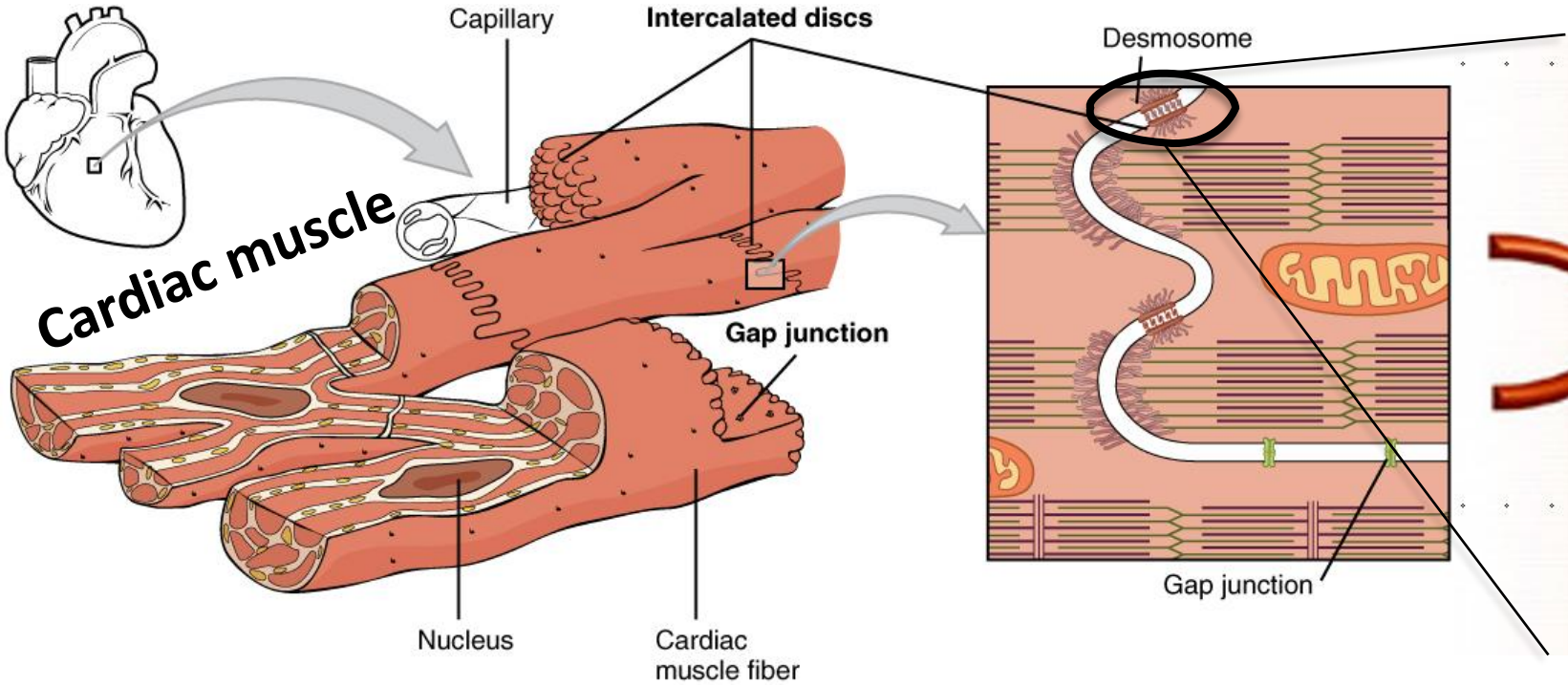
Marzia De Bortoli

A Rossini

Rare pathogenic variants most frequently located in **desmosomal genes**

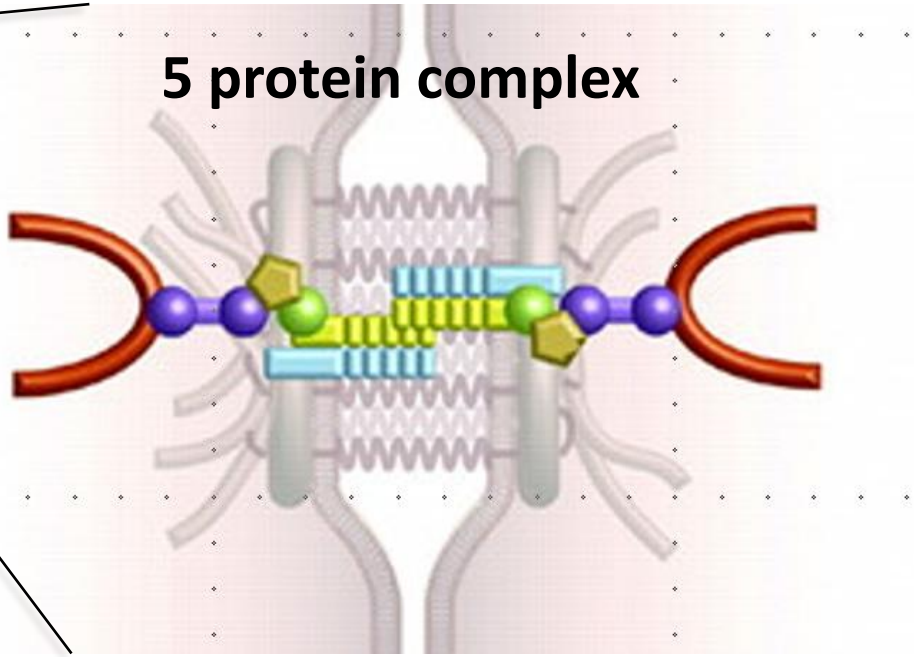
Cardiac desmosome

cell–cell adhesion structure between cardiomyocytes



Desmosome

5 protein complex



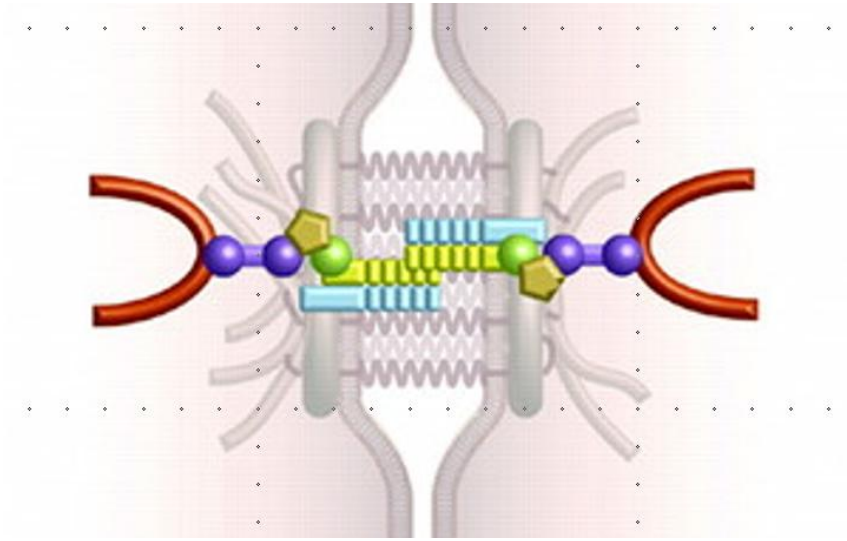
Delmar, McKenna, *Circ Res* 2010

https://commons.wikimedia.org/wiki/File:1020_Cardiac_Muscle.jpg

Cardiac desmosome

cell–cell adhesion structure between cardiomyocytes

5 protein complex



Delmar, McKenna, *Circ Res* 2010

Plakoglobin *JUP*

Desmoplakin *DSP*

Plakophilin-2 *PKP2*

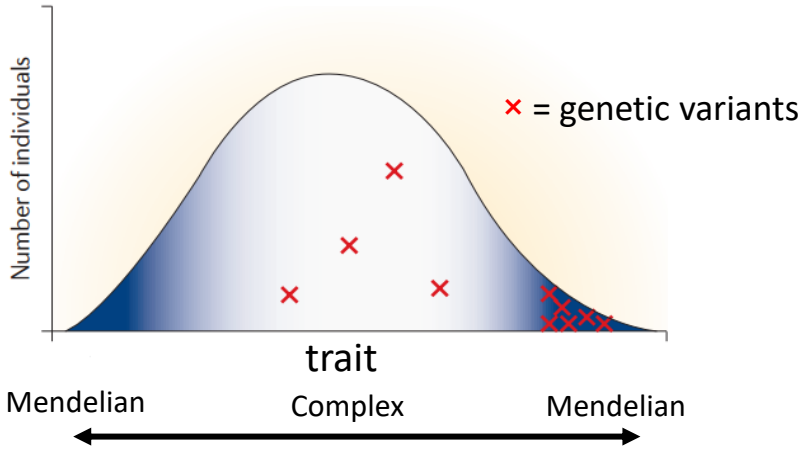
Desmocollin-2 *DSC2*

Desmoglein-2 *DSG2*

Scientific question #1

Luisa Foco
Senior Researcher, Eurac Research

Mendelian and complex traits share genetic architecture *(Blair 2013, Cell)*



Variants in desmosomal genes

rare → **ACM**

common → **Atrial conduction within normal ranges** ?

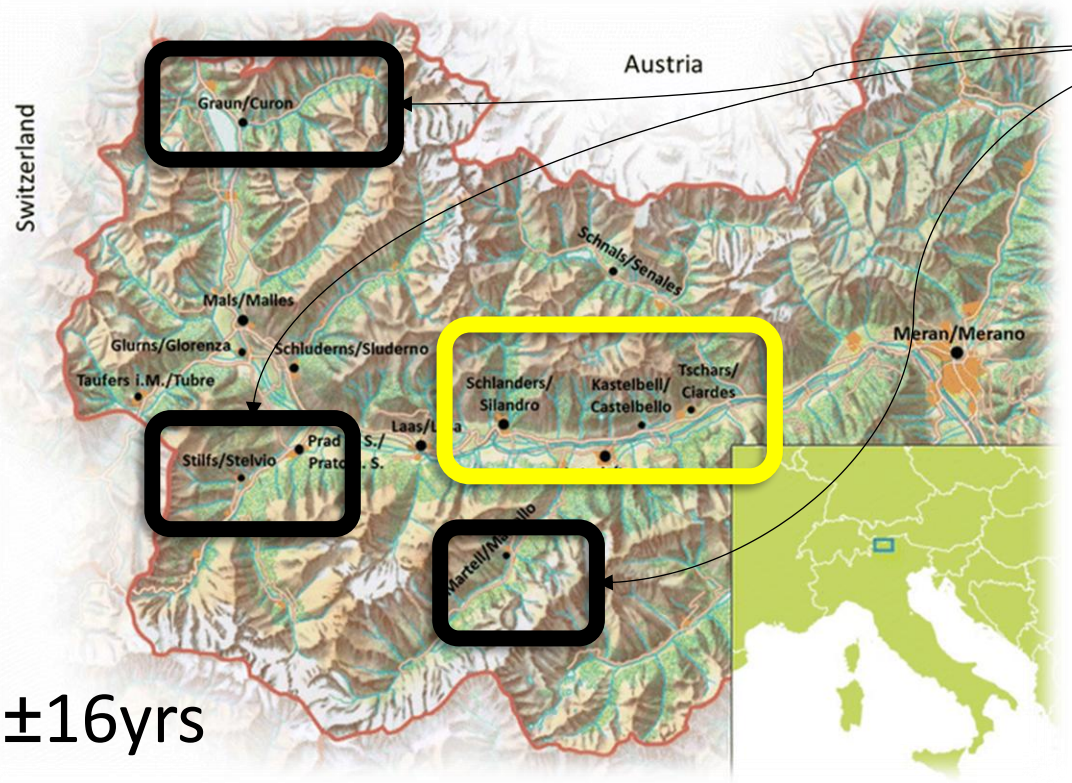
Q1 Are common variants at desmomal genes associated with ECG trait variability in the general population?

Cross-sectional studies

DISCOVERY

Cooperative Health Research In South Tyrol (CHRIS) study

CHRIS
eurac research



N=4338

56%F, 46±16yrs

Pattaro et al, *J Transl Med* 2015; Lundin et al, *Int J Epidemiol* 2025

REPLICATION

MICROisolates In South Tyrol (MICROS) study

N=636, 51%F, 44±17yrs

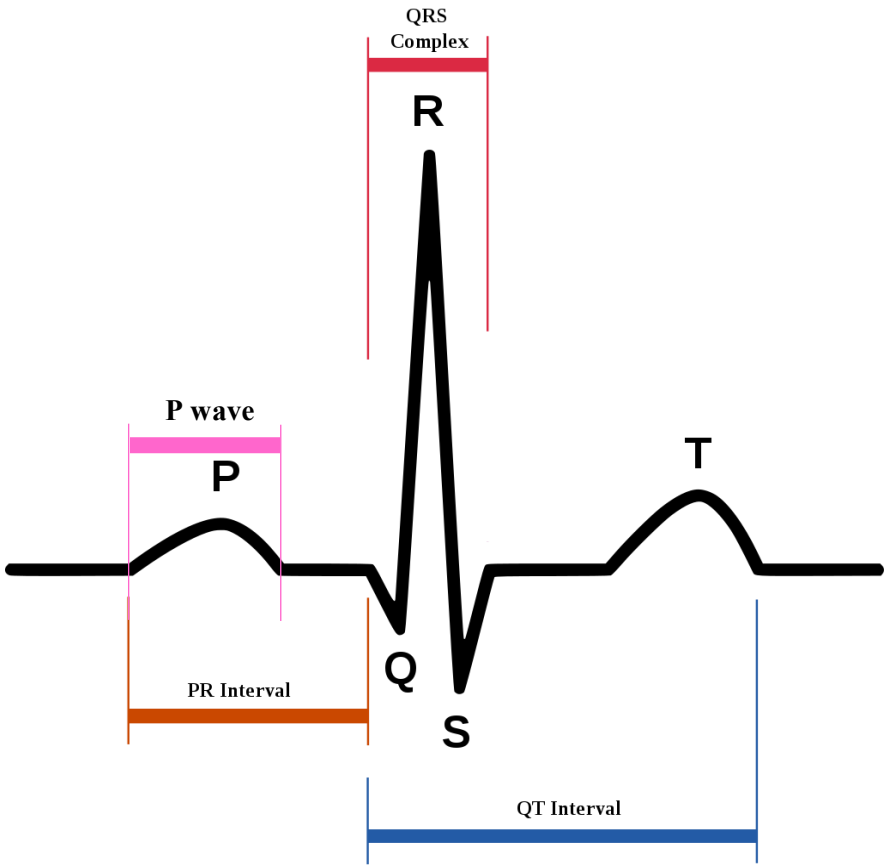
Pattaro et al, *BMC Med Genet* 2008

SHIP and SHIP-TREND

N=3779, 52%F, 48±16yrs

Völzke et al, *Int J Epidemiol* 2022

Outcomes



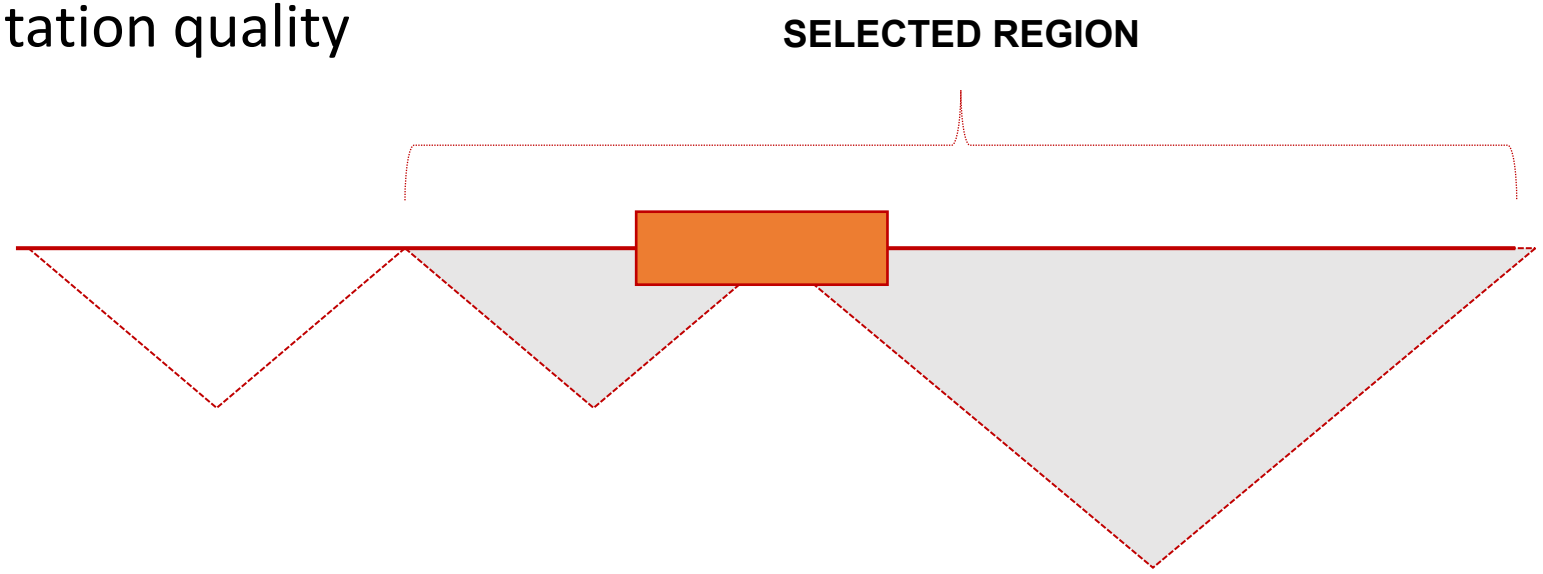
P-wave, PR, QRS, QT intervals

- from standard 10” 12-lead ECGs
- excl. history of atrial fibrillation, myocardial infarction, heart failure, Wolff-Parkinson-White syndrome, assuming class I and III antiarrhythmics and/or digoxin, pacemaker carriers, and pregnant women
- excl. outliers > | IQR ± 3 × IQR |
- ~normally distributed

Exposures

Imputed, single nucleotide polymorphism (SNP) dosage levels

- Illumina Human OmniExpress Exome array imputed on 1000 Genome Phase 1
- Selected at DSP, PKP2, JUP, DSC2, DSG2
- Total: **2742 SNPs**, **spanning 570 kb**
- with high imputation quality



Association testing

Linear mixed models

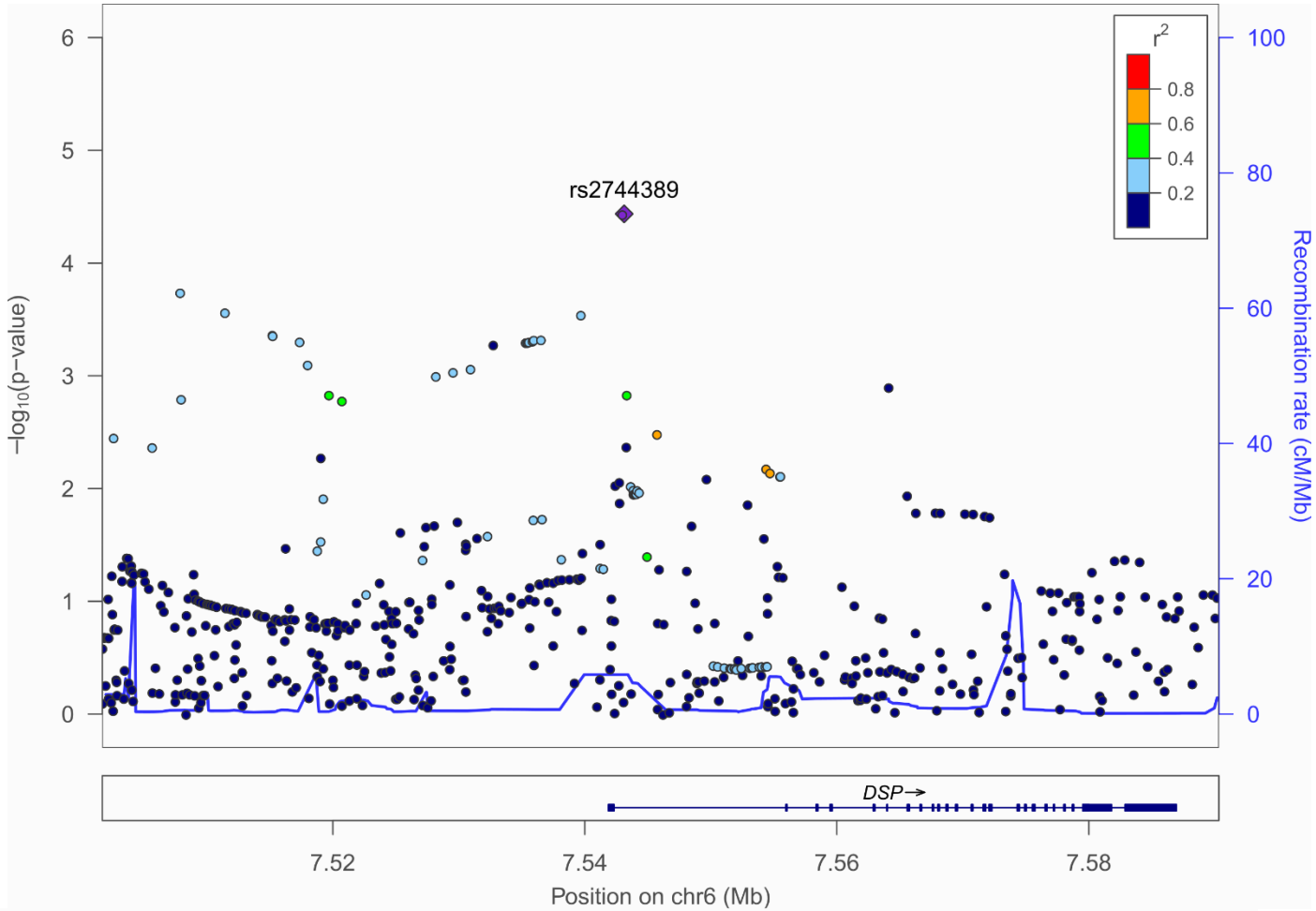
$$\boldsymbol{ECG\ trait} \sim \boldsymbol{SNP_i} + Fixed\ Effects + Random\ Intercepts + \varepsilon$$

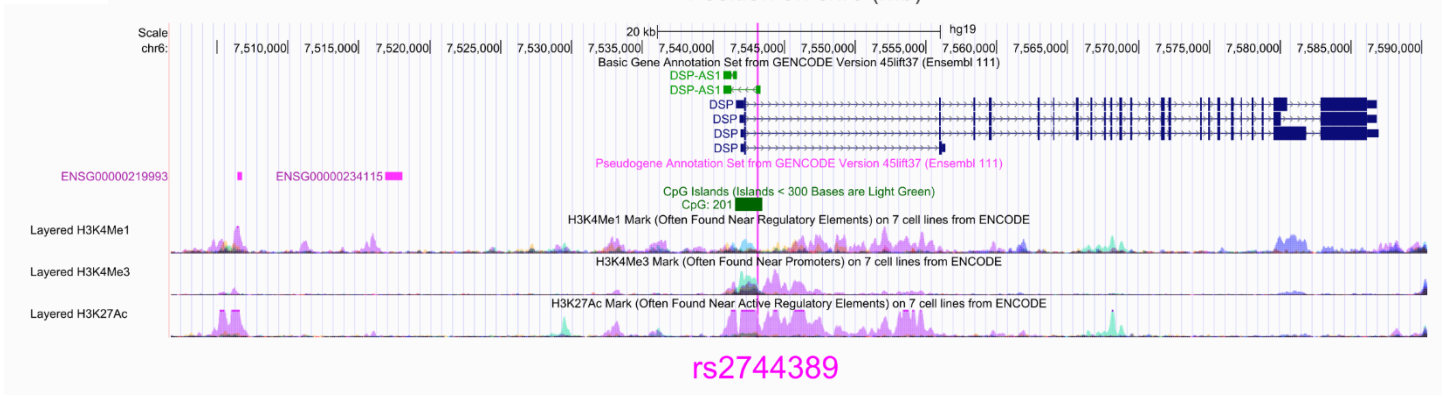
$$i = 1..2742; \qquad \varepsilon \sim \mathrm{N}(0, \sigma_K^2 K + \sigma^2 I)$$

- **Fixed effect covariates**: age, sex
- **Random intercept**: day of participation
- **Variance component**: kinship matrix
- **Significance level**: $\alpha_{discovery} = 2\times10^{-4}$; $\alpha_{replication} = 0.017$ (1-sided test / 3)
- **Software**: EMMAX and *coxme* in R

Results

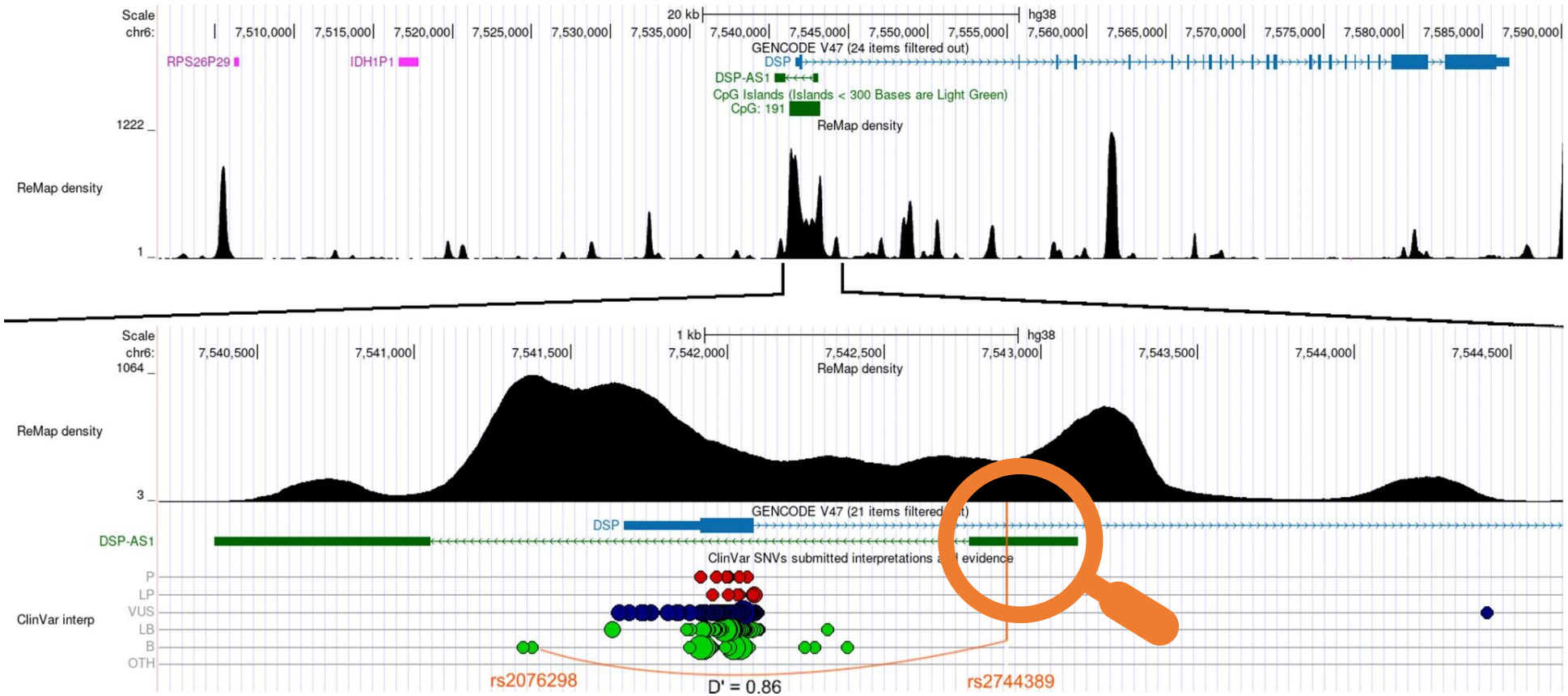
			CHRIS, n=4338			MICROS, n=636			SHIP, n=3779		
Trait	SNP, gene	Alleles	EAF	Beta(SE)	P	EAF	Beta(SE)	P‡	EAF	Beta(SE)	P‡
P-wave	rs115171396, *JUP*	C/**T**	0.02	4.87(1.08)	6.6×10^{-6}	0.02	-0.19(3.12)	0.525	0.01	1.59(1.45)	0.136
P-wave	rs72835665, *JUP*	G/**A**	0.51	-1.10(0.27)	4.5×10^{-5}	0.56	0.18(0.87)	0.585	0.53	-0.06(0.29)	0.416
QRS, ms	rs2744389, ***DSP***	A/**C**	0.18	-1.10(0.24)	3.5×10^{-6}	0.18	-1.47(0.64)	0.010	0.16	0.21(0.32)	0.748

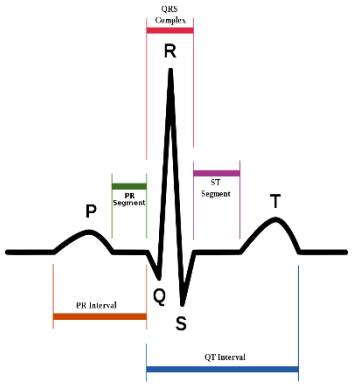

rs2744389
r²
0.8
0.6
0.4
0.2
−log10(p−value)
Recombination rate (cM/Mb)
DSP
7.52
7.54
7.56
7.58
Position on chr6 (Mb)


Scale
chr6:
20 kb
hg19
Basic Gene Annotation Set from GENCODE Version 45lift37 (Ensembl 111)
DSP-AS1
DSP-AS1
DSP
DSP
DSP
DSP
Pseudogene Annotation Set from GENCODE Version 45lift37 (Ensembl 111)
ENSG00000219993
ENSG00000234115
CpG Islands (Islands < 300 Bases are Light Green)
CpG: 201
H3K4Me1 Mark (Often Found Near Regulatory Elements) on 7 cell lines from ENCODE
Layered H3K4Me1
H3K4Me3 Mark (Often Found Near Promoters) on 7 cell lines from ENCODE
Layered H3K4Me3
H3K27Ac Mark (Often Found Near Active Regulatory Elements) on 7 cell lines from ENCODE
Layered H3K27Ac
rs2744389

rs2744389 on DSP associated with QRS interval

1) located **in** DSP and in DSP-AS1 (anti-sense long-non-coding RNA)
2) **in LD w functional** causal mutations

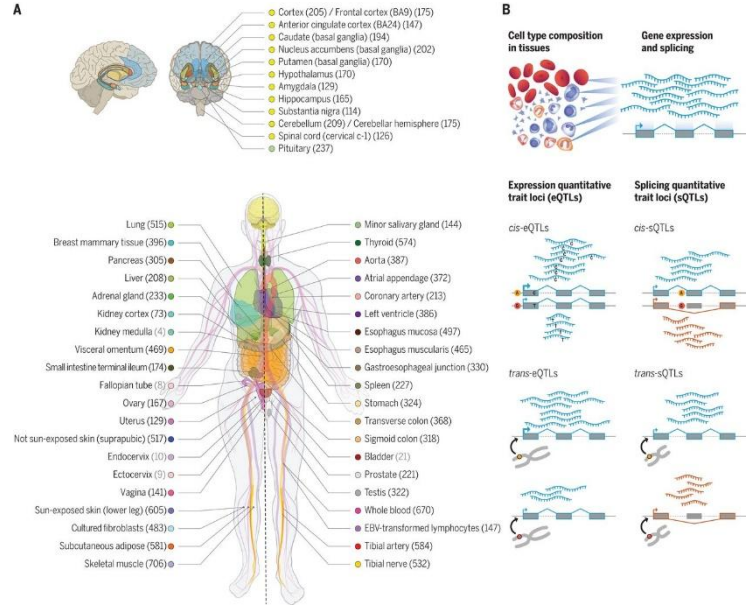




Results

GTEx Portal

https://www.gtexportal.org/home/snp/rs2744389



rs2744389 is associated with

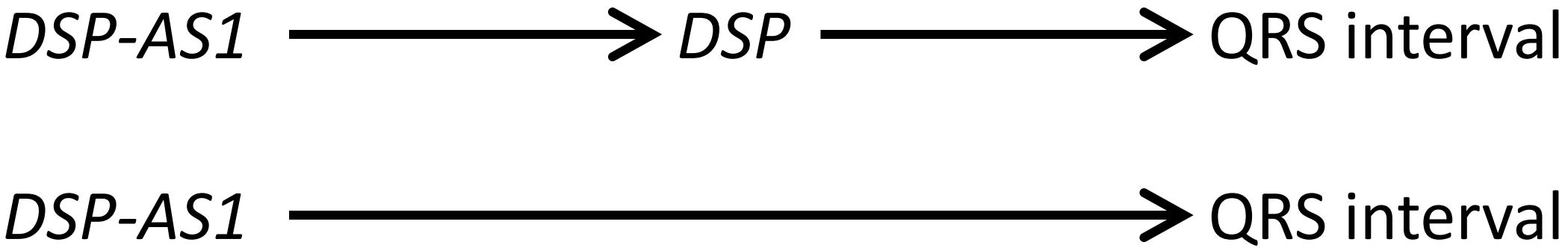
- **expression of *DSP-AS1* lncRNA**

(p = 3.1e-14 liver; 2.3e-7 heart atrial appendage)

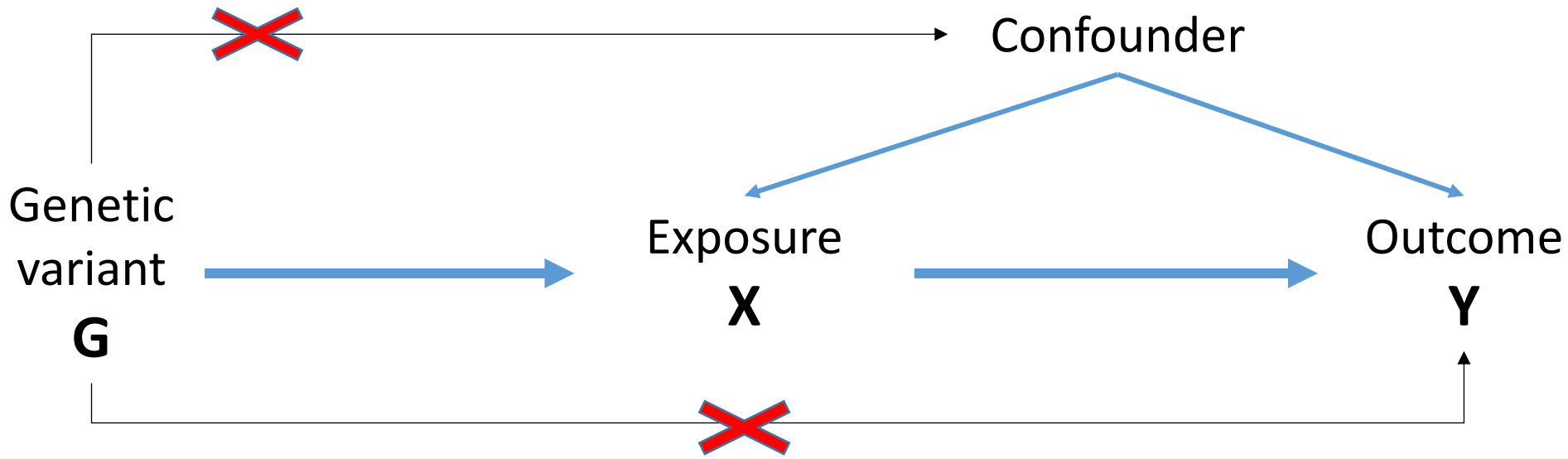
- but not with expression of *DSP*

Scientific question #2

Q2 is *DSP-AS1* expression causally associated with *DSP* expression and with the ECG QRS interval?



2-sample MR



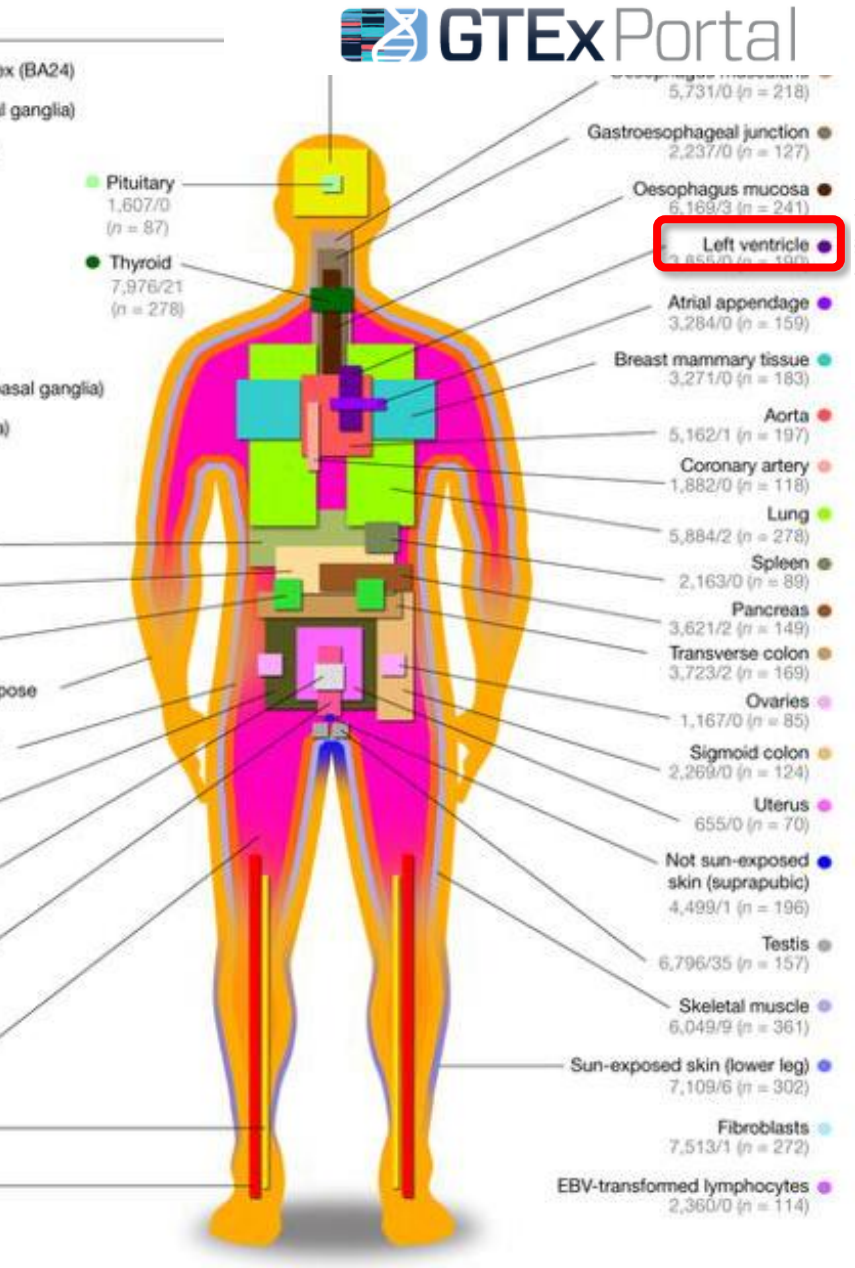
Wald-ratio estimator

$$b = \frac{effect\ of\ G\ on\ Y}{effect\ of\ G\ on\ X}$$

if **multiple variants**: meta-analysis; regression techniques; polygenic score

Core assumptions

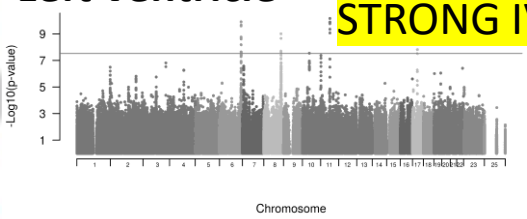
1) G is strongly associated with X
2) The G-X association is not confounded by hidden factors (eg: population stratification)
3) The is no other pathway from G to Y other than through X



DSP-AS1 RNA → DSP mRNA → QRS duration

Left ventricle

N=386
STRONG IV



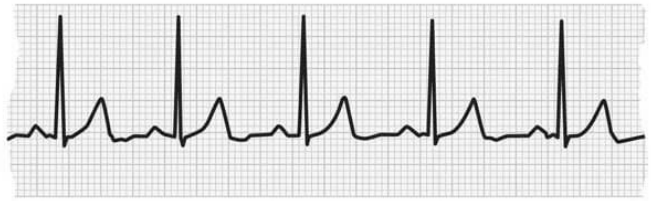
perfect tissue ☺

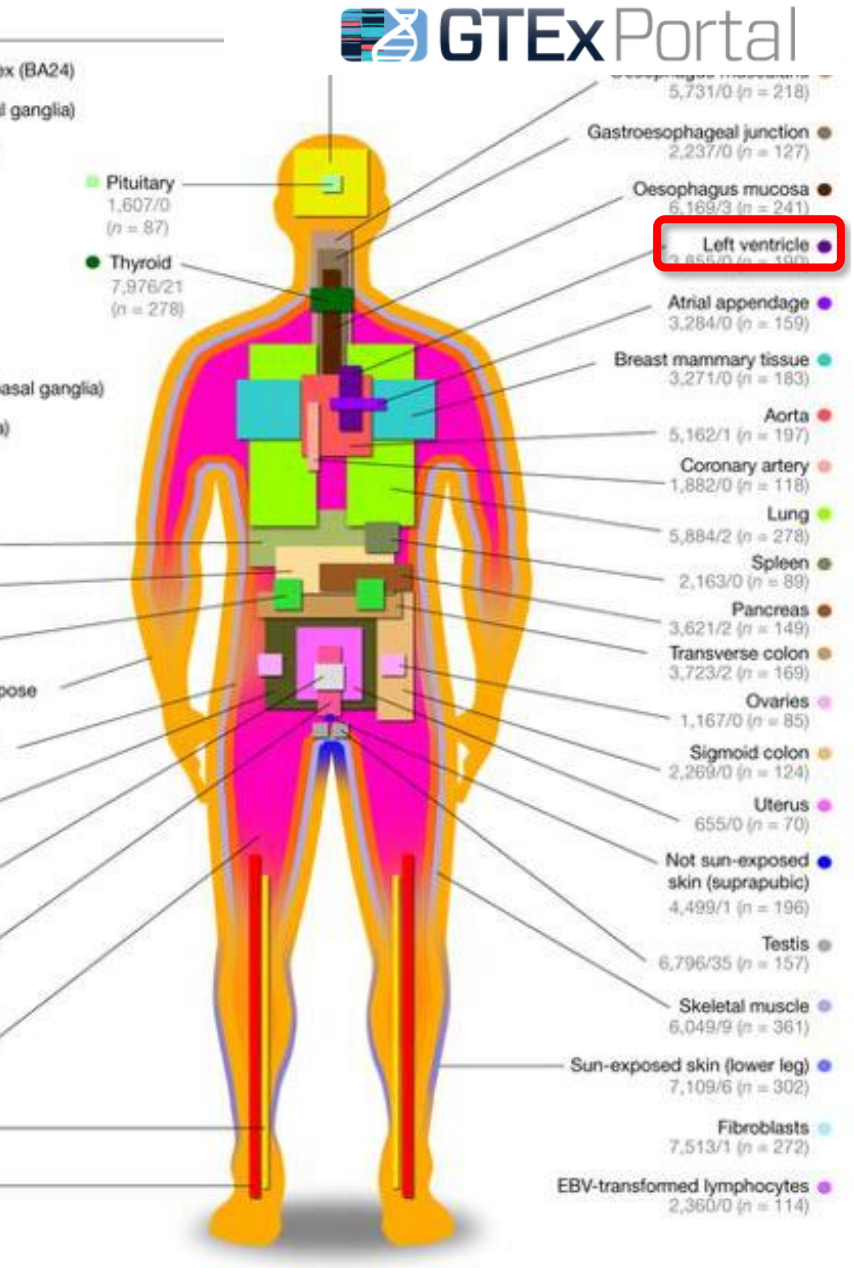
DSP-AS1 RNA
GTEx
left ventricle
N=386

D → cg02643433 methylation
Framingham Heart study
N=4170

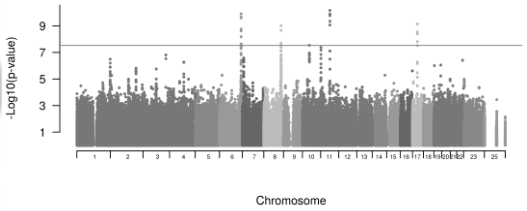
E* → DSP mRNA
GTEx, left ventricle
N=386
b=neg
$p=6.33\times10^{-5}$

F* → QRS duration
CHRIS study
N=4259
b=neg
p=0.015





Left ventricle



perfect tissue ☺

DSP → QRS interval

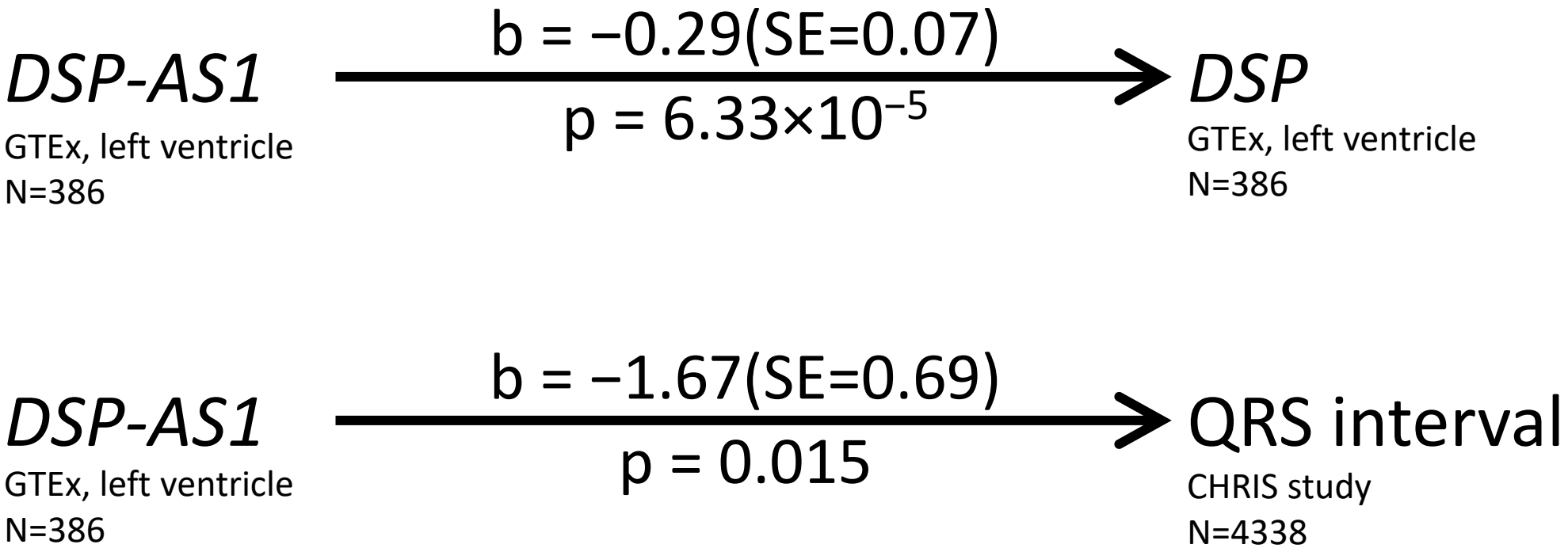
GTEx, left ventricle

NO VALID IVs

CHRIS study

N=4338

Results

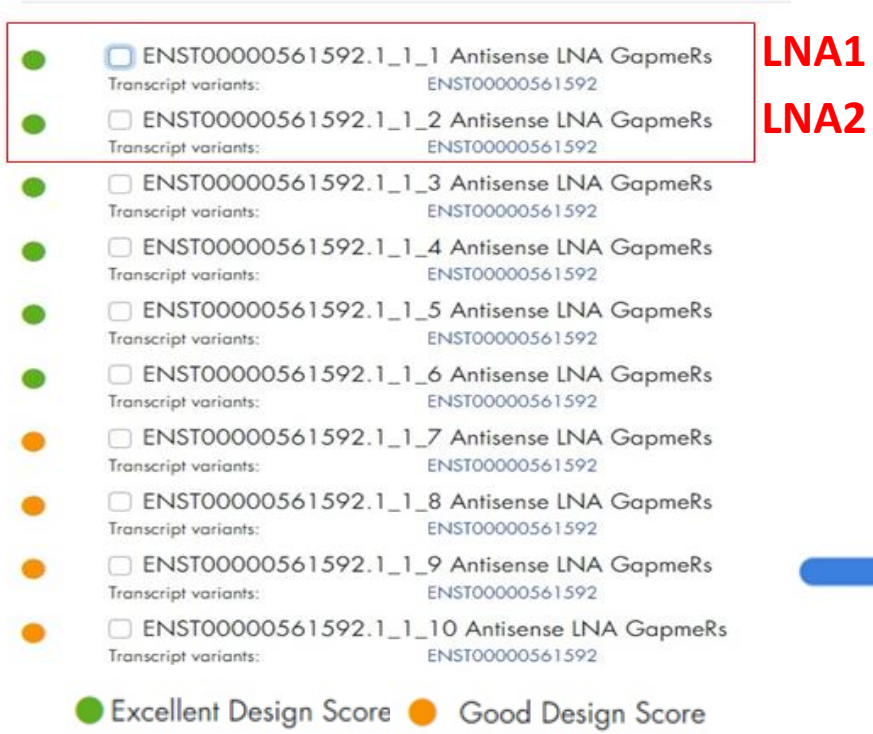

DSP-AS1
GTEx, left ventricle
N=386
b = −0.29(SE=0.07)
p = 6.33×10−5
DSP
GTEx, left ventricle
N=386
DSP-AS1
GTEx, left ventricle
N=386
b = −1.67(SE=0.69)
p = 0.015
QRS interval
CHRIS study
N=4338

Scientific question #3

Q3 Can we confirm *in vitro* the causal, negative effect of DSP-AS1 on DSP?

In vitro confirmation

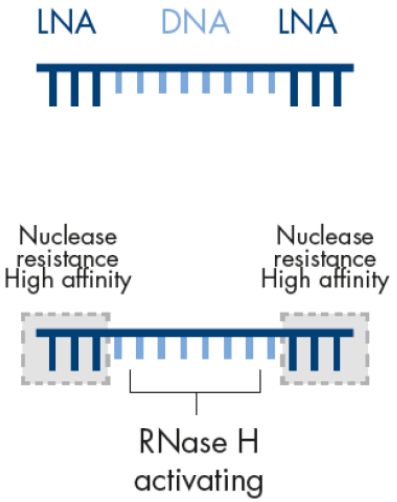
Antisense LNA GapmeRs design to downregulate DSP-AS1 lncRNA

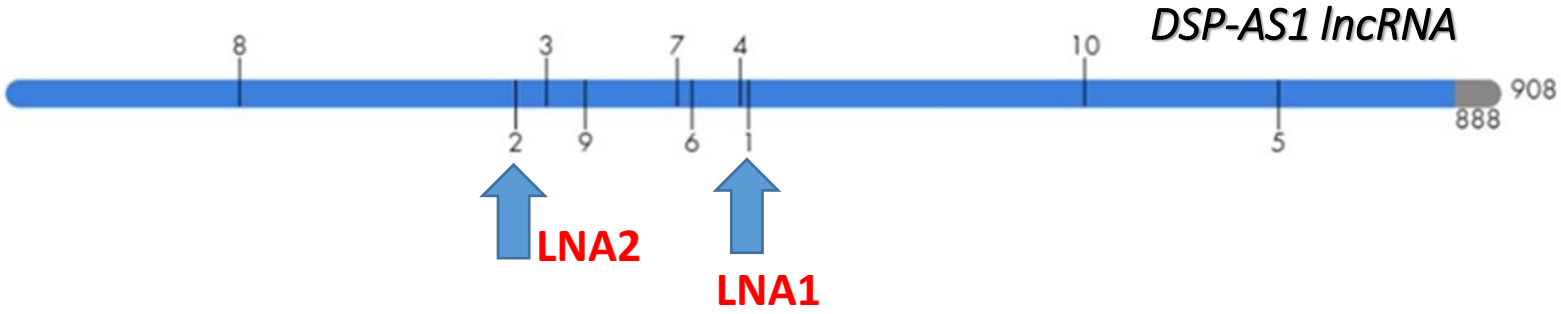


LNA1

LNA2

16 nucleotide sequence enriched with locked nucleic acids (LNAs) in the flanking regions, to increase affinity to the specific target
→ hybdrid DNA/RNA double-strand



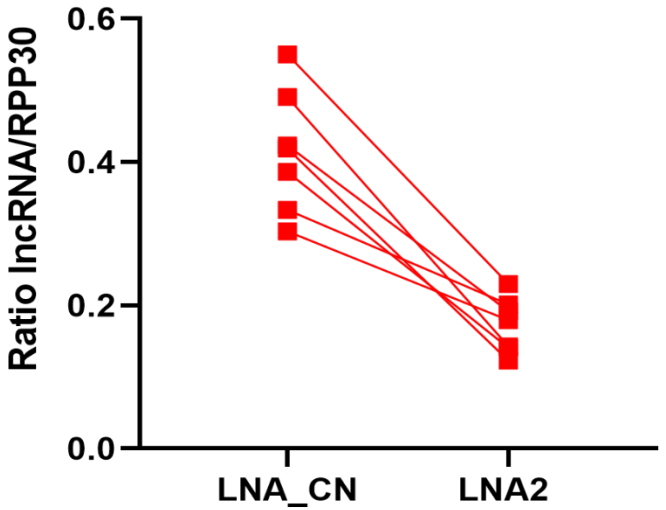


In vitro confirmation

in iPSC-derived cardiomyocytes (CMs)

lncRNA downregulation by LNA2 ↑ DSP mRNA expression

***DSP-AS1* expression**

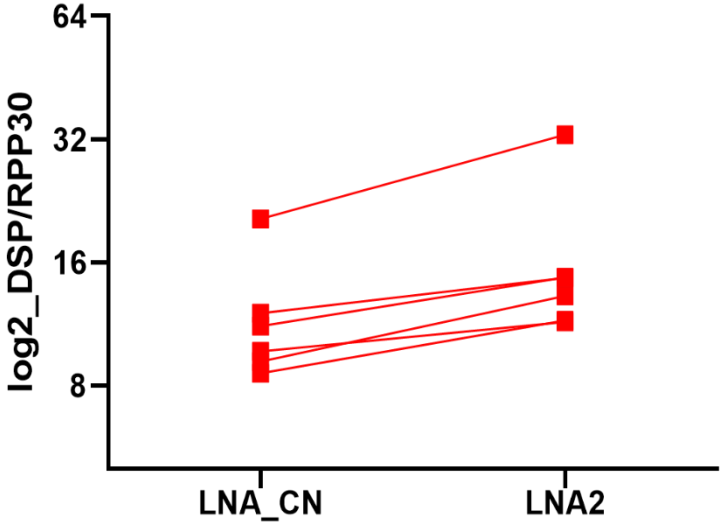


Wilcoxon matched-pairs signed rank test

** p=0.0156*

***DSP* expression**

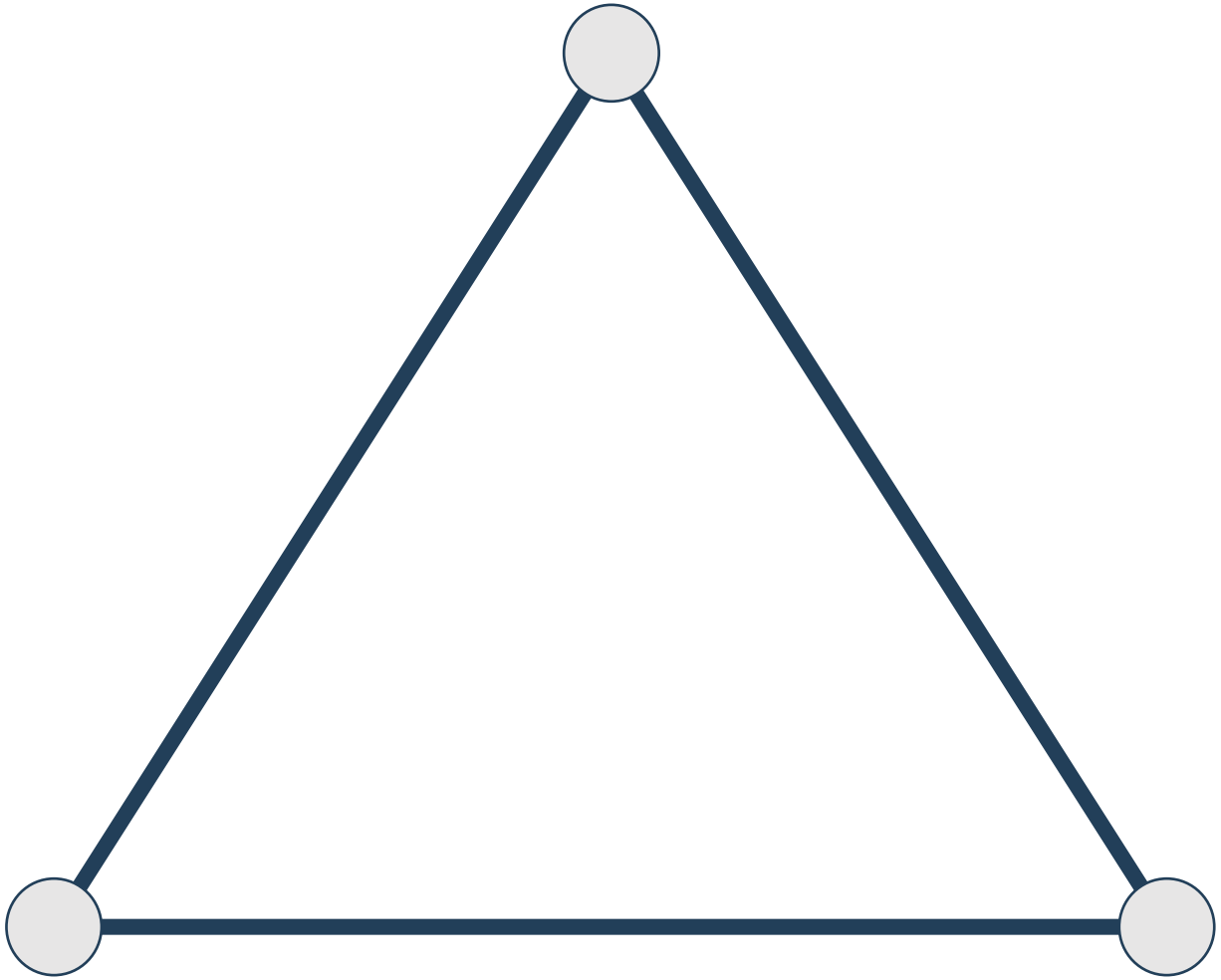
Treatment: 1000nM of Gapmers for 10days



Wilcoxon matched-pairs signed rank test

** p=0.0313*

Observational association (SNP is an eQTL for DSP-AS1 but not for DSP)



Statistical **causal** inference (MR)

In **vitro** demonstration

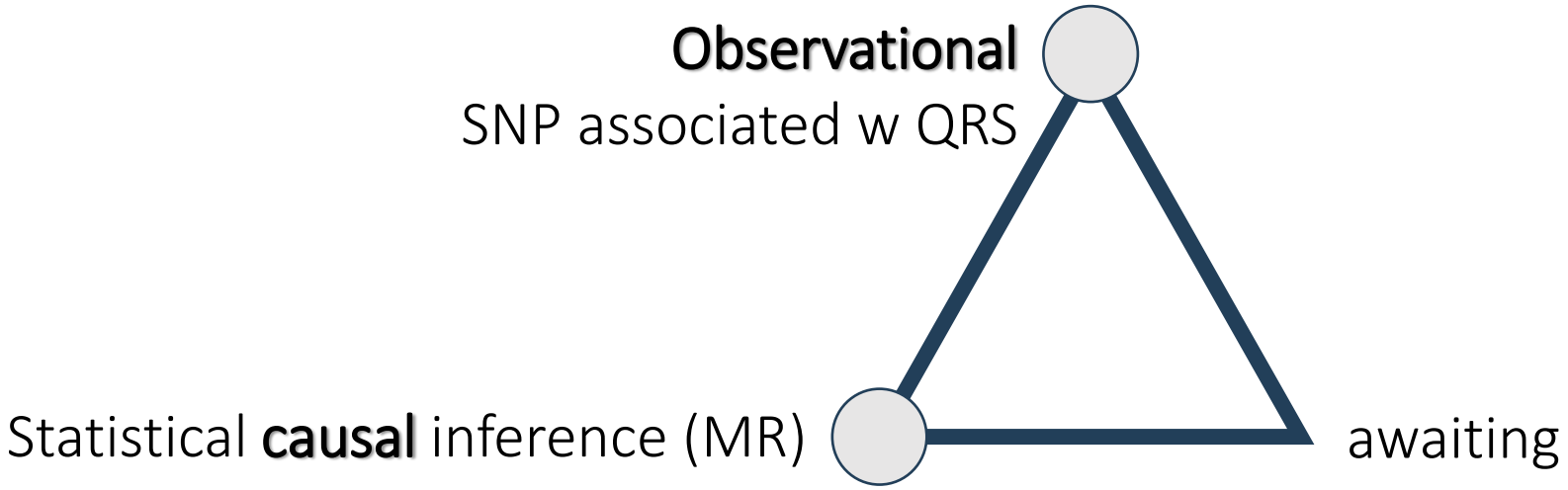
? Is DSP-AS1 causal to QRS ?

Mendelian randomization experiment using GTEx and GWAS data **confirmed**

DSP-AS1
GTEx, left ventricle
N=386

b = −1.67(SE=0.69)
p = 0.015

QRS interval
CHRIS study
N=4338

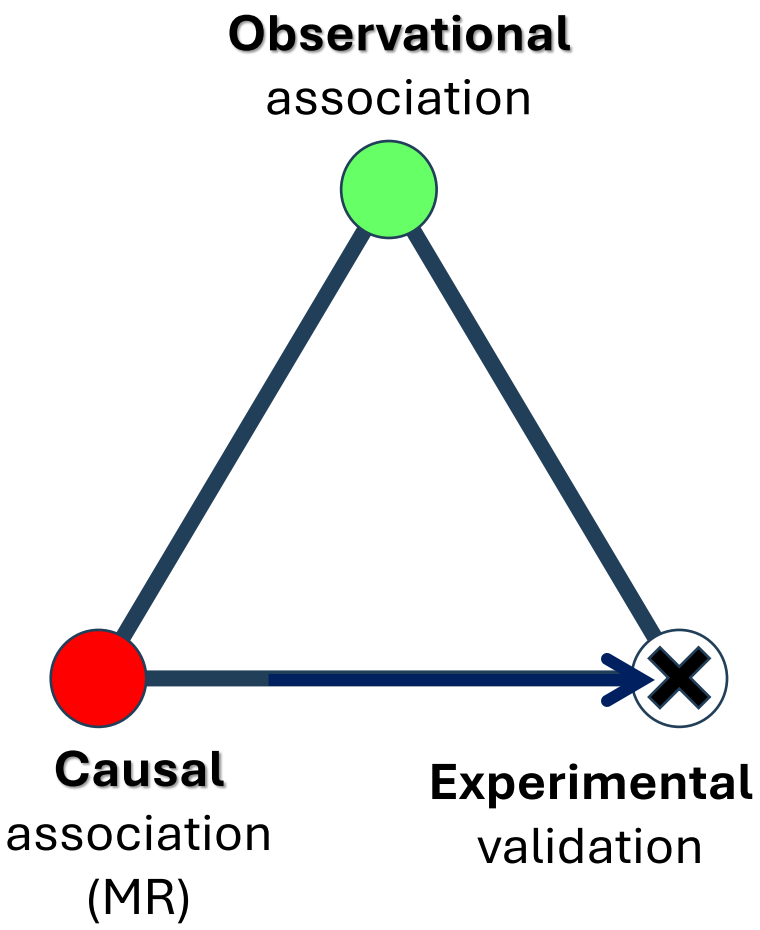


Conclusions

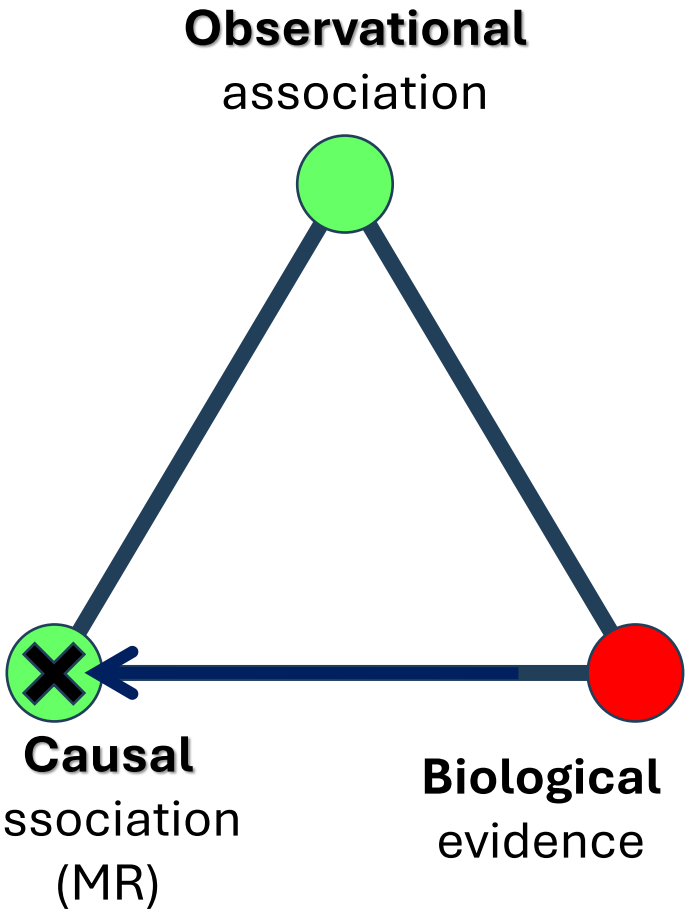
- First time a mechanism for DSP regulation was identified in cardiomyocytes
  - LNA2 Gapmer induces ~60% *DSP-AS1* downregulation in hiPSC-CMs
  - *DSP-AS1* downregulation increases *DSP* mRNA expression in hiPSC-CMs
- Potential to develop therapeutic strategies targeting *DSP* through *DSP-AS1*
- Further investigations warranted to assess results in ACM patients carrying *DSP* mutations

Conclusions

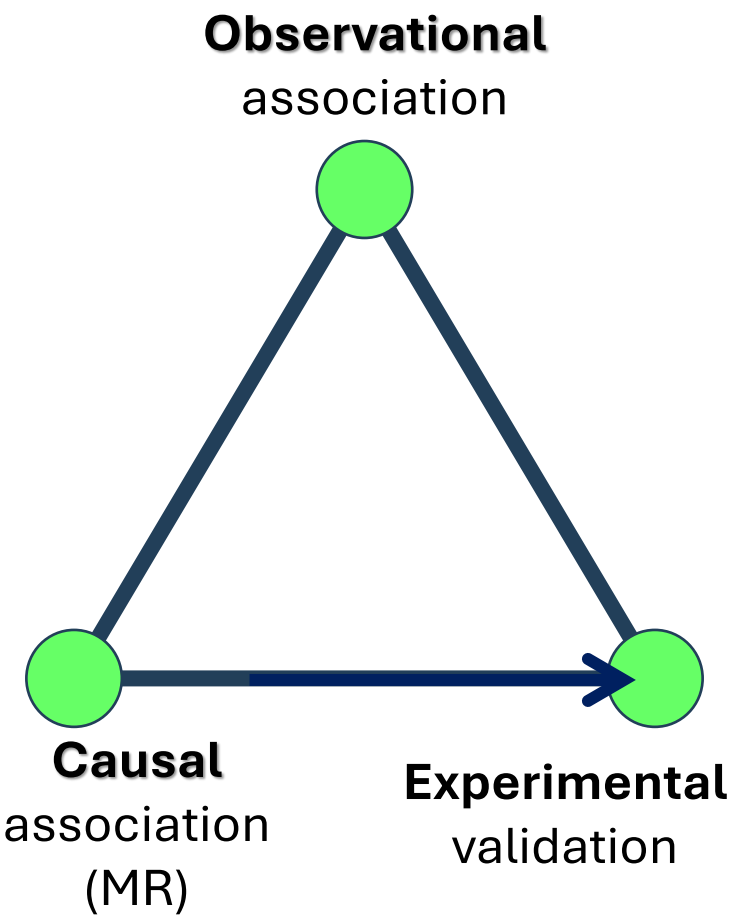
1
lipidomics



2
complement



3
desmoplakin



Take-homes

1. Population-based studies with appropriate omics data can provide valuable knowledge to identify therapeutic targets, also for rare conditions not usually detectable in cohort studies

2. Despite recent abuses, Mendelian randomization remains a solid method to prioritize functional experiments. Appropriate use of MR can help efficient resource allocation in scientific laboratories.

3. Causal mechanisms can never be demonstrated using only statistical methods and data: functional experiment or randomized trails are eventually necessary.

4. Triangulation of evidence should become a standard approach to causality

Acknowledgments

Group of Biostatistics & Epidemiology


L Foco


F Del Greco


R Melotti


M Gögele


R Lundin


S Lago


N Alipour


M Banerjee

Former members


D Noce

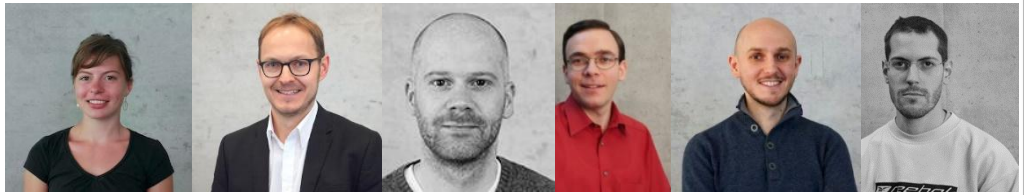

G Barbieri


D Ghasemi

INSTITUTE FOR
BIOMEDICINE


P Pramstaller


M De Bortoli A Rossini



Scientific collaborators


R Würzner F Kronenberg C Lass-Flörl D Orth-Höller
Medical University of Innsbruck


A. Butterworth
(U Cambridge)


M. Pietzner, C. Langenberg
(Queen Mary U London)


J Wilson
U Edinburgh


A Teumer
U Greifswald

Funding & support

- Department of Innovation, Research, University & Museums of the Autonomous Province of Bolzano
- European Regional Development Fund (FESR1157)
- EU's Horizon 2020 research & innovation programme - Marie Skłodowska-Curie grant agreement No 860977 (TrainCKDis)
- Joint Project Alto Adige-SNSF (Italy-Switzerland) to MDB
- Azienda Sanitaria Alto-Adige

Autonome Provinz Bozen
Provincia autonoma di Bolzano
Provinzia autonoma de Bulsan
SÜDTIROL · ALTO ADIGE

Thanks

cristian.pattaro@eurac.edu

@cpattaro.bsky.social