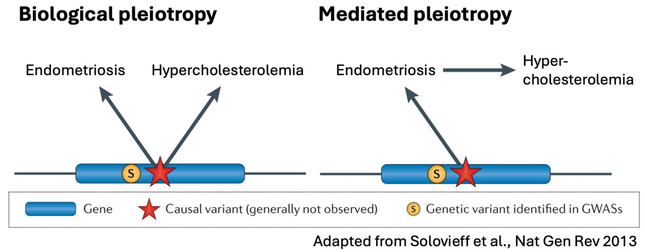
Elucidating the shared genetic architecture of hypercholesterolemia and endometriosis through a mediation framework

Ananya Rajagopalan[1], Lindsay Guare[1], Aude Ikuzwe[1], Lanna Caruth[1], Craig Teerlink[2], Danielle Candelieri[2],
Penn Medicine BioBank, Regeneron Genetics Center, Marijana Vujković[3], Julie Lynch[4], Shefali Setia-Verma[1]

[1]Department of Pathology and Laboratory Medicine, Perelman School of Medicine, Philadelphia, PA, United States [2]Departments Primary - Internal Medicine, Adjunct - Family & Preventive Medicine, University of Utah, UT, United States
[3]Division of Translational Medicine and Human Genetics, Perelman School of Medicine, Philadelphia, PA, United States [4]VA Informatics and Computing Infrastructure, VA Salt Lake City Healthcare System (VINCI), Salt Lake City, UT, USA

Background

- Hypercholesterolemia is a complex cardiovascular disorder shaped by genetic, environmental, and lifestyle factors.
- Individuals with endometriosis are at 1.25-fold increased risk of hypercholesterolemia, with previous work indicating significant genetic correlation between the two conditions.
- Current genome-wide association studies (GWAS) often overlook underlying mediating pathways, treating SNP effects on hypercholesterolemia as purely direct.
- **We hypothesize that accounting for endometriosis as a potential confounder among women would improve detection of relevant genetic associations with hypercholesterolemia**, by mitigating bias from hormone-related metabolic differences.

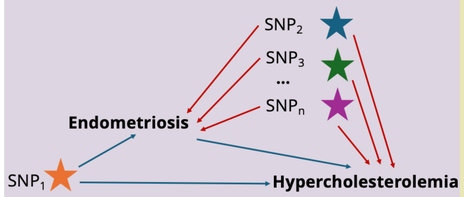


Adapted from Solovieff et al., Nat Gen Rev 2013

Graphical Abstract

To what extent does endometriosis mediate genetic effects on hypercholesterolemia?

Conceptual Framework

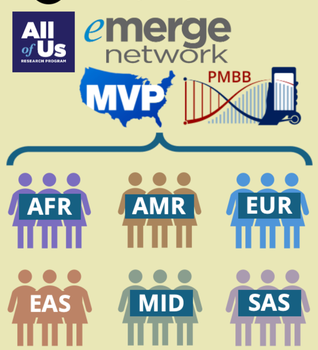


Adapted from Dong, Z., Zhao, H. & DeWan, A.T. *J Hum Genet* 2024

Methodology

Endometriosis-mediated GWAS for hypercholesterolemia

1 Study Population



2 Statistical Analysis

- Age, first 10 PCs, Endometriosis used as covariates
- Minimum case count of 50 for per-biobank, per-ancestry GWAS
- Proportion mediation calculation

Methods

Phenotyping: ICD-9 and ICD-10 codes for endometriosis and hypercholesterolemia.

Table 1: Distribution of Case/Controls Across Biobank and Ancestry

Ancestry	All of Us	emerge network	MVP	PMBB
AFR	n = 48,653 n_{case} = 12,254	n = 7,224 n_{case} = 3,254	n = 17,021 n_{case} = 9,905	n = 7,070 n_{case} = 3,063
AMR	n = 51,742 n_{case} = 9,973	-	n = 6,217 n_{case} = 3,140	n = 395 n_{case} = 125
EAS	n = 6,323 n_{case} = 1,055	-	-	-
EUR	n = 139,592 n_{case} = 41,689	n = 34,397 n_{case} = 18,790	n = 33,593 n_{case} = 21,664	n = 13,571 n_{case} = 5,524
MID	n = 817 n_{case} = 170	-	-	-
SAS	n = 2942 n_{case} = 445	-	-	n = 275 n_{case}= 81

Mediated GWAS: We leveraged microarray genotyping array data and short-read WGS data from the 4 biobanks to run the mediated GWAS. Age and the first 10 principal components were used as covariates. We enforced a minimum threshold of 50 cases per ancestry group. We added the mediating phenotype (endometriosis) as a covariate. GWAMA was used for meta-analysis. **369,832 individuals (131,132 cases)**

Results: Hypercholesterolemia Mediated GWAS

We identified **34 genome-wide significant loci**, **20 of which are unreported in a previous hypercholesterolemia GWAS**.

Among the most significant were loci within the *GCKR* and *HNF1A* genes. While not reported in prior hypercholesterolemia GWAS, they are captured in GWAS of lipid traits and metabolic regulation.

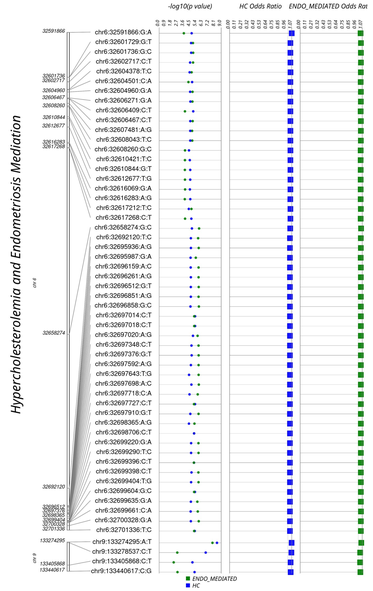
GCKR regulates glucose metabolism levels in liver and pancreatic islet cells and is implicated in previous GWAS for insulin and lipid levels. HNF1A encodes a liver-specific transcription factor, with gene defects associated with maturity-onset diabetes of the young.

Furthermore, 5 of the genome-wide significant loci are unreported in prior GWAS of lipid panel measurements and hypercholesterolemia: *LINC01899, LOC102546299, DENND1A, LOC105372096, and RNA5SP189/LOC345571*.

Notably, DENND1A is associated with polycystic ovary syndrome and shows evidence of regulatory variation (eQTLs) in ovarian tissue, while *LOC102546299* is linked to female coronary artery disease survival.

LINC01899 and *RNA5SP189/LOC345571* are associated with smoking initiation, a risk factor for elevated cholesterol.

Results: Mediated Genetic Effects on Hypercholesterolemia



Comparing Direct vs Indirect Effects

In the synthesis view plot on the left, we explore how the genetic effects of SNPs on hypercholesterolemia are explained by mediation.

We compare the mediated GWAS results to a hypercholesterolemia GWAS (Verma et al., 2024). All GWAS are meta-analyzed.

Proportion Mediation

We calculate the percentage mediation for SNPs associated independently with hypercholesterolemia and endometriosis.

We observe partial mediation (with an average of 50%) of endometriosis on hypercholesterolemia for 56 SNPs. Signals include *ADAMTS13* (9:133440617), which is a significant regulator of blood clotting, as well as *HLA* (6: 32658274).

Key Takeaways: Direct versus Indirect Genetic Effects

- In this work, we developed and applied an innovative mediated GWAS methodology that distinguishes between direct genetic effects on hypercholesterolemia, and indirect effects mediated through endometriosis.
- These findings collectively broaden our understanding of the genetic architecture of hypercholesterolemia, highlight potential sex-specific regulatory mechanisms, and point to previously unreported loci for future functional investigation.
- Our future work includes mediated GWAS of endometriosis and additional cardiometabolic indicators, given their posited interconnected biological mechanisms.

Acknowledgements

- Research reported in this poster was supported by the Eunice Kennedy Shriver National Institute of Child Health and Human Development of the National Institutes of Health under award number R01HD110567-01A0.
- We gratefully acknowledge All of Us, eMERGE, and Million Veteran Program participants for their contributions, without whom this research would not have been possible. We also thank the NIH's All of Us Research Program for making available the participant data examined in this study.
- We acknowledge the Penn Medicine BioBank (PMBB) for providing data and thank the patient-participants of Penn Medicine who consented to participate in this research program. We would also like to thank the Penn Medicine BioBank team for providing genetic variant data for analysis. The PMBB is approved under IRB protocol #813913 and supported by Perelman School of Medicine at University of Pennsylvania, a gift from the Smilow family, and the National Center for Advancing Translational Sciences of the National Institutes of Health under CTSA award number UL1TR001878.